Optimal Learning-Augmented Algorithm for Online Bidding

Changyeol Lee[*]
Dept. of Computer Science and Engineering
Yonsei University
Seoul, South Korea
777john@yonsei.ac.kr

Dahoon Lee[*†]
Dept. of Computer Science
New York University
New York, NY, USA
dahoon.lee@nyu.edu

Jongseo Lee[*]
School of Computing
KAIST
Daejeon, South Korea
leejseo@kaist.ac.kr

Yongho Shin[*]
Institute of Computer Science
University of Wrocław
Wrocław, Poland
yongho@cs.uni.wroc.pl

Changki Yun[*†]
Dept. of Computer Science
University of Maryland
College Park, MD, USA
ckyun@umd.edu

Abstract

Recent advances in machine learning have spurred significant interest in learning-augmented algorithms, particularly for online optimization. A growing body of work has studied *online bidding* in this framework, aiming to characterize the trade-off between robustness and consistency. While this trade-off is fully understood for deterministic algorithms, a gap between upper and lower bounds remains in the randomized setting. In this paper, we close this gap by presenting a Pareto-optimal randomized learning-augmented algorithm for this problem. Our approach introduces the notion of a *bidding profile*, an alternative representation of the probability distribution over bids generated by an algorithm in terms of a non-decreasing function. We show that any bidding algorithm can be reduced, without loss of generality, to one driven by a bidding profile, and we characterize the optimal profile via a system of delayed differential equations. Finally, we demonstrate the broader applicability of our approach by extending it to the *linear search* problem, yielding a significant improvement over prior learning-augmented algorithms for linear search.

1 Introduction

Recent advances in machine learning (ML) have spurred active research on *learning-augmented algorithms* [MV22]. These algorithms take predictions as additional inputs and leverage them to solve computational problems. A key advantage of this paradigm is that it exploits the empirically high quality of ML predictions to achieve strong performance, while still providing robust guarantees regardless of prediction accuracy. Another distinguishing feature is that the prediction model is treated as a complete black box, making the framework independent of any specific prediction mechanism.

Research on learning-augmented algorithms has been particularly active in the context of *online optimization*, where predictions can naturally serve as surrogates for unknown future inputs. Since the seminal work of Lykouris and Vassilvitskii [LV21], a substantial body of work, including [PSK18, BMS20, APT22, EADL24, SLLA25, CJS25], has studied a variety of online optimization problems in the learning-augmented setting, demonstrating the effectiveness of this paradigm.

[*]Equal contribution
[†]Work done while was at the Department of Mathematical Sciences, Seoul National University, South Korea.

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

The *online bidding* problem is a canonical online optimization problem in which an algorithm sequentially submits bids until reaching an unknown target value $T > 0$, with the objective of minimizing the total sum of submitted bids. This problem succinctly captures the core challenge of searching for an unknown target in online optimization. For example, it models the design of *anytime algorithms* or *interruptible algorithms*—which arise frequently in intelligent systems and real-time applications [Zil96, KWP$^+$11]—through the scheduling of *contract algorithms* [RZ91, LOAH14].[3] In this setting, one schedules a sequence of contract algorithm executions with prescribed running times so as to achieve the best possible performance guarantee under an unknown interruption time T. The resulting trade-off between solution quality and execution time is precisely captured by online bidding.

The standard performance measure for online bidding algorithms is the *competitive ratio*, defined as the worst-case ratio between the (expected) total sum of submitted bids and the optimal bid with hindsight (i.e., the target value T). A classical folklore algorithm for this problem is the *doubling method*, which doubles the bid in each round until termination. Chrobak, Kenyon, Noga, and Young [CKNY08] showed that this algorithm is 4-competitive and that no deterministic algorithm can achieve a better competitive ratio. For randomized algorithms, one can randomly shift the initial bid to have e^U with $U \sim \mathrm{Unif}(0, 1]$, and then increase the bid by a factor of e in every round. This yields an e-competitive algorithm, which was also shown to be optimal among all randomized algorithms [CKNY08].

Following the growing interest in learning-augmented algorithms, online bidding has been extensively studied in this framework, particularly in the setting where a single value prediction P is provided as an estimate of the unknown target value T [AGKP21, IMXZ23, ADJ$^+$24, SLLA25, AS25]. This setting arises naturally in various applications; for instance, in the above example of designing anytime algorithms, a prediction of the interruption time could be obtained from an ML model trained on historical execution data. The availability of such a prediction motivates the study of learning-augmented algorithms that use it to guide their scheduling decisions while retaining desired performance guarantees. A central focus of this line of research is the trade-off between *robustness* and *consistency* [LV21, PSK18]. Informally, robustness measures the competitive ratio guaranteed regardless of the prediction quality, whereas consistency captures the performance when the prediction is accurate, i.e., when $P = T$.

One might expect that a Pareto-optimal learning-augmented algorithm can be obtained by adapting an optimal competitive algorithm without learning augmentation. This is indeed the case for deterministic online bidding: Angelopoulos, Dürr, Jin, Kamali, and Renault [ADJ$^+$24] and Shin, Lee, Lee, and An [SLLA25][4] independently showed that a simple modification of the doubling method yields a deterministic Pareto-optimal learning-augmented algorithm. In particular, the algorithm achieves strong consistency by deterministically including the prediction P into its sequence of geometrically increasing bids.

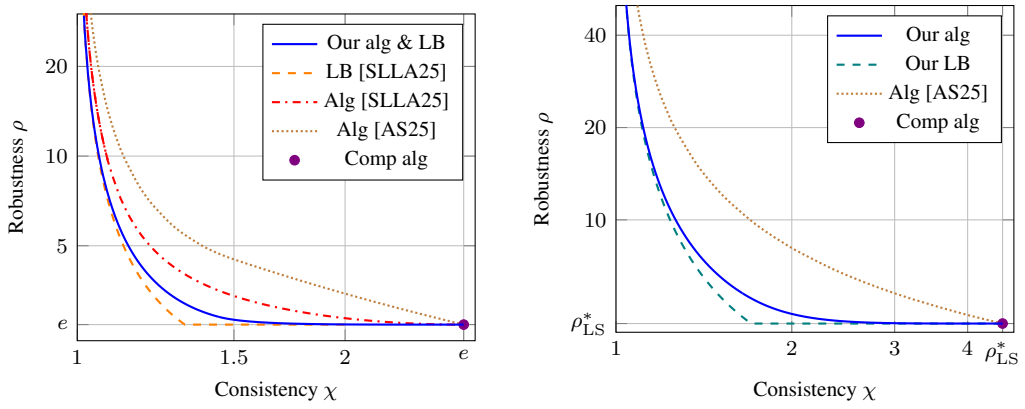
In contrast, a straightforward extension is less apparent for the randomized e-competitive algorithm, since the random shift in the bid sequence makes it difficult to deterministically incorporate the prediction P. To address this issue, Shin et al. [SLLA25] introduced a "prediction-following" phase, in which the algorithm explicitly submits the bid P during the execution of the randomized algorithm, yielding the currently best known trade-off. Alternatively, Angelopoulos and Simon [AS25] carefully interpolated between the optimal randomized competitive algorithm and the optimal deterministic learning-augmented algorithm through controlled randomization. Despite these advances, a gap remains between the best-known upper and lower bounds on the robustness-consistency trade-off [SLLA25], leaving the characterization of the Pareto-optimal trade-off as an open question. See Figure 1a for a comparison of the known trade-offs.

1.1 Our contributions

In this paper, we resolve this open question by presenting a randomized learning-augmented algorithm that achieves the Pareto-optimal robustness-consistency trade-off (see Figure 1a).

[3] A contract algorithm receives its execution time as part of the input and provides a performance guarantee only if allowed to run for the prescribed duration, whereas an anytime (or interruptible) algorithm can be interrupted at any time and return its best current solution.

[4] Shin et al. [SLLA25] studied multi-option ski rental, but their techniques extend naturally to online bidding. See Appendix B for the equivalence between the two problems.



(a) Online bidding

(b) Linear search

Figure 1: Robustness-consistency trade-offs of **(a) online bidding and (b) linear search**. All axes use logarithmic scaling. **(a)** The solid blue curve depicts our Pareto-optimal trade-off (Theorem 1), while the dash-dotted red and dashed orange curves represent the upper and lower bounds from [SLLA25], respectively. The dotted brown curve shows the trade-off achieved by [AS25]. The violet point corresponds to the classical e-competitive algorithm without learning augmentation. **(b)** The solid blue curve depicts the trade-off achieved by our algorithm (Theorem 17), while the dashed teal curve represents our lower bound (Theorem 25). The dotted brown curve shows the upper bound from [AS25]. The violet point corresponds to the classical linear search algorithm with competitive ratio $\rho^*_{\mathrm{LS}} := 1 + 1/W_0(e^{-1}) \approx 4.59$.

Theorem 1. *For any trade-off parameter $s \in (0, 1]$, there exists a randomized $\rho(s)$-robust $\chi(s)$-consistent online bidding algorithm with*

$$\rho(s) := \frac{e^s}{s} \quad \textit{and} \quad \chi(s) := \begin{cases} (e^s - 1)/s, & 0 < s < \ln 2, \\ \xi(s)/s, & \ln 2 \le s \le 1, \end{cases}$$

where $\xi(s)$ is the unique solution $\xi \in [1, e]$ of $\xi(2 - \ln \xi) = e^s$ for $s \in [\ln 2, 1]$. Moreover, the robustness-consistency trade-off of this algorithm is Pareto-optimal.

To achieve this goal, we introduce the notion of a *bidding profile*, an alternative representation of the probability distribution over the bids generated by an algorithm in terms of a non-decreasing function. Intuitively, a bidding profile is specified by a non-decreasing function $G : \mathbb{R} \to (0, \infty)$, from which the bid sequence is generated through a random shift. To gain intuition for the role of a bidding profile, consider the classical randomized e-competitive algorithm. Recall that the algorithm chooses the initial bid as e^U with $U \sim \mathrm{Unif}(0, 1]$, and subsequently multiplies the bid by a factor of e in every round until termination. This execution can be naturally associated with the exponential function $G(t) := e^t$, in the sense that the bids generated by the algorithm form a subsequence of the bi-infinite random sequence $\{G(n + U)\}_{n \in \mathbb{Z}}$, where $U \sim \mathrm{Unif}(0, 1]$.

Motivated by this perspective, we study the family of online bidding algorithms induced by bidding profiles G and identify the profile whose corresponding algorithm achieves the optimal robustness-consistency trade-off within this family via a tailored system of delayed differential equations. We then show that any online bidding algorithm can be transformed into one induced by a bidding profile without loss in performance guarantees, thereby establishing the Pareto-optimality of our bidding-profile-based algorithm.

We further demonstrate the versatility of our profile-based approach by extending it to the *linear search* problem, yielding a significant improvement in the robustness–consistency trade-off for randomized learning-augmented linear search algorithms over the previous state of the art [AS25]. However, our lower-bound argument for online bidding does not extend naturally to the linear search setting, and thus establishing the Pareto-optimality of the resulting algorithm remains an open problem. Instead, we derive a nearly tight lower bound using Yao's principle, building on ideas from prior work [AS25, SLLA25]. See Figure 1b for a comparison of our upper and lower bounds.

During the preparation of this paper, we became aware of concurrent and independent work by Degryse, Zaakour, Dürr, and Angelopoulos [DZDA26] on learning-augmented online bidding. They also establish the Pareto-optimal robustness-consistency trade-off for robustness ratios at least 2.885, covering almost the entire range of robustness. Their bidding function is essentially equivalent to our bidding profile, and both works combine this representation with a random shift to generate bid sequences. However, they derive their lower bounds via an LP-based approach, whereas we prove optimality within the class of profile-driven algorithms and show that restricting attention to this class incurs no loss in robustness or consistency. In addition, they extend their results to the incremental median problem, whereas we extend our profile-based approach to learning-augmented linear search and obtain an improved algorithm.

Further related work is discussed in Appendix A.

2 Preliminaries

This section formally defines the problems considered in this paper, along with the performance measures and the notation used throughout this paper.

Notation For a real function f and a real number x, we denote $f(x+) = \lim_{y \downarrow x} f(y)$, and $f(x-) = \lim_{y \uparrow x} f(y)$. For an event $\mathcal{E}$, we denote the indicator function with $\mathbf{1}_{\mathcal{E}}$. For a function G and a subset of its domain I, we denote the restriction of G over I as $G|_I$. We denote by W_0 the principal branch of the Lambert W function.

Problem definitions In the *online bidding* problem, there is an unknown target value $T > 0$. A bidder repeatedly submits non-decreasing[5] bids $b \in \mathbb{R}_{>0}$ until a bid $b \geq T$ is submitted. Prior to termination, the bidder receives no information about T other than the fact that all previous bids were smaller than T. The cost is defined as the sum of all submitted bids up to and including the first bid of value at least T. Naturally, the objective is to minimize the cost.

In the *linear search* problem, there is an unknown target point $T \in \mathbb{R} \setminus \{0\}$, and the searcher starts at the origin. One can view the search as a sequence of signed *excursion endpoints*. In a single excursion to a signed endpoint $z \in \mathbb{R} \setminus \{0\}$, the searcher starts at the origin and continuously moves toward z. If the searcher reaches the target on the way, then the search immediately stops there; otherwise, the searcher returns to the origin, ending this excursion. The cost is defined as the total distance traveled until termination. It is also without loss of generality to assume that consecutive excursions have opposite directions (i.e., signs) and that the magnitudes of excursion endpoints are non-decreasing in each direction.

In the *learning-augmented* setting, the instance additionally contains a prediction for the unknown target. For online bidding, the prediction is a value $P > 0$ on the target value T; for linear search, the prediction is a point $P \in \mathbb{R} \setminus \{0\}$ on the target point T. Note that the prediction can be inaccurate.

From this point, we focus on online bidding due to space constraints. The analogous discussion for linear search is deferred to Appendix E.1.

Algorithms and strategies Note that an algorithm for online bidding can be represented by an infinite sequence of non-decreasing real numbers. However, since the algorithm should be competitive against every possible $T > 0$ which can be any positive real number, it is difficult to define the *first* bid submitted by the algorithm. To address this issue, one standard assumption is to bound T from below by some constant (say, 1) [ADJ$^+$24, AS25, SLLA25].

In this paper, however, we introduce a surrogate model for an algorithm, called a *strategy*, defined as a bi-infinite sequence $b = \{b_n\}_{n \in \mathbb{Z}}$ of non-decreasing bids $0 < \cdots \leq b_{-1} \leq b_0 \leq b_1 \leq \cdots$. The cost incurred by a deterministic strategy b given a target value $T > 0$, denoted by $\mathsf{cost}_b(T)$, is then defined as the prefix sum of bids up to and including the first bid at least T, i.e.,

$$\mathsf{cost}_b(T) := \textstyle\sum_{n \leq n_*} b_n, \text{ where } n_* := \min\{n \in \mathbb{Z} : b_n \geq T\}.$$

A randomized strategy B is naturally defined as a probability distribution over deterministic strategies. The expected cost incurred by B given T is denoted by $\mathsf{cost}_B(T)$.

[5] It is without loss of generality to assume that bids are strictly increasing, but we allow them to be non-decreasing. This helps simplify the presentation of the analysis below.

Normalization As the universe of T is $(0, \infty)$, the problem is homogeneous under scaling. We therefore assume from now that the prediction is given as 1 (by dividing the whole universe by P). Moreover, as a strategy is bi-infinite, we can also freely shift the indices of the strategy. Hence, we assign index 0 to the first bid of value at least $1 = P$.

Performance measures To measure the performance of learning-augmented algorithms, we adopt the *robustness-consistency analysis* [LV21, MV22]. We say that a (randomized) algorithm is *ρ-robust* if, for every possible target T, its expected cost is at most ρT regardless of prediction accuracy, and *χ-consistent* if its expected cost is at most χ when the prediction is correct, i.e., $T = 1$. Similarly, we say that a (randomized) strategy B is *ρ-robust* if $\mathsf{cost}_B(T) \le \rho T$ for every $T > 0$ and *χ-consistent* if $\mathsf{cost}_B(1) \le \chi$. A pair (ρ, χ) is *Pareto-optimal* for algorithms and strategies, respectively, if there exists a ρ-robust χ-consistent algorithm and strategy, respectively, but no ρ'-robust χ-consistent algorithm and strategy, respectively, for any $\rho' < \rho$.

In what follows, we focus on the performance guarantees of strategies. The main reason is that it provides substantial simplicity in our analysis. In fact, even though a strategy cannot serve as an algorithm by itself due to its bi-infinite nature, one can easily generate from a strategy an algorithm which incurs no more than the strategy (in expectation) in the setting where the target value is bounded from below by some constant. Moreover, we can also argue that any lower bound on the robustness-consistency trade-off for strategies also serves as a valid lower bound for algorithms since a strategy can be viewed as an algorithm given a sufficiently large prediction.

Lemma 2. *For online bidding, a strategy induces an algorithm whose (expected) cost is no greater than that of the strategy. Conversely, any lower bound on the robustness-consistency trade-off for strategies for online bidding also applies to algorithms.*

A detailed discussion on this "equivalence" is provided in Appendix C.

3 Pareto-optimal trade-off for online bidding

In this section, we present a Pareto-optimal learning-augmented strategy for online bidding. To this end, we introduce in Section 3.1 an alternative representation called a *bidding profile* for representing the distribution over bids made by a strategy. In Section 3.2, we construct a bidding profile such that the strategy driven by the profile attains the best-possible trade-off among all profile-driven strategies. Finally in Section 3.3, we prove that any bidding strategy can be reduced to a bidding profile, implying the Pareto-optimality of the strategy driven by our bidding profile.

3.1 Bidding profiles

To motivate our alternative representation, we revisit the well-known e-competitive algorithm for online bidding [CKNY08]. Under the assumption that $T \ge 1$, it samples $U \sim \mathrm{Unif}(0, 1]$ at the beginning, submits the first bid e^U, and repeatedly increases the bid by a factor e until termination. Note that we can represent the random execution of the strategy corresponding to this algorithm by a single non-decreasing function $G(x) := e^x$ as follows: For a shared random source $U \sim \mathrm{Unif}(0, 1]$, the strategy is defined as $\{G(n + U)\}_{n \in \mathbb{Z}}$.

Let us now analyze the competitive ratio of this strategy, with the assumption lifted. For the target $T > 0$, we define $x := G^{-1}(T) = \ln T$ and write $x = q + r$ for some (uniquely determined) $q \in \mathbb{Z}$ and $r \in (0, 1]$. Note that the last bid submitted by the strategy is $G(q+U)$ if $U \ge r$, and $G(q+1+U)$ if $U < r$. Hence, the expected sum of total bids submitted by the strategy is computed as follows:

$$\begin{aligned}\int_0^r \sum_{n \le q+1} G(n+u)\,\mathrm{d}u + \int_r^1 \sum_{n \le q} G(n+u)\,\mathrm{d}u &= \int_0^1 \sum_{n \le q} G(n+u)\,\mathrm{d}u + \int_0^r G(q+1+u)\,\mathrm{d}u \\ &= \int_{-\infty}^{q+1+r} G(t)\,\mathrm{d}t = \int_{-\infty}^{x+1} G(t)\,\mathrm{d}t.\end{aligned}$$

Observe that, for any $T > 0$ with $x = G^{-1}(T)$, we have

$$\int_{-\infty}^{x+1} G(t)\,\mathrm{d}t = e\,G(x) = e\,T \tag{1}$$

by the choice of function G, showing the e-competitiveness of the strategy.

In what follows, we generalize G in the above strategy to an arbitrary non-decreasing positive function, and focus on the family of bidding strategies driven by such functions G. More precisely, given a non-decreasing function G, the strategy is defined as $\{G(n+U)\}_{n\in\mathbb{Z}}$ for a shared random source $U \sim \mathrm{Unif}(0,1]$. We highlight that this strategic representation already captures many existing algorithms with or without learning augmentation [CKNY08, AGKP21, ADJ$^+$24, SLLA25, AS25]. For a simple example, notice that the strategy corresponding to the classical deterministic 4-competitive algorithm can be represented by $G(x) := 2^{\lceil x\rceil - 1}$.

In the learning-augmented setting, we are given a prediction P on T. Recall we assume $P = 1$. Moreover, we horizontally shift G so that $G|_{(0,1]}$ represents the strategy's bid that becomes at least the prediction for the first time. The robustness and consistency of the strategy driven by G can be expressed analogously to (1), leading us to the following definition of a *bidding profile*:

Definition 3 (Bidding profile)**.** Given ρ and χ such that $1 < \chi \le \rho$, a non-decreasing, left-continuous function $G : \mathbb{R} \to (0,\infty)$ is called a (ρ,χ)*-bidding profile* if

- (Offset condition) $G(x) < 1$ for every $x < 0$, and $G(x) \ge 1$ for every $x > 0$;
- (Robustness condition) $\int_{-\infty}^{x+1} G(t)\mathrm{d}t \le \rho\, G(x)$ for every $x \in \mathbb{R}$;
- (Consistency condition) $\int_{-\infty}^{1} G(t)\mathrm{d}t \le \chi$.

Remark 4. Note that $G(0) \le 1$, from its left-continuity and the offset condition. Also, the strict positivity of G can be deduced from the robustness condition.

The proof of Remark 4 is deferred to Appendix D.1.

Given a bidding profile G, the strategy driven by G is expressed by $\{G(n+U)\}_{n\in\mathbb{Z}}$, where $U \sim \mathrm{Unif}(0,1]$ is a random source shared across $n \in \mathbb{Z}$.

As G is non-decreasing, this bidding sequence is indeed well defined. Denoting by $\mathsf{cost}_G(T)$ the expected cost incurred by the strategy driven by G given T, the next lemma provides a formula for $\mathsf{cost}_G(T)$, implying the robustness and consistency of this strategy.

Lemma 5. *Let G denote a (ρ,χ)-bidding profile. Then, for any $T > 0$,*

$$\tau(T) := \sup\{t \in \mathbb{R} : G(t) < T\} < \infty, \quad \textit{and} \quad \mathsf{cost}_G(T) = \textstyle\int_{-\infty}^{\tau(T)+1} G(t)\mathrm{d}t.$$

Following is a direct corollary from the above lemma and the definition of a bidding profile.

Corollary 6. *The strategy driven by a (ρ,χ)-bidding profile is ρ-robust and χ-consistent.*

3.2 Pareto-optimal bidding profile

We now construct a bidding profile which gives the Pareto-optimal trade-off among profile-driven strategies. We present the following theorem, implying the first half of Theorem 1 together with Corollary 6.

Theorem 7. *For any $s \in (0,1]$, there exists a $(\rho(s),\chi(s))$-bidding profile, where*

$$\rho(s) := \frac{e^s}{s} \quad \textit{and} \quad \chi(s) := \begin{cases} (e^s-1)/s & (0 < s \le \ln 2), \\ \xi(s)/s & (\ln 2 \le s \le 1), \end{cases}$$

where $\xi(s)$ is the unique solution $\xi \in [1,e]$ satisfying $\xi(2-\ln\xi) = e^s$ for any $s \in [\ln 2, 1]$.

In the interest of space, we sketch the proof here, while deferring the full proof to Appendix D.2.

Proof sketch. For $s = 1$, it is trivial to see that the bidding profile $G(x) = e^x$ induces the classical e-competitive algorithm. We thus fix any $s \in (0,1)$, and write $\rho := \rho(s)$ and $\chi := \chi(s)$ for simplicity. We first define a bidding profile G for $x > 0$ as follows:

$$G(x) := \begin{cases} \max(1, s\chi \cdot e^{s(x-1)}), & \text{for } 0 < x \le 1, \\ \max(1, s\chi) \cdot e^{s(x-1)}, & \text{for } x > 1. \end{cases} \tag{2}$$

It remains to define G on $x \leq 0$, which is defined inductively as follows. Let $G(0) := \frac{\chi}{\rho} \leq 1$. For any non-positive integer $k \in \mathbb{Z}_{\leq 0}$, given G defined on $[k, k+1)$, we define G on $[k-1, k)$ as

$$G(x) := G(k) - \frac{1}{\rho} \int_{x+1}^{k+1} G(t)\mathrm{d}t \quad \text{for any } x \in [k-1, k). \tag{3}$$

Intuitively, the above expression computes, given the distribution of bid $G(k+U)$, the optimal distribution of bid $G(k-1+U)$ satisfying the desired robustness and consistency. See Figure 2 for illustrations of the constructed profile G for several representative values of the parameter s.

The key technical part is to show that G is non-negative. For the moment, let us assume that this statement is true. We can then show that G is indeed a (ρ, χ)-bidding profile. Observe that G is non-decreasing due to (2) and (3) along with the non-negativity assumption and $G(0) \leq 1$, where the offset condition also follows. The left-continuity of G is guaranteed by the fact that G is continuous on $x > 0$ and $x \leq 0$ again due to (2) and (3), respectively. For the remaining two conditions, let us first argue the robustness condition for $x < 0$. For any non-positive integer $k \in \mathbb{Z}_{\leq 0}$ and $x \in [k-1, k)$, we derive from (3)

$$G(x) = G(x) - G(k) + \textstyle\sum_{n \in \mathbb{Z}_{\leq k}} (G(n) - G(n-1)) = \frac{1}{\rho} \int_{-\infty}^{x+1} G(t)\mathrm{d}t,$$

which shows the robustness condition for $x < 0$. Due to the left-continuity, the equation also holds for $x = 0$, implying the consistency condition together with the choice of $G(0) = \frac{\chi}{\rho}$. The robustness condition for $x > 0$ can be shown analogously to the standard analysis of the classical e-competitive algorithm.

To show the non-negativity of G, we define the following functional operator.

Definition 8. Given $\rho > 0$, let $\mathcal{F}$ be a functional operator on an integrable, non-decreasing, non-negative function $H : (-\infty, 1] \to [0, \infty)$, giving a function $\mathcal{F}H : (-\infty, 1] \to [0, \infty)$ such that

$$\mathcal{F}H(x) := \begin{cases} H(x), & 0 < x \leq 1, \\ \frac{1}{\rho} \int_{-\infty}^{x+1} H(t)\mathrm{d}t, & x \leq 0. \end{cases}$$

We claim that our bidding function G satisfies $G|_{(-\infty,0]} \equiv G^*|_{(-\infty,0]}$, where $G^* : (-\infty, 1] \to [0, \infty)$ is a function satisfying (i) $G^* = \mathcal{F}G^*$, (ii) $\int_{-\infty}^{1} G^*(t)\mathrm{d}t = \chi$, and (iii) $G^*|_{(0,1]} \equiv \phi|_{(0,1]}$ where $\phi(x) := \max(1, s\chi \cdot e^{s(x-1)})$. First, we can derive $G(0) = G^*(0) = \frac{\chi}{\rho}$ from (i) and (ii). Next, for any non-positive integer $k \in \mathbb{Z}_{\leq 0}$ and $x \in [k-1, k)$,

$$\begin{aligned} G^*(x) &= \tfrac{1}{\rho} \textstyle\int_{-\infty}^{x+1} G^*(t)\mathrm{d}t = \tfrac{1}{\rho} \left(\int_{-\infty}^{k+1} G^*(t)\mathrm{d}t - \int_{x+1}^{k+1} G^*(t)\mathrm{d}t \right) \\ &= G^*(k) - \tfrac{1}{\rho} \textstyle\int_{x+1}^{k+1} G^*(t)\mathrm{d}t = G(x), \end{aligned}$$

where the first and third equalities are from (i), and the last equality is by the inductive hypothesis. This completes the proof of the claim if G^* exists.

It thus remains to show the existence of G^*. We exploit the *order-preserving* property of $\mathcal{F}$: if $A \leq B$ for two functions A, B (i.e., $A(x) \leq B(x)$ for all $x \in (-\infty, 1]$), then $\mathcal{F}A \leq \mathcal{F}B$. Let $G^{(0)}$ be a function on $(-\infty, 1]$ such that $G^{(0)}(x) := 0$ for $x \leq 0$ and $G^{(0)}(x) := \phi(x)$ for $0 < x \leq 1$, and for $n \in \mathbb{Z}_{\geq 0}$, let $G^{(n+1)}(x) := \mathcal{F}G^{(n)}(x)$.

Since $G^{(0)} \leq G^{(1)}$, the order-preserving property implies $G^{(n)} \leq G^{(n+1)}$ for all $n \in \mathbb{Z}_{\geq 0}$. We then show that, for all $n \in \mathbb{Z}_{\geq 0}$, $G^{(n)}$ is bounded from above by H defined as $H(x) := e^{cx}$ for $x \leq 1$ for some $c \in (s, 1)$. Indeed, $\mathcal{F}H \leq H$, which implies $G^{(n)} \leq \mathcal{F}^{(n)}G^{(0)} \leq \mathcal{F}^{(n)}H \leq H$, where the second inequality follows from the definition of $G^{(0)}$ and H. By the monotone convergence theorem, $\{G^{(n)}\}_{n \in \mathbb{Z}_{\geq 0}}$ converges. Simple calculation shows that the converged function satisfies the remaining conditions for G^*, completing the proof. □

We now argue that the trade-off of Theorem 7 is indeed the best possible among profile-driven strategies. Let us say a bidding profile G is *induced by* $\phi : (0, 1] \to [1, \infty)$ if $G|_{(0,1]} \equiv \phi$. We further define a *tight* bidding profile.

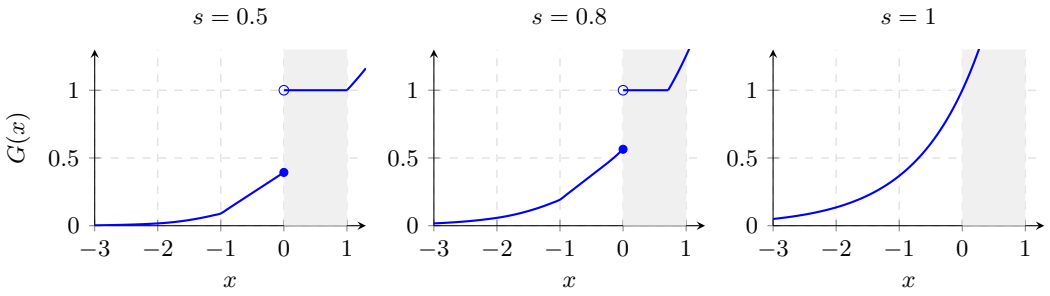


Figure 2: Representative bidding profiles G for three regimes of s which determines the trade-off. The shaded strip highlights the interval $(0, 1]$.

Definition 9 (Tight bidding profile)**.** We say a (ρ, χ)-bidding profile G is *tight* if it further satisfies

$$\int_{-\infty}^{x+1} G(t)\mathrm{d}t = \rho\, G(x) \quad \text{for all } x \le 0, \quad \text{and} \quad \int_{-\infty}^{1} G(t)\mathrm{d}t = \chi.$$

In other words, we call a bidding profile *tight* if both the robustness condition for $x \le 0$ and the consistency condition hold with equality. We then show that it suffices to consider tight bidding profiles.

Lemma 10. *Suppose there exists a (ρ, χ)-bidding profile G for fixed positive numbers $\rho > \chi > 1$. Let $\phi : (0, 1] \to [1, \infty)$ denote the function that induces G. Then, there exists a tight (ρ, χ')-bidding profile G' induced by ϕ such that $\chi' \le \chi$.*

To prove the above lemma, recall the functional operator $\mathcal{F}$ defined in the proof sketch of Theorem 7. (See Definition 8). Intuitively, $\mathcal{F}$ uses the slack in the robustness condition to "lower" the bidding profile, thereby improving consistency while preserving robustness. We show that iteratively applying this operator yields a sequence of bidding profiles converging to a tight one.

Finally, a calculation shows that if a tight bidding profile has robustness $\rho(s)$, then its consistency must be at least $\chi(s)$. This establishes the optimality of our bidding profile.

Theorem 11. *For any $0 < s \le 1$, if there exists a tight $(\rho(s), \chi)$-bidding profile, then $\chi \ge \chi(s)$.*

Again, all proofs are deferred to Appendix D.2 due to the space constraint.

3.3 Generality of bidding profiles

To complete the proof of Theorem 1, we show the following lemma.

Lemma 12. *For any fixed positive real numbers ρ and χ such that $\rho > \chi > 1$, there exists a ρ-robust χ-consistent online bidding strategy if and only if there exists a (ρ, χ)-bidding profile G.*

Note that the second half of Theorem 1 then immediately follows from Theorem 11. Due to the space constraint, we only provide a sketch of its proof, while deferring the full proof to Appendix D.3.

Proof sketch. Recall that a bidding strategy is a probability distribution over bi-infinite sequences in $(0, \infty)$, which we may assume to be strictly increasing without loss of generality. Hence, given a ρ-robust χ-consistent bidding strategy $\mathcal{A}$ with prediction $P = 1$, by simply accumulating the probability measure value-wise, we can obtain an *aggregate measure* μ; informally speaking, for each value $t \in (0, \infty)$, $\mu(t)$ corresponds to the marginal probability that $\mathcal{A}$ ever submits t as its bid value when $\mathcal{A}$ never terminates.

We then define the *inverse profile* $G : \mathbb{R} \to (0, \infty)$ from μ as follows:

$$G(x) := \begin{cases} \sup\{0 < t < 1 : \ \mu([t, 1)) \ge -x\}, & \text{if } x < 0; \\ \inf\{t \ge 1 : \ \mu([1, t]) \ge x\}, & \text{if } x \ge 0. \end{cases}$$

Intuitively, G can be seen as a bidding profile whose driven strategy $\mathcal{B}$ induces the same aggregate measure μ as $\mathcal{A}$.

To show that G is actually a (ρ, χ)-bidding profile, we argue that the expected cost incurred by $\mathcal{B}$ is always no worse than that by $\mathcal{A}$ for every target value $T > 0$. On the one hand, the expected costs incurred before submitting at least T are the same between $\mathcal{A}$ and $\mathcal{B}$ since they have the same aggregate measure. On the other hand, the distribution of the last bid (which is at least T) can be viewed as a packing of size 1 in $\mu([T, \infty))$, and by the definition of G, $\mathcal{B}$ packs this distribution as small as possible with respect to $\mu([T, \infty))$. These together imply that $\mathcal{B}$ incurs no worse than $\mathcal{A}$ in expectation for every $T > 0$. □

4 Extension to linear search

In this section, we present a learning-augmented strategy for the linear search problem. We introduce in Section 4.1 an alternative representation called an *excursion profile*, adapted from a bidding profile to describe a sequence of excursions. In Section 4.2, we present a near-optimal excursion profile whose driven strategy shows the robustness-consistency trade-off presented in Figure 1b.

4.1 Excursion profiles

Since linear search is also scale-invariant, we assume that the prediction is located at coordinate $+1$. Instead of a single profile, we assign a pair $G_\pm = (G_+, G_-)$ of non-decreasing, non-negative functions as an *excursion profile*. The strategy driven by $G_\pm$ naturally corresponds to the following excursion sequence with a common random source $U \sim \mathrm{Unif}(0, 1]$:

$$\cdots + G_+(n + U) \to -G_-(n + U) \to +G_+(n + 1 + U) \to -G_-(n + 1 + U), \cdots$$

Moreover, we define $G_+|_{(0,1]}$ as the first excursion passing through the prediction point. This leads us to the following definition.

Definition 13 (Excursion profile)**.** For ρ and χ with $\rho \geq \chi > 0$, a pair $G_\pm$ of non-decreasing, left-continuous functions $G_+, G_- : \mathbb{R} \to (0, \infty)$ is called a (ρ, χ)-*excursion profile* if

- (Plus-offset condition) $G_+(x) \geq 1$ for any $x > 0$, and $G_+(x) < 1$ for any $x < 0$;
- (Robustness condition) $C_+(x) \leq \rho\, G_+(x)$ and $C_-(x) \leq \rho\, G_-(x)$ for all $x \in \mathbb{R}$;
- (Consistency condition) $C_+(0) \leq \chi$,

where $C_+(x) := \int_{-\infty}^{x} G_+(t)\mathrm{d}t + \int_{-\infty}^{x} G_-(t)\mathrm{d}t$ and $C_-(x) := \int_{-\infty}^{x+1} G_+(t)\mathrm{d}t + \int_{-\infty}^{x} G_-(t)\mathrm{d}t$.

We can easily adapt the proof of Lemma 5 to derive the following lemma whose proof is deferred to Section E.2.

Lemma 14. *The strategy driven by a (ρ, χ)-excursion profile $G_\pm$ is $(1 + 2\rho)$-robust and $(1 + 2\chi)$-consistent.*

4.2 Near-optimal excursion profile

In this section, we construct a near-optimal excursion profile, using the methods derived from Section 3.2. The proofs can be found in Appendix E.3.

Let us say an excursion profile $G_\pm$ is *induced by* $\psi : (0, 1] \to [1, \infty)$ if $G_+|_{(0,1]} \equiv \psi$. We will consider *tight* excursion profiles induced by ψ without degrading performance.

Definition 15 (Tight excursion profile)**.** We say a (ρ, χ)-excursion profile $G_\pm$ is *tight* if it further satisfies $C_+(x) = \rho\, G_+(x)$ and $C_-(x) = \rho\, G_-(x)$ for all $x \leq 0$, and $C_+(0) = \chi$.

The next lemma essentially shows that one can construct a tight excursion profile that is pointwise minimal on $(-\infty, 0]$ among all excursion profiles induced by ψ with the same robustness, whenever such profiles exist, and its consistency is solely determined by a weighted integral of ψ.

Lemma 16. *Fix $\rho > 1/(2W_0(e^{-1}))$. Given a bounded and integrable function $\psi : (0, 1] \to [1, \infty)$, there exist functions $G_+ : (-\infty, 1] \to [0, \infty)$ and $G_- : (-\infty, 0] \to [0, \infty)$ such that*

(i) $G_+ \equiv \psi$ on $(0, 1]$.

(ii) For all $x \le 0$, we have

$$\int_{-\infty}^{x} G_+(t)\mathrm{d}t + \int_{-\infty}^{x} G_-(t)\mathrm{d}t = \rho G_+(x),$$
$$\int_{-\infty}^{x+1} G_+(t)\mathrm{d}t + \int_{-\infty}^{x} G_-(t)\mathrm{d}t = \rho G_-(x).$$

(iii) For any pair $G'_\pm$ with the same domains and codomains satisfying (i) and the inequalities obtained by replacing $=$ with $\le$ in (ii), we have $G_+(x) \le G'_+(x)$ and $G_-(x) \le G'_-(x)$ for all $x \le 0$.

If ρ is parameterized by $s \in (0, s_*)$ as $\rho = (1+e^s)/(2s)$, where $s_* := 1 + W_0(e^{-1})$, we have

$$\chi := \rho G_+(0) = \frac{1}{1+e^s}\int_0^1 \left(e^{2s(1-x)} - 1\right)\psi(x)\mathrm{d}x. \tag{4}$$

Building upon the above relation between ρ and χ with respect to ψ, we derive our main result on learning-augmented linear search.

Theorem 17. *Let $s_K \approx 0.5878$ denote the unique positive solution of $e^s(e^{2s}-1+2se^s) = (1+e^s)^2$, and define $K : (0, s_*] \to (0, e^{2s_*}]$ as*

$$K(s) := \begin{cases} \frac{e^s(e^{2s}-1+2se^s)}{(1+e^s)^2}, & \text{for } s \in (0, s_K]; \\ \xi(s), & \text{for } s \in (s_K, s_*], \end{cases}$$

where $\xi(s)$ is the unique solution $\xi \in [1, e^{2s}]$ of $(e^s - \xi)\ln\xi + \xi(3+e^{-s}) = e^s(e^s+2s-1)$. Then, for every trade-off parameter $s \in (0, s_*]$, there exists a $(\rho(s), \chi(s))$-excursion profile with

$$\rho(s) = \frac{1+e^s}{2s}, \quad \chi(s) = \begin{cases} \frac{e^{2s}-1-2s}{2s(1+e^s)} = \rho K(s)e^{-s} - 1, & \text{for } s \in (0, s_K], \\ \frac{e^{2s}+1+(1+K(s))\log K(s) - 2K(s) - 2s}{2s(1+e^s)}, & \text{for } s \in (s_K, s_*]. \end{cases}$$

Remark 18. The profile of Theorem 17 yields a ρ_{LS}-robust χ_{LS}-consistent strategy with $\rho_{\mathrm{LS}} = \frac{2}{\chi_{\mathrm{LS}}-1} + \frac{7}{3} + O(\chi_{\mathrm{LS}} - 1)$. In Appendix E.4, we present a lower bound of $\rho_{\mathrm{LS}} = \frac{2}{\chi_{\mathrm{LS}}-1} + \frac{3}{2} + \Omega(\chi_{\mathrm{LS}} - 1)$, demonstrating our strategy is near-optimal up to sub-leading constant. Both trade-off bounds are depicted in Figure 1b.

5 Conclusion

We study learning-augmented online bidding and completely characterize the robustness–consistency trade-off for randomized algorithms via bidding profiles. We further extend our approach to the linear search problem, demonstrating its versatility.

Several interesting directions arise from this work. A natural next step is to close the remaining gap in the robustness–consistency trade-off for linear search; we conjecture that our profile-based approach leads to a Pareto-optimal strategy.

Another important direction is to fully understand the *smoothness* of Pareto-optimal algorithms. Let $\eta := |T - P|$ denote the prediction error. For our online bidding strategy, when $T \le P$, monotonicity and consistency give an expected cost of at most $\chi P = \chi(T + \eta)$. In contrast, when T is infinitesimally larger than P, the expected competitive ratio jumps from χ to 2χ for $0 < s < \ln 2$ and to ρ for $\ln 2 \le s < 1$. This *brittleness* arises from the plateau at the predicted value in our bidding profile (when $\rho > e$). We note that the plateau can be removed by simply taking a convex combination of our profile and the profile of the classical e-competitive algorithm. Specifically, for any $0 < \varepsilon < e - \chi$, define

$$G_{\mathrm{new}}(x) := (1-c)G(x) + ce^x \text{ with } c := \frac{\varepsilon}{e-\chi}.$$

The resulting profile is strictly increasing and induces a strategy with a continuous expected competitive ratio, while preserving ρ-robustness and guaranteeing $(\chi+\varepsilon)$-consistency. However, the ratio can still increase steeply just above the prediction when ε is small. It would be interesting to determine what stronger smoothness guarantees are achievable for online bidding and linear search.

Finally, it would also be interesting to find other problems in which our profile-based approach can be effectively applied.

Acknowledgments and Disclosure of Funding

CL: This work was partly supported by an IITP grant funded by the Korean Government (MSIT) (No. RS-2020-II201361, Artificial Intelligence Graduate School Program (Yonsei University)). This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (RS-2025-00563707). DL and CY: The work of Dahoon Lee and Changki Yun was supported in part by the National Research Foundation of Korea (NRF) under Grant No. RS-2026-25484051 and by the New Faculty Startup Fund of Seoul National University. YS: This work was supported by the NCN grant no. 2020/39/B/ST6/01641. The authors acknowledge the use of LLMs for assisting with some proofs, editing the manuscript, and generating figures. All LLM-generated content has been carefully verified by the authors.

References

[ABDS24] Spyros Angelopoulos, Marcin Bienkowski, Christoph Dürr, and Bertrand Simon. Contract scheduling with distributional and multiple advice. In *The 33rd International Joint Conference in Artificial Intelligence (IJCAI-24)*, pages 3652–3660, 2024.

[ADJ+24] Spyros Angelopoulos, Christoph Dürr, Shendan Jin, Shahin Kamali, and Marc Renault. Online computation with untrusted advice. *Journal of Computer and System Sciences*, 144:103545, 2024.

[AGKP21] Keerti Anand, Rong Ge, Amit Kumar, and Debmalya Panigrahi. A regression approach to learning-augmented online algorithms. *Advances in Neural Information Processing Systems*, 34:30504–30517, 2021.

[Ang23] Spyros Angelopoulos. Online search with a hint. *Information and Computation*, 295:105091, 2023.

[APT22] Yossi Azar, Debmalya Panigrahi, and Noam Touitou. Online graph algorithms with predictions. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 35–66. SIAM, 2022.

[AS25] Spyros Angelopoulos and Bertrand Simon. Learning-augmented online bidding in stochastic settings. *Advances in Neural Information Processing Systems*, 38:67649–67679, 2025.

[Bec64] Anatole Beck. On the linear search problem. *Israel Journal of Mathematics*, 2(4):221–228, 1964.

[Bel63] Richard Bellman. An optimal search. *SIAM Review*, 5(3):274, 1963.

[BHLL25] Tian Bai, Zhiyi Huang, Chui Shan Lee, and Dongchen Li. Optimal stopping with a predicted prior. *arXiv preprint arXiv:2511.03289*, 2025.

[BMM25] Omar Besbes, Will Ma, and Omar Mouchtaki. Beyond IID: Data-driven decision making in heterogeneous environments. *Management Science*, 71(12):10538–10555, 2025.

[BMS20] Etienne Bamas, Andreas Maggiori, and Ola Svensson. The primal-dual method for learning augmented algorithms. *Advances in Neural Information Processing Systems*, 33:20083–20094, 2020.

[BN70] Anatole Beck and Donald J Newman. Yet more on the linear search problem. *Israel journal of mathematics*, 8(4):419–429, 1970.

[CCM25] Clément L Canonne, Kenny Chen, and Julián Mestre. With a little help from my friends: Exploiting probability distribution advice in algorithm design. *arXiv preprint arXiv:2505.04949*, 2025. To appear in ISAAC 2026.

[CD26] Qiming Cui and Michael Dinitz. Ski rental with distributional predictions of unknown quality. In *Proceedings of the 43rd International Conference on Machine Learning*, volume 306, pages 21805–21828. PMLR, 2026.

[CJS25] Davin Choo, Billy Jin, and Yongho Shin. Learning-augmented online bipartite fractional matching. *Advances in Neural Information Processing Systems*, 38:7671–7686, 2025.

[CKM06] Marek Chrobak and Claire Kenyon-Mathieu. Sigact news online algorithms column 10: Competitiveness via doubling. *ACM SIGACT News*, 37(4):115–126, 2006.

[CKNY08] Marek Chrobak, Claire Kenyon, John Noga, and Neal E Young. Incremental medians via online bidding. *Algorithmica*, 50(4):455–478, 2008.

[DFG06] Erik D Demaine, Sándor P Fekete, and Shmuel Gal. Online searching with turn cost. *Theoretical computer science*, 361(2-3):342–355, 2006.

[DIL+24] Michael Dinitz, Sungjin Im, Thomas Lavastida, Benjamin Moseley, Aidin Niaparast, and Sergei Vassilvitskii. Binary search with distributional predictions. *Advances in Neural Information Processing Systems*, 37:90456–90472, 2024.

[DZDA26] Mathis Degryse, Imrane Zaakour, Christoph Dürr, and Spyros Angelopoulos. The Pareto frontier of randomized learning-augmented online bidding. *arXiv preprint arXiv:2605.06106*, 2026. To appear in NeurIPS 2026.

[EADL24] Alex Elenter, Spyros Angelopoulos, Christoph Dürr, and Yanni Lefki. Overcoming brittleness in Pareto-optimal learning augmented algorithms. *Advances in Neural Information Processing Systems*, 37:9329–9357, 2024.

[IMXZ23] Sungjin Im, Benjamin Moseley, Chenyang Xu, and Ruilong Zhang. Online state exploration: Competitive worst case and learning-augmented algorithms. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 333–348. Springer, 2023.

[KF26] Jihwan Kim and Chenglin Fan. Robust and consistent ski rental with distributional advice. In *Proceedings of the 43rd International Conference on Machine Learning*, volume 306, pages 57411–57444. PMLR, 2026.

[KK25] Gregory Kehne and Thomas Kesselheim. Prophet and secretary at the same time. *arXiv preprint arXiv:2511.09531*, 2025. Appeared in EC 2026.

[Kra24] Evangelos Kranakis. A survey of the impact of knowledge on the competitive ratio in linear search. In *International Symposium on Stabilizing, Safety, and Security of Distributed Systems*, pages 23–38. Springer, 2024.

[KWP+11] Sertac Karaman, Matthew R Walter, Alejandro Perez, Emilio Frazzoli, and Seth Teller. Anytime motion planning using the RRT*. In *2011 IEEE international conference on robotics and automation*, pages 1478–1483. IEEE, 2011.

[LOAH14] Alejandro López-Ortiz, Spyros Angelopoulos, and Angele M Hamel. Optimal scheduling of contract algorithms for anytime problem-solving. *Journal of Artificial Intelligence Research*, 51:533–554, 2014.

[LV21] Thodoris Lykouris and Sergei Vassilvitskii. Competitive caching with machine learned advice. *Journal of the ACM (JACM)*, 68(4):1–25, 2021.

[MV22] Michael Mitzenmacher and Sergei Vassilvitskii. Algorithms with predictions. *Communications of the ACM*, 65(7):33–35, 2022.

[PSK18] Manish Purohit, Zoya Svitkina, and Ravi Kumar. Improving online algorithms via ML predictions. *Advances in Neural Information Processing Systems*, 31, 2018.

[RZ91] Stuart J Russell and Shlomo Zilberstein. Composing real-time systems. In *IJCAI*, volume 91, pages 212–217, 1991.

[SLLA25] Yongho Shin, Changyeol Lee, Gukryeol Lee, and Hyung-Chan An. Improved learning-augmented algorithms and (tight) lower bounds for multi-option ski rental problem. *ACM Trans. Algorithms*, 22(1), November 2025.

[Zil96] Shlomo Zilberstein. Using anytime algorithms in intelligent systems. *AI magazine*, 17(3):73–83, 1996.

A Further related work

Online bidding and linear search (also known as the cow-path problem) are classical problems in online optimization [CKNY08, Bel63, Bec64, BN70, DFG06, Kra24]. Techniques developed for these problems have played a fundamental role in the design and analysis of online algorithms, with applications including contract scheduling, incremental clustering, hierarchical clustering, and list scheduling. They have also influenced the study of offline optimization problems; for example, they have been used to establish connections between the minimum latency problem and k-TSP. We refer the interested reader to the survey of Chrobak and Kenyon-Mathieu [CKM06] and the references therein.

We briefly review previous work on learning-augmented algorithms for linear search. Angelopoulos [Ang23] proposed a deterministic Pareto-optimal learning-augmented algorithm for the problem. Later, Angelopoulos and Simon [AS25] studied randomized learning-augmented algorithms for linear search alongside their results for online bidding.

Recent work on learning-augmented algorithms has also shown growing interest in *distributional predictions* [DIL$^+$24, BHLL25, CCM25, KK25, BMM25, CD26, KF26]. In this direction, online bidding and linear search with distributional predictions have been studied in [ABDS24, AS25].

Another important aspect of learning-augmented algorithms is *smoothness*, which relates algorithmic performance to prediction error [MV22]. A substantial body of work has focused on designing algorithms whose performance degrades smoothly as the prediction error increases [IMXZ23, ABDS24, EADL24]. In contrast, this paper focuses on characterizing the trade-off between robustness and consistency, leaving the question of optimal smoothness for these problems as an interesting direction for future work.

For a comprehensive collection of work on learning-augmented algorithms (or algorithms with predictions), we refer the reader to the following repository.[6]

B Equivalence between multi-option ski rental and online bidding

It suffices to show a reduction from multi-option ski rental to online bidding with performance guarantees preserved, because the other direction has already been shown by [SLLA25].[7]

Lemma 19 (cf. Lemma 5.1 of [SLLA25])**.** *If there exists a ρ-robust, χ-consistent algorithm for multi-option ski rental, then for any constant $\varepsilon \in (0,1)$, there exists a $(\rho+\varepsilon)$-robust, $(\chi+\varepsilon)$-consistent algorithm for online bidding.*

In the multi-option ski rental problem, we are given a set of renting options, each of which is defined by a tuple $(c,d) \in \mathbb{R}_{>0} \times (\mathbb{Z}_{\geq 1} \cup \{\infty\})$, representing that we are allowed to go skiing for d days at cost c. The objective is to cover at the minimum total cost the whole skiing days $T_{\mathrm{ski}} \in \mathbb{Z}_{\geq 1}$, which is unknown until the last day. For the learning-augmented setting, we are additionally given a prediction $P_{\mathrm{ski}} \in \mathbb{Z}_{\geq 1}$ for the last day T_{ski}. For each $t \in \mathbb{Z}_{\geq 1}$, let $\mathrm{OPT}_{\mathrm{ski}}(t)$ denote the minimum cost of a multiset of renting options to span at least t days. For any $x \in \mathbb{R}_{\geq 0}$, let $B_{\mathrm{ski}}(x)$ denote a multiset of renting options that span the most number of days within the budget x. If every option costs more than x, we let $B_{\mathrm{ski}}(x)$ be the empty multiset.

We are now ready to describe the reduction of multi-option ski rental to online bidding. Let $\mathcal{A}$ denote a learning-augmented algorithm for online bidding. Given a multi-option ski rental instance (except unknown T_{ski}), we first feed to $\mathcal{A}$ the prediction $P := \mathrm{OPT}_{\mathrm{ski}}(P_{\mathrm{ski}})$. We then follow $\mathcal{A}$ by, if $\mathcal{A}$ makes a bid of x, appending $B_{\mathrm{ski}}(x)$ into our solution for multi-option ski rental. Whenever we are informed of the last day T_{ski}, we terminate.

For the analysis, let us fix T_{ski} and let T denote $\mathrm{OPT}_{\mathrm{ski}}(T_{\mathrm{ski}})$. Under the same random choices, the reduced algorithm for multi-option ski rental uses a prefix of the bids that $\mathcal{A}$ would make on target T with prediction $P = \mathrm{OPT}_{\mathrm{ski}}(P_{\mathrm{ski}})$, since any bid of at least T buys a multiset of renting options

[6] Alexander Lindermayr and Nicole Megow. `https://algorithms-with-predictions.github.io/`, Accessed: 2026-05-05.

[7] In fact, Shin et al. [SLLA25] established an (almost) factor-preserving reduction from the *button* problem, a finite variant of online bidding in which the minimum and maximum values are given in advance. One can then derive Lemma 19 by letting the maximum value tend to ∞.

sufficient for all T_{ski} days. As each chosen multiset costs at most its corresponding bid, the expected rental cost is at most the expected bidding cost on target T, which is at most ρT. If $T_{\text{ski}} = P_{\text{ski}}$, then $T = P$ and the expected cost is at most χT.

Lemma 20. *If there exists a ρ-robust, χ-consistent algorithm for online bidding, there exists a ρ-robust, χ-consistent algorithm for multi-option ski rental.*

C Sufficiency of strategies for online bidding

In this appendix, we discuss that it is sufficient to study strategies in analyzing the robustness-consistency trade-off of learning-augmented algorithms for online bidding.

We first argue that any bidding strategy can be converted into an algorithm for the setting where the target value T is guaranteed to be at least some constant $c_0 > 0$, with the same performance guarantees. We can still assume that $P = 1$ by normalizing the universe $[c_0, \infty)$ by $P \geq c_0$ from the beginning. Let $c := c_0/P$ denote the lower bound on the target after the normalization. Note that any reasonable deterministic strategy $x = \{x_n\}_{n\in\mathbb{Z}}$ must satisfy

$$\lim_{n\to-\infty} x_n = 0; \tag{5}$$

otherwise, the strategy would be non-robust. Therefore, we can always find the minimum index $n_c(x)$ at which the sequence has a value at least c, i.e., $n_c(x) := \min\{n \in \mathbb{Z} : x_n \geq c\}$.

Consider the bidding algorithm represented by the suffix of x from index $n_c(x)$, i.e., the algorithm sequentially submits bids $x_{n_c(x)}, x_{n_c(x)+1}, x_{n_c(x)+2}, \ldots$ in this order until bidding at least T. It is easy to see that the algorithm always incurs less cost than the strategy by the cut-off prefix sum $\sum_{n<n_c(x)} x_n$, implying that the robustness and consistency follow from the strategy.

For the randomized setting, note again that any reasonable randomized strategy satisfies (5) almost surely to guarantee a bounded expected robustness. Hence, sampling a deterministic strategy from the randomized strategy and executing the above process with this sampled strategy would yield a randomized bidding algorithm with the same robustness and consistency guarantees.

It remains to justify that the lower bound on the robustness-consistency trade-off for bidding strategies also applies to bidding algorithms. Let us assume $c_0 = 1$. We replicate the proof of Lemma 12, but subject to a ρ-robust, χ-consistent algorithm $\mathcal{A}$. For each P, denote the aggregate measure of (normalized) bids of $\mathcal{A}$ equipped with prediction P as μ_P. Without loss of generality, we discard the bids below $1/P$. From the robustness of $\mathcal{A}$, the local bounds are reproduced in μ_P as well.

$$\frac{1}{2} \leq \mu_P([T, (2\rho-1)T)) \leq \rho(2\rho-1), \quad T \geq \frac{1}{P}. \tag{6}$$

One can define the inverse 'profile' $G_P : \mathbb{R} \to [0, \infty)$ in the same way. G_P is still non-decreasing, left-continuous, and satisfies the offset and consistency condition. What lacks to be a bidding profile is that G_P may be zero, and the robustness condition is restricted to its support. Denote $l_P := \sup\{x : G_P(x) = 0\}$, then ρ-robustness of the algorithm implies

$$\int_{l_P}^{x+1} G_P(t)\mathrm{d}t \leq \rho G_P(x), \quad x > l_P.$$

Now we take G_P to the limit of $P \to \infty$. Note that l_P is also unbounded. From (6),

$$-l_P = m((l_P, 0)) = \mu_P([1/P, 1)) \geq \frac{1}{2}\lfloor \log_{(2\rho-1)} P \rfloor \to \infty,$$

as $P \to \infty$. Therefore, define

$$G_\infty(x) := \liminf_{P\to\infty} G_P(x).$$

Therefore, for any $N > 0$, $l_P < -N$ for sufficiently large P. Note that G_∞ is still non-decreasing, but not necessarily left-continuous.

From Fatou's lemma, for any infinitesimal $\varepsilon > 0$ and $x > -N + \varepsilon$,

$$\begin{aligned}\int_{-N}^{x+1-\varepsilon} G_\infty(t)\mathrm{d}t &\le \liminf_{P\to\infty} \int_{-N}^{x+1-\varepsilon} G_P(t)\mathrm{d}t \\ &\le \liminf_{P\to\infty} \int_{l_P}^{x+1-\varepsilon} G_P(t)\mathrm{d}t \\ &\le \liminf_{P\to\infty} \rho G_P(x-\varepsilon) = \rho G_\infty(x-\varepsilon).\end{aligned}$$

Letting $N \to \infty$ and $\varepsilon \downarrow 0$, we obtain

$$\int_{-\infty}^{x+1} G_\infty(t)\mathrm{d}t \le \rho G_\infty(x-),$$

for any $x \in \mathbb{R}$. Therefore, the left-continuous representative of G_∞ preserves the ρ-robustness condition. The χ-consistency condition follows analogously. Note that, from Remark 4, G_∞ is globally positive. One subtle issue is the offset condition; $G_\infty(x)$ might be 1 for some $x < 0$, but the robustness condition implies that $\tau := \sup\{x : G_\infty(x) < 1\}$ is nonpositive and finite. Putting

$$G(x) := G_\infty(\tau + x)$$

satisfies the offset condition additionally, securing all the other conditions.

Therefore, G is a (ρ, χ)-bidding profile, yielding a ρ-robust, χ-consistent strategy by Corollary 6.

D Additional details on online bidding

D.1 Deferred proofs from Section 3.1

Remark 4. Note that $G(0) \le 1$, from its left-continuity and the offset condition. Also, the strict positivity of G can be deduced from the robustness condition.

Proof. If G is zero at some point, take the supremum point x_0 for the zero set. The left-continuity implies that $G(x_0) = 0$, and from the robustness condition,

$$\int_{-\infty}^{x_0+1} G(t)\mathrm{d}t \le \rho G(x_0) = 0,$$

implying that $G(t) \equiv 0$ for $(-\infty, x_0 + 1)$, contradicting the maximality of x_0. □

Lemma 5. *Let G denote a (ρ, χ)-bidding profile. Then, for any $T > 0$,*

$$\tau(T) := \sup\{t \in \mathbb{R} : G(t) < T\} < \infty, \quad \textit{and} \quad \mathsf{cost}_G(T) = \textstyle\int_{-\infty}^{\tau(T)+1} G(t)\mathrm{d}t.$$

Proof. We first claim that $\tau(T) \le \rho T - 1 < \infty$. Note that

$$G(\rho T - 1) \ge \frac{1}{\rho}\int_{-\infty}^{\rho T} G(t)\,dt \ge \frac{1}{\rho}\int_0^{\rho T} 1\,dt = T,$$

where the first inequality follows from the robustness condition and the second inequality from the offset condition. Since G is non-decreasing, $G(t) \ge T$ if $t \ge \rho T - 1$, hence $\tau(T) \le \rho T - 1$.

Next, we analyze the expected cost $\mathsf{cost}_G(T)$. After $U \sim \mathrm{Unif}(0, 1]$ is sampled, note that for any $n \in \mathbb{Z}$, the algorithm makes the bid $G(n + U)$ if and only if $G(n - 1 + U) < T$. Therefore,

$$\begin{aligned}\mathsf{cost}_G(T) &= \mathbb{E}_{U\sim\mathrm{Unif}(0,1]} \sum_{n\in\mathbb{Z}} G(n+U)\mathbf{1}_{G(n-1+U)<T} \\ &= \sum_{n\in\mathbb{Z}} \int_0^1 G(n+u)\mathbf{1}_{G(n-1+u)<T}\,\mathrm{d}u \\ &= \int_{\{t:G(t-1)<T\}} G(t)\mathrm{d}t \\ &= \int_{-\infty}^{\tau(T)+1} G(t)\mathrm{d}t,\end{aligned}$$

which concludes the proof. □

D.2 Deferred proofs in Section 3.2

Recall that

$$\rho(s) := \frac{e^s}{s}, \qquad \chi(s) := \begin{cases} (e^s - 1)/s & (0 < s \le \ln 2), \\ \xi(s)/s & (\ln 2 \le s \le 1), \end{cases}$$

where $\xi(s)$ is the unique solution $\xi \in [1, e]$ satisfying $\xi(2 - \ln \xi) = e^s$.

We show the lemma that will be useful in the subsequent proof.

Lemma 21. *Fix a positive number $0 < s \le 1$. For some χ, let G be a tight $(\rho(s), \chi)$-bidding profile induced by an integrable function $\phi : (0, 1] \to [1, \infty)$. Then*

$$\chi \ge e^s \int_0^1 e^{-sx} \phi(x) \mathrm{d}x.$$

Proof. Let us use ρ to denote $\rho(s)$ for simplicity. For $x \in \mathbb{R}$, let

$$A(x) := \int_{-\infty}^{x} G(t) \mathrm{d}t.$$

Note that $A'(x) = G(x)$ a.e. (almost everywhere) as G is non-decreasing. From the "tightness" of a tight bidding profile G for $x \le 0$, we obtain a delayed-differential equation (which holds in equality):

$$\rho A'(x) = A(x + 1), \qquad x \le 0, \text{ a.e.}$$

Let us assume that $e^{-sx} A(x)$ is non-decreasing. Then the following is well-defined:

$$L := \lim_{t \to -\infty} e^{-st} A(t) \ge 0.$$

Consider any $\kappa < -1$. By tightness, we have

$$\int_{\kappa}^{0} e^{-sx} \big(\rho A'(x) - A(x + 1) \big) \mathrm{d}x = 0.$$

Integrating by parts,

$$\begin{aligned} \int_{\kappa}^{0} e^{-sx} A(x + 1) \mathrm{d}x &= e^s \int_{\kappa+1}^{1} e^{-sx} A(x) \mathrm{d}x \\ &= e^s \left(-\frac{e^{-sx} A(x)}{s} \bigg|_{\kappa+1}^{1} + \int_{\kappa+1}^{1} \frac{e^{-sx} A'(x)}{s} \mathrm{d}x \right) \\ &= -\frac{\chi}{s} + \rho e^{-s(\kappa+1)} A(\kappa + 1) + \rho \int_{\kappa+1}^{1} e^{-sx} A'(x) \mathrm{d}x, \end{aligned}$$

where we use $A(1) = \chi$ and $\rho = e^s/s$ for the last equality. We thus obtain

$$\begin{aligned} 0 &= \frac{\chi}{s} - \rho e^{-s(\kappa+1)} A(\kappa + 1) + \rho \int_{\kappa}^{0} e^{-sx} A'(x) \mathrm{d}x - \rho \int_{\kappa+1}^{1} e^{-sx} A'(x) \mathrm{d}x \\ &= \frac{\chi}{s} - \rho e^{-s(\kappa+1)} A(\kappa + 1) + \rho \int_{\kappa}^{\kappa+1} e^{-sx} A'(x) \mathrm{d}x - \rho \int_{0}^{1} e^{-sx} A'(x) \mathrm{d}x. \end{aligned}$$

For the integral from κ to $\kappa + 1$, we have

$$\int_{\kappa}^{\kappa+1} e^{-sx} A'(x) \mathrm{d}x = e^{-s(\kappa+1)} A(\kappa + 1) - e^{-s\kappa} A(\kappa) + s \int_{\kappa}^{\kappa+1} e^{-sx} A(x) \mathrm{d}x.$$

Substituting this into the preceding equation, we obtain

$$0 = \frac{\chi}{s} - \rho e^{-s\kappa} A(\kappa) + \rho s \int_{\kappa}^{\kappa+1} e^{-sx} A(x) \mathrm{d}x - \rho \int_{0}^{1} e^{-sx} A'(x) \mathrm{d}x.$$

Note that $\lim_{\kappa\to-\infty}\int_\kappa^{\kappa+1} e^{-sx}A(x)\mathrm{d}x = L$, since the monotonicity of $e^{-sx}A(x)$ gives

$$e^{-s\kappa}A(\kappa) \le \int_\kappa^{\kappa+1} e^{-sx}A(x)\mathrm{d}x \le e^{-s(\kappa+1)}A(\kappa+1)$$

and both the lower and upper bounds converge to L as $\kappa \to -\infty$. Thus, taking $\kappa \to -\infty$ yields

$$0 = \frac{\chi}{s} - \rho(1-s)L - \rho\int_0^1 e^{-sx}A'(x)\mathrm{d}x.$$

Since $A'(x) = \phi(x)$ for almost every $x \in (0,1]$ and $\rho = e^s/s$, we obtain

$$\chi = e^s(1-s)L + e^s\int_0^1 e^{-sx}\phi(x)\mathrm{d}x \ge e^s\int_0^1 e^{-sx}\phi(x)\mathrm{d}x,$$

where the inequality follows from $L \ge 0$ and $0 < s \le 1$.

It remains to show $e^{-sx}A(x)$ is non-decreasing. Since A is locally absolutely continuous and $x \mapsto e^{-sx}$ is smooth, their product is locally absolutely continuous. Therefore,

$$e^{-sx}A(x) \text{ is non-decreasing on } \mathbb{R} \iff A'(x) \ge sA(x) \text{ a.e.} \tag{7}$$

To this end, we show that $A'(x) \ge sA(x)$ a.e. First observe that,

$$\begin{aligned}
\rho A'(x) &\ge A(x+1)\\
&= A(x) + \int_x^{x+1} A'(t)\mathrm{d}t\\
&\ge A(x) + \int_x^{x+1} A'(x)\mathrm{d}t\\
&= A(x) + A'(x),
\end{aligned}$$

where the inequality is due to the fact that $A'(t) = G(t)$ is non-decreasing. Thus,

$$A'(x) \ge \frac{1}{\rho-1}A(x).$$

Now define $c_0 := \frac{1}{\rho-1}$, and $c_n := e^{c_{n-1}}/\rho$ for $n \in \mathbb{Z}_{\ge 1}$. We show $A'(x) \ge c_nA(x)$. From definition, $A'(x) \ge c_0A(x)$. Let us assume $A'(x) \ge c_{n-1}A(x)$. Since

$$\ln A(x+1) - \ln A(x) = \int_x^{x+1}\frac{A'(t)}{A(t)}\mathrm{d}t \ge c_{n-1},$$

we have that

$$A'(x) \ge \frac{1}{\rho}A(x+1) \ge \frac{1}{\rho}e^{c_{n-1}}A(x) = c_nA(x).$$

We now show $\{c_n\}$ converges to s as $n \to \infty$. We first show that it is increasing and bounded from above by s. Note that $c_0 < s$. Suppose $c_{n-1} < s$. It is easy to see that $c_n = e^{c_{n-1}}e^{-s}s < s$. Since $f(z) := e^z/z$ strictly decreases on $z \in (0,1]$,

$$\frac{c_n}{c_{n-1}} = \frac{e^{c_{n-1}}/\rho}{c_{n-1}} = \frac{f(c_{n-1})}{f(s)} > 1,$$

implying $c_n > c_{n-1}$. By the monotone convergence theorem, $c_n \to c^*$. Since $e^{c^*} = \rho c^*$, $c^* = s$. Therefore, $A'(x) \ge c^*A(x) = sA(x)$, showing (7). □

We show a useful fact.

Fact 22. *Given $s \in (0,1]$, $s > \ln 2 \iff s\chi(s) > 1$. Moreover, if $s\chi(s) > 1$, $0 \le x_0 < 1$ where $x_0 := 1 - \frac{\ln(s\chi(s))}{s}$.*

Proof. We first prove the first equivalence, beginning with the if part. Suppose toward contradiction that $s \le \ln 2$. Then $s\chi(s) = e^s - 1$. On the other hand, $e^s - 1 > 1 \iff s > \ln 2$.

For the only if part, fix any $s > \ln 2$. From $e^s > 2$ and the fact that $x \mapsto x(2 - \ln x)$ strictly increases on $[1, e]$, $\xi(s)(2 - \ln \xi(s)) = e^s > 2 = 1 \cdot (2 - \ln 1)$, we obtain $\xi(s) > 1$. Observe $s\chi(s) = \xi(s)$ from the definition of $\chi(s)$.

For the second statement, $x_0 < 1$ follows from $s\chi(s) > 1$. Note that $s\chi(s) = \xi(s)$ since $s > \ln 2$. Furthermore, $2 - \ln \xi(s) \ge 1$ since $\xi(s) \in [1, e]$. Therefore,

$$s\chi(s) = \xi(s) \le \xi(s)(2 - \ln \xi(s)) = e^s.$$

The following equivalence completes the proof. $x_0 \ge 0 \iff s\chi(s) \le e^s$. □

D.2.1 Optimal bidding profile

Theorem 7. *For any $s \in (0, 1]$, there exists a $(\rho(s), \chi(s))$-bidding profile, where*

$$\rho(s) := \frac{e^s}{s} \quad \textit{and} \quad \chi(s) := \begin{cases} (e^s - 1)/s & (0 < s \le \ln 2), \\ \xi(s)/s & (\ln 2 \le s \le 1), \end{cases}$$

where $\xi(s)$ is the unique solution $\xi \in [1, e]$ satisfying $\xi(2 - \ln \xi) = e^s$ for any $s \in [\ln 2, 1]$.

Observe that $\xi(s)$ is well defined since $\xi \mapsto \xi(2 - \ln \xi)$ is increasing on $\xi \in [1, e]$ with $1 \mapsto 2$ and $e \mapsto e$.

Proof. It is trivial for $s = 1$; the classical (e, e)-bidding profile $G(x) = e^x$ suffices.

Now fix any $0 < s < 1$. Let us use ρ and χ to denote $\rho(s)$ and $\chi(s)$, respectively. Throughout the proof let $\phi(x) := \max(1, s\chi \cdot e^{s(x-1)})$ be a function defined on $(0, 1]$.

We first define G when $x > 0$:

$$G(x) := \begin{cases} \phi(x), & 0 < x \le 1, \\ \max(1, s\chi) \cdot e^{s(x-1)}, & x > 1. \end{cases}$$

It remains to define G on $x \le 0$; we inductively define the remaining G. Let $G(0) = \frac{\chi}{\rho}$. Given G defined on $[k, k+1)$ for an integer $k \le 0$, we define G on $[k-1, k)$ by

$$G(x) = G(k) - \frac{1}{\rho} \int_{x+1}^{k+1} G(t) \mathrm{d}t.$$

Note that G is continuous on $x \le 0$.

Suppose G is non-negative and non-decreasing. We show that G is a (ρ, χ)-bidding profile. The offset condition trivially holds. From the definition, for any integer $k \le 0$ and $x \in [k-1, k)$,

$$G(x) \ge G(x) - G(k) + \sum_{n \le k} (G(n) - G(n-1)) = \frac{1}{\rho} \int_{-\infty}^{x+1} G(t) \mathrm{d}t, \tag{8}$$

which shows the robustness when $x < 0$. The above also holds when $x = 0$ by the choice of $G(0)$, showing the consistency condition and the robustness condition when $x = 0$.

We show the robustness condition when $x > 0$. Suppose $s\chi \le 1$. From Fact 22, $s \le \ln 2$. Then,

$$\chi = \frac{e^s - 1}{s}, \qquad G(x) = 1 \quad (0 < x \le 1), \qquad G(x) = e^{s(x-1)} \quad (x \ge 1).$$

Note that

$$\int_{-\infty}^{x+1} G(t) \mathrm{d}t \le \chi + \frac{e^{sx} - 1}{s}.$$

If $0 < x \le 1$, then

$$\chi + \frac{e^{sx} - 1}{s} = \frac{e^s + e^{sx} - 2}{s} \le \frac{e^s}{s} = \rho G(x),$$

because $s \le \ln 2$ implies $e^{sx} \le 2$. If $x \ge 1$, then

$$\chi + \frac{e^{sx} - 1}{s} \le \frac{e^{sx}}{s} = \rho e^{s(x-1)} = \rho G(x),$$

because $\chi \le 1/s$.

Now suppose $s\chi > 1$. Then $G(x) = s\chi \cdot e^{s(x-1)}$ for $x \ge 1$. Hence

$$\int_{-\infty}^{x+1} G(t)\mathrm{d}t \le \chi + \int_{1}^{x+1} s\chi \cdot e^{s(t-1)}\mathrm{d}t = \chi e^{sx}.$$

It is easy to see that the robustness condition holds when $x > 1$. Consider any $x \in (0, 1]$. If $G(x) = s\chi \cdot e^{s(x-1)}$, then

$$\rho G(x) = \frac{e^s}{s} \cdot s\chi \cdot e^{s(x-1)} = \chi e^{sx}.$$

If $G(x) = 1$, which implies $s\chi \cdot e^{s(x-1)} \le 1$, then the above equation shows $\chi e^{sx} \le \rho = \rho G(x)$.

Note that (8) holds in equality. Suppose toward contradiction that $\lim_{n\to-\infty} G(n) = \varepsilon$ for some $\varepsilon > 0$. Then, for any integer $N < 0$, $G(0) - G(N) \ge \frac{|N|\varepsilon}{\rho}$ since G is non-decreasing. Sufficiently small N leads to $G(N) < 0$, contradicting the non-negativity of G. Therefore, G is a tight (ρ, χ)-bidding profile.

In what follows, we show G is non-negative and non-decreasing. Recall the definition of the functional operator $\mathcal{F}$.

Definition 8. Given $\rho > 0$, let $\mathcal{F}$ be a functional operator on an integrable, non-decreasing, non-negative function $H : (-\infty, 1] \to [0, \infty)$, giving a function $\mathcal{F}H : (-\infty, 1] \to [0, \infty)$ such that

$$\mathcal{F}H(x) := \begin{cases} H(x), & 0 < x \le 1, \\ \frac{1}{\rho}\int_{-\infty}^{x+1} H(t)\mathrm{d}t, & x \le 0. \end{cases}$$

Observe that our bidding profile G on $x \le 0$ can be equivalently represented as the $(-\infty, 0]$ region of a non-negative, non-decreasing function G^* defined on $(-\infty, 1]$ such that $G^* = \mathcal{F}G^*$, $\int_{-\infty}^{1} G^*(t)\mathrm{d}t = \chi$ and $G^* \equiv \phi$ on $(0, 1]$.This can be seen from $G^*(0) = \chi/\rho = G(0)$, and for $x \in [k-1, k)$ for some integer $k \le 0$,

$$\begin{aligned} G^*(x) = \frac{1}{\rho}\int_{-\infty}^{x+1} G^*(t)\mathrm{d}t &= \frac{1}{\rho}\left(\int_{-\infty}^{k+1} G^*(t)\mathrm{d}t - \int_{x+1}^{k+1} G^*(t)\mathrm{d}t\right) \\ &= G^*(k) - \frac{1}{\rho}\int_{x+1}^{k+1} G^*(t)\mathrm{d}t = G(x) \end{aligned}$$

where the first and the third equalities are from $G^* = \mathcal{F}G^*$ and the last equality is from the inductive hypothesis.

We show such G^* exists. We exploit the *order-preserving* property of $\mathcal{F}$: if $A \le B$ (i.e., $A(x) \le B(x)$ for all x in the domain) for two functions A, B, then $\mathcal{F}A \le \mathcal{F}B$. Let $G^{(0)}$ be a function defined on $x \le 1$ such that $G^{(0)}(x) = 0$ for $x \le 0$ and $G^{(0)}(x) = \phi(x)$ for $0 < x \le 1$. For $n \in \mathbb{Z}_{\ge 0}$, $G^{(n+1)}(x) := \mathcal{F}G^{(n)}(x)$. Note that the order-preserving property shows the following implication:

$$G^{(n)} \le G^{(n+1)} \Longrightarrow \mathcal{F}G^{(n)} \le \mathcal{F}G^{(n+1)} \Longrightarrow G^{(n+1)} \le G^{(n+2)}.$$

Since $G^{(0)} \le G^{(1)}$, we have $G^{(n)} \le G^{(n+1)}$ for all $n \in \mathbb{Z}_{\ge 0}$.

We show for all $n \in \mathbb{Z}_{\ge 0}$, $G^{(n)}$ is bounded from above by $H(x) := e^{cx}$ for $x \le 1$ where $c \in (s, 1)$.

First we show $\mathcal{F}H \le H$. For $0 < x \le 1$, we have $(\mathcal{F}H)(x) = H(x)$ by definition. For $x \le 0$, we have

$$(\mathcal{F}H)(x) = \frac{1}{\rho}\int_{-\infty}^{x+1} e^{ct}\mathrm{d}t = \frac{e^c}{\rho c}H(x) \le H(x),$$

where the inequality follows from $e^c/c < e^s/s = \rho$ because $s < c < 1$.

Note that, for every positive integer k, both $\mathcal{F}^{(k)}G^{(0)}$ and $\mathcal{F}^{(k)}H$ lie in the domain of $\mathcal{F}$. Along with $G^{(0)} \leq H$ (which holds from the definition), we obtain the following:

$$G^{(n)} \leq \mathcal{F}^{(n)}G^{(0)} \leq \mathcal{F}^{(n)}H \leq H.$$

By the monotone convergence theorem, $\{G^{(n)}\}$ converges pointwise. We show that $G^* := \lim_{n\to\infty} G^{(n)}$ satisfies the required conditions. It is non-negative and non-decreasing, with $G^* \equiv \phi$ on $(0, 1]$ and, by the monotone convergence theorem, $G^* = \mathcal{F}G^*$.

It remains to show that $\chi^* := \int_{-\infty}^{1} G^*(t)\mathrm{d}t$ equals χ. Recall that $G^*(x) \leq e^{cx}$ for $x \leq 0$, where $c > s$. Thus,

$$L := \lim_{x\to-\infty} e^{-sx} \int_{-\infty}^{x} G^*(t)\mathrm{d}t = 0.$$

For $x \leq 0$, the equality $G^* = \mathcal{F}G^*$ gives $\rho G^*(x) = \int_{-\infty}^{x+1} G^*(t)\mathrm{d}t$. Following the proof of Lemma 21, we obtain

$$\chi^* = e^s(1-s)L + e^s \int_0^1 e^{-sx}\phi(x)\mathrm{d}x = e^s \int_0^1 e^{-sx}\phi(x)\mathrm{d}x.$$

Suppose $s\chi \leq 1$. From Fact 22, $s \leq \ln 2$, which implies $\phi(x) = 1$ for all $0 < x \leq 1$, and thus $\chi^* = \chi$. Otherwise, let $x_0 := 1 - \ln(s\chi)/s \in [0, 1)$. Then

$$\begin{aligned}
\chi^* &= e^s \left(\int_0^{x_0} e^{-sx} \cdot 1\mathrm{d}x + \int_{x_0}^1 e^{-sx} s\chi \cdot e^{s(x-1)}\mathrm{d}x \right) \\
&= \frac{e^s - s\chi}{s} + \chi \ln(s\chi) \\
&= \frac{\xi(2 - \ln\xi) - s\chi}{s} + \chi \ln(s\chi) \\
&= \frac{s\chi(2 - \ln(s\chi)) - s\chi}{s} + \chi \ln(s\chi) \\
&= \chi.
\end{aligned}$$

□

D.2.2 Lower bounds

Lemma 10. *Suppose there exists a (ρ, χ)-bidding profile G for fixed positive numbers $\rho > \chi > 1$. Let $\phi : (0, 1] \to [1, \infty)$ denote the function that induces G. Then, there exists a tight (ρ, χ')-bidding profile G' induced by ϕ such that $\chi' \leq \chi$.*

Proof. Recall the functional operator $\mathcal{F}$ defined in the proof of Theorem 7. At a high level, we show that applying $\mathcal{F}$ to G yields a (ρ, χ')-bidding profile with $\chi' \leq \chi$. We then show that by iteratively applying $\mathcal{F}$ to G, the process converges to a tight bidding profile. Let $G^{(0)} := G|_{(-\infty,1]}$, and for $n \in \mathbb{Z}_{\geq 0}$, $G^{(n+1)}(x) := \mathcal{F}G^{(n)}(x)$. Note that the order-preserving property of $\mathcal{F}$ shows the following implication:

$$G^{(n)} \geq G^{(n+1)} \Longrightarrow \mathcal{F}G^{(n)} \geq \mathcal{F}G^{(n+1)} \Longrightarrow G^{(n+1)} \geq G^{(n+2)}.$$

From the robustness condition of bidding profiles and $\mathcal{F}$, for $x \leq 0$,

$$G^{(0)}(x) \geq \frac{1}{\rho} \int_{-\infty}^{x+1} G^{(0)}(t)\mathrm{d}t = \mathcal{F}G^{(0)}(x) = G^{(1)}(x).$$

Since $G^{(0)}(x) = \mathcal{F}G^{(0)}(x)$ for $x \in (0, 1]$, $G^{(0)} \geq G^{(1)}$ and thus we have $G^{(n)} \geq G^{(n+1)}$ for all $n \in \mathbb{Z}_{\geq 0}$. Since applying $\mathcal{F}$ to a non-negative function always yields a non-negative function, $\{G^{(n)}\}$ converges to a function, say G^*, as $n \to \infty$ by the monotone convergence theorem.

To see that G^* glued with G for $x > 1$ is a tight (ρ, χ')-bidding profile for some $\chi' \leq \chi$, we simply follow the proof of Theorem 7. In particular,

$$\chi' := \int_{-\infty}^{1} G^*(t)\mathrm{d}t \leq \int_{-\infty}^{1} G(t)\mathrm{d}t \leq \chi.$$

This completes the proof. □

We show the following lemma before proving Theorem 11.

Lemma 23. *Fix $0 < s \leq 1$, and an integrable function $\phi : (0, 1] \to [1, \infty)$. If there exists a tight $(\rho := \rho(s), \chi)$-bidding profile induced by some ϕ, then*

$$\phi(x) \geq \max(1, s\chi e^{s(x-1)}).$$

Proof. As in the proof of Lemma 21, define $A(x) := \int_{-\infty}^{x} G(t)\mathrm{d}t$. Recall from (7) in the proof of Lemma 21 that $A'(x) \geq sA(x)$ a.e. Therefore, for $x > 0$,

$$\ln A(x+1) - \ln \chi = \ln A(x+1) - \ln A(1) = \int_1^{x+1} \frac{A'(t)}{A(t)}\mathrm{d}t \geq \int_1^{x+1} s\mathrm{d}t = sx.$$

From the robustness condition for $0 < x \leq 1$,

$$\phi(x) = G(x) \geq \frac{1}{\rho}A(x+1) \geq \frac{1}{\rho}A(1)e^{sx} = s\chi e^{s(x-1)}.$$

Meanwhile, we have $\phi(x) \geq 1$ from the definition of a bidding profile, completing the proof. □

We are now ready to prove the following theorem.

Theorem 11. *For any $0 < s \leq 1$, if there exists a tight $(\rho(s), \chi)$-bidding profile, then $\chi \geq \chi(s)$.*

Proof. Let G be a tight $(\rho(s), \chi)$-bidding profile, and let ϕ be a function on $(0, 1]$ where G is induced. From Lemmas 21 and 23,

$$\chi \geq e^s \int_0^1 e^{-sx}\phi(x)\mathrm{d}x \geq e^s \int_0^1 e^{-sx} \max(1, s\chi e^{s(x-1)})\mathrm{d}x. \tag{9}$$

Note that the following holds in general.

$$\chi \geq e^s \int_0^1 e^{-sx}\mathrm{d}x = \frac{e^s - 1}{s}.$$

Suppose toward contradiction that $\chi < \chi(s)$. If $s \leq \ln 2$, then $\chi(s) = (e^s - 1)/s$ which contradicts $\chi \geq (e^s - 1)/s$. Thus, it must be the case that $s > \ln 2$. Note $s\chi \geq e^s - 1 > 1$.

From Fact 22, $s\chi(s) > 1$ and $x_0 \in [0, 1)$ where

$$x_0 := 1 - \frac{\ln(s\chi(s))}{s}.$$

Let

$$y_0 := 1 - \frac{\ln(s\chi)}{s}.$$

Note that $s\chi > 1$ implies $y_0 < 1$. Since $s\chi < s\chi(s)$, we have $y_0 > x_0$. Thus, $y_0 \in (x_0, 1) \subseteq (0, 1)$.

By further bounding (9) from below,

$$\chi \geq e^s \left(\int_0^{y_0} e^{-sx}\mathrm{d}x + \int_{y_0}^1 e^{-sx} s\chi e^{s(x-1)}\mathrm{d}x \right) = \frac{e^s - s\chi}{s} + \chi \ln(s\chi),$$

implying

$$s\chi(2 - \ln(s\chi)) \geq e^s.$$

Let $f(x) := x(2 - \ln x)$. Recall that $f(x)$ is strictly increasing on $[1, e]$. By definition, $\xi(s) = s\chi(s) \in [1, e]$ satisfies $f(s\chi(s)) = e^s$. Recall $1 < s\chi < s\chi(s) \leq e$. The strict monotonicity of f implies

$$f(s\chi) < f(s\chi(s)) = e^s,$$

which directly contradicts $f(s\chi) = s\chi(2 - \ln(s\chi)) \geq e^s$. □

D.3 Deferred proofs from Section 3.3

Lemma 12. *For any fixed positive real numbers ρ and χ such that $\rho > \chi > 1$, there exists a ρ-robust χ-consistent online bidding strategy if and only if there exists a (ρ, χ)-bidding profile G.*

Proof. The if part is proved in Corollary 6. We prove the only-if part.

Let any ρ-robust χ-consistent (randomized) strategy $\mathcal{A}$ be given, whose prediction is given as $P = 1$. Recall $\mathcal{A}$ is a probability distribution over bi-infinite sequences; let $(Q_n)_{n\in\mathbb{Z}}$ denote a random sequence. Since the strategy is ρ-robust and χ-consistent,

$$\mathsf{cost}_{\mathcal{A}}(T) = \mathbb{E}\left[\sum_{n\in\mathbb{Z}} Q_n \mathbf{1}_{Q_{n-1}<T}\right] \le \rho T, \quad \forall T > 0, \tag{10}$$

and

$$\mathsf{cost}_{\mathcal{A}}(1) = \mathbb{E}\left[\sum_{n\le 0} Q_n\right] \le \chi. \tag{11}$$

Let μ_n be the marginal probability distribution of Q_n, and μ be its aggregate measure.

$$\mu := \sum_{n\in\mathbb{Z}} \mu_n.$$

In other words, for every bounded interval $I \subseteq (0,\infty)$,

$$\mu(I) = \sum_{n\in\mathbb{Z}} \mathbb{E}\left[\mathbf{1}_{Q_n\in I}\right] = \mathbb{E}\left[\sum_{n\in\mathbb{Z}} \mathbf{1}_{Q_n\in I}\right].$$

Clearly, μ is supported on $(0,\infty)$. Now we verify that μ is locally finite on $(0,\infty)$ while $\mu((0,1)) = \mu([1,\infty)) = \infty$. For any $T \in (0,\infty)$, consider a set $[T,(2\rho-1)T)$ and let $x := \mu([T,(2\rho-1)T))$. Since $\mathsf{cost}_{\mathcal{A}}((2\rho-1)T) \le \rho(2\rho-1)T$, we have

$$\begin{aligned} Tx &= \mathbb{E}\left[\sum_{n\in\mathbb{Z}} T\cdot \mathbf{1}_{Q_n\in[T,(2\rho-1)T)}\right] \\ &\le \mathbb{E}\left[\sum_{n\in\mathbb{Z}} Q_n\cdot \mathbf{1}_{Q_n\in[T,(2\rho-1)T)}\right] \\ &\le \mathbb{E}\left[\sum_{n\in\mathbb{Z}} Q_n\cdot \mathbf{1}_{Q_{n-1}<(2\rho-1)T}\right] \le \rho(2\rho-1)T, \end{aligned}$$

thus $x \le \rho(2\rho-1)$. On the other hand, from $\mathsf{cost}_{\mathcal{A}}(T) \le \rho T$, we have

$$\begin{aligned} \rho T &\ge \mathbb{E}\left[\sum_{n\in\mathbb{Z}} Q_n\cdot \mathbf{1}_{Q_{n-1}<T}\right] \\ &\ge \mathbb{E}\left[\sum_{n\in\mathbb{Z}} Q_n\cdot \mathbf{1}_{Q_{n-1}<T\le Q_n}\right] \\ &\ge \mathbb{E}\left[\sum_{n\in\mathbb{Z}} Q_n\cdot \mathbf{1}_{Q_{n-1}<T\le Q_n<(2\rho-1)T}\right] + \mathbb{E}\left[\sum_{n\in\mathbb{Z}} Q_n\cdot \mathbf{1}_{Q_{n-1}<T<(2\rho-1)T\le Q_n}\right] \\ &\ge T\cdot\min\{1,x\} + (2\rho-1)T\cdot(1-\min\{1,x\}), \end{aligned} \tag{12}$$

To check the inequality (12), note that $\sum_{n\in\mathbb{Z}} Q_n\cdot\mathbf{1}_{Q_{n-1}<T\le Q_n}$ is a random variable that represents the stopping bid of $\mathcal{A}$ when the target is at T. If $x < 1$, then the stopping bid is larger than or equal to $(2\rho-1)T$ with probability $\ge 1-x$. Therefore,

$$2(\rho-1)\min\{1,x\} \ge (\rho-1), \qquad x \ge \min\{1,x\} \ge \frac{1}{2}.$$

Overall, we have

$$\frac{1}{2} \le \mu([T,(2\rho-1)T)) \le \rho(2\rho-1),$$

where the upper bound implies the local finiteness while the lower bound implies $\mu((0,1)) = \mu([1,\infty)) = \infty$.

We now define $G : \mathbb{R} \to (0,\infty)$ as follows:

$$G(x) := \begin{cases} \sup\{0 < t < 1 :\ \mu([t,1)) \ge -x\}, & \text{if } x < 0; \\ \inf\{t \ge 1 :\ \mu([1,t]) \ge x\}, & \text{if } x \ge 0. \end{cases}$$

Since $\mu([t,1)) < \infty$ for all $0 < t < 1$ by the local finiteness of μ and $\lim_{t\downarrow 0} \mu([t,1)) = \infty$, $G(x)$ is well-defined, finite, and strictly positive for all $x < 0$. Similarly, since $\mu([1,t]) < \infty$ for all $t > 1$ and $\lim_{t\to\infty} \mu[1,t] = \infty$, the same properties hold for all $x \ge 0$.

Note that for every bounded interval $I \subseteq (0,\infty)$,

$$m(G^{-1}(I)) = \mu(I), \tag{13}$$

where m denotes the Lebesgue measure. In other words, μ is a pushforward measure of the Lebesgue measure by G.

We show that G satisfies all the conditions of a (ρ,χ)-bidding profile. G is non-decreasing by its definition. We may assume that G is left-continuous; otherwise, we can modify G to be left-continuous without changing the value of its integral. This is justified because G is monotone, implying it has only countably many discontinuities, hence the integration of G over the Lebesgue measure like (15) and (20) stays intact after the change. The offset condition is clear by the definition of G:

$$G(x) < 1 \quad (x < 0), \qquad G(x) \ge 1 \quad (x > 0).$$

Finally, we show that the strategy driven from G has no larger expected cost at any target, which naturally implies both robustness and consistency conditions. Let any target $T > 0$ be given. From the original strategy, let us decompose the bidding cost for T into

$$\begin{aligned} \mathsf{cost}_{\mathcal{A}}(T) &= \mathbb{E}\left[\sum_{n\in\mathbb{Z}} Q_n \mathbf{1}_{Q_{n-1}<T}\right] \\ &= \mathbb{E}\left[\sum_{n\in\mathbb{Z}} Q_n \mathbf{1}_{Q_n<T}\right] + \mathbb{E}\left[\sum_{n\in\mathbb{Z}} Q_n \mathbf{1}_{Q_{n-1}<T\le Q_n}\right]. \end{aligned} \tag{14}$$

The first term of (14) simplifies to

$$\begin{aligned} \mathbb{E}\left[\sum_{n\in\mathbb{Z}} Q_n \mathbf{1}_{Q_n<T}\right] &= \sum_{n\in\mathbb{Z}} \mathbb{E}\left[Q_n \mathbf{1}_{Q_n<T}\right] \\ &= \sum_{n\in\mathbb{Z}} \int_{(0,T)} t\, d\mu_n(t) \\ &= \int_{(0,T)} t\, d\mu(t) = \int_{-\infty}^{\tau(T)} G(t)\, dt, \end{aligned} \tag{15}$$

where $\tau(T) := \sup\{t \in \mathbb{R} : G(t) < T\}$ is defined as in Lemma 5. The last equality is due to the change of variable formula with (13), since $G^{-1}(0,T) = (-\infty,\tau(T))$ or $(-\infty,\tau(T)]$.

We now consider the second term of (14). Note that the second term is just an expectation of the stopping bid. Denote R as a random variable representing the stopping bid from $\mathcal{A}$, and R^G as that from a strategy $\mathcal{B}$ induced by G. Note that R^G has the same distribution as $G(\tau(T)+U)$, $U \sim \mathrm{Unif}(0,1]$. Defining $T^* := G(\tau(T)+1)$, we have

$$\begin{aligned} \mathbb{E}\left[\sum_{n\in\mathbb{Z}} Q_n \mathbf{1}_{Q_{n-1}<T\le Q_n}\right] &= \mathbb{E}\left[R\right] = \mathbb{E}\left[R \cdot (\mathbf{1}_{T\le R<T^*} + \mathbf{1}_{R=T^*} + \mathbf{1}_{R>T^*})\right] \\ &= \mathbb{E}\left[(R-T^*) \cdot (\mathbf{1}_{T\le R<T^*} + \mathbf{1}_{R=T^*} + \mathbf{1}_{R>T^*})\right] + T^*. \end{aligned} \tag{16}$$

To bound the first term, we note that for every interval $I \subseteq [T, T^*)$,

$$\Pr[R \in I] \leq \Pr\left[\exists n, Q_n \in I\right] \leq \sum_{n \in \mathbb{Z}} \Pr[Q_n \in I] = \mu(I) = \Pr[R^G \in I].$$

Since $x - T^* < 0$ for $T \leq x < T^*$, the above implies that

$$\mathbb{E}\left[(R - T^*)\mathbf{1}_{T \leq R < T^*}\right] \geq \mathbb{E}\left[(R^G - T^*)\mathbf{1}_{T \leq R^G < T^*}\right]. \tag{17}$$

Moreover,

$$\mathbb{E}\left[(R - T^*)\mathbf{1}_{R = T^*}\right] = 0 = \mathbb{E}\left[(R^G - T^*)\mathbf{1}_{R^G = T^*}\right], \tag{18}$$

$$\mathbb{E}\left[(R - T^*)\mathbf{1}_{R > T^*}\right] \geq 0 = \mathbb{E}\left[(R^G - T^*)\mathbf{1}_{R^G > T^*}\right]. \tag{19}$$

Combining (17), (18) and (19) to (16) yields

$$\begin{aligned}
\mathbb{E}[R] &= \mathbb{E}\left[(R - T^*)(\mathbf{1}_{T \leq R < T^*} + \mathbf{1}_{R = T^*} + \mathbf{1}_{R > T^*})\right] + T^* \\
&\geq \mathbb{E}\left[(R^G - T^*)(\mathbf{1}_{T \leq R^G < T^*} + \mathbf{1}_{R^G = T^*} + \mathbf{1}_{R^G > T^*})\right] + T^* \\
&= \mathbb{E}\left[R^G\right] = \int_0^1 G(\tau(T) + t)\, dt = \int_{\tau(T)}^{\tau(T)+1} G(t)\, dt.
\end{aligned} \tag{20}$$

Overall, applying (15) and (20) to (14), we have

$$\begin{aligned}
\mathsf{cost}_{\mathcal{A}}(T) &= \mathbb{E}\left[\sum_{n \in \mathbb{Z}} Q_n \mathbf{1}_{Q_n < T}\right] + \mathbb{E}\left[\sum_{n \in \mathbb{Z}} Q_n \mathbf{1}_{Q_{n-1} < T \leq Q_n}\right] \\
&\geq \int_{-\infty}^{\tau(T)} G(t)\, dt + \int_{\tau(T)}^{\tau(T)+1} G(t)\, dt \\
&= \int_{-\infty}^{\tau(T)+1} G(t)\, dt = \mathsf{cost}_{\mathcal{B}}(T).
\end{aligned}$$

Combining the above with (10) and (11), we have for any $\varepsilon > 0$:

$$\int_{-\infty}^{x+1} G(t)\, dt \leq \int_{-\infty}^{\tau(G(x)+\varepsilon)+1} G(t)\, dt = \mathsf{cost}_{\mathcal{B}}(G(x) + \varepsilon) \leq \mathsf{cost}_{\mathcal{A}}(G(x) + \varepsilon) \leq \rho(G(x) + \varepsilon).$$

Letting $\varepsilon \to 0$, we get that $\int_{-\infty}^{x+1} G(t)\mathrm{d}t \leq \rho G(x)$ for all x. For consistency, we use $\tau(1) = 0$:

$$\int_{-\infty}^{1} G(t)\, dt = \int_{-\infty}^{\tau(1)+1} G(t)\, dt = \mathsf{cost}_{\mathcal{B}}(1) \leq \mathsf{cost}_{\mathcal{A}}(1) \leq \chi.$$

Therefore, G is a (ρ, χ)-bidding profile, which completes the proof. □

E Additional details on linear search

E.1 Preliminaries on linear search

We continue with preliminary background on linear search.

As in online bidding, we introduce a strategy for linear search as a surrogate model for an algorithm. A deterministic strategy is formally defined as a bi-infinite sequence $y = \{y_n\}_{n \in \mathbb{Z}}$ of signed excursion endpoints such that the signs of consecutive elements are alternating (i.e., $\mathrm{sgn}(y_n) \neq \mathrm{sgn}(y_{n+1})$ for all $n \in \mathbb{Z}$), and the magnitudes of every other element are non-decreasing (i.e., $|y_n| \leq |y_{n+2}|$ for all $n \in \mathbb{Z}$). Given a deterministic strategy y and a target $T \in \mathbb{R} \setminus \{0\}$, the cost incurred by y given T is defined as

$$\mathsf{cost}_y(T) := |T| + 2\sum_{n < n_*} |y_n|,$$

where $n_* := \min\{n \in \mathbb{Z} : \mathrm{sgn}(y_n) = \mathrm{sgn}(T) \text{ and } |y_n| \geq |T|\}$. A randomized strategy Y is defined as a probability distribution over deterministic strategies, and the expected cost incurred by Y given T is denoted by $\mathsf{cost}_Y(T)$.

Note that, by adapting the first half of the argument in Section C, one can easily convert a ρ-robust χ-consistent strategy into a ρ-robust χ-consistent algorithm for the setting where $|T|$ is bounded below by some constant $c_0 > 0$. It is uncertain whether the second half of the argument also extends to linear search, but it is unnecessary in this paper as we present a lower bound using Yao's lemma instead in Section E.4.

As the universe of T is $\mathbb{R} \setminus \{0\}$, we can see that the problem is homogeneous under scaling and negating. We thus assume that the prediction is given as $P := +1$ by dividing the whole universe by P. Moreover, we can freely shift the indices of a strategy since it is bi-infinite. We may assign index 0 to the first excursion point where $+1 = P$ is passed through for the first time. Finally, let c denote the lower bound on the target magnitude after the normalization and negation.

A (randomized) strategy Y is called ρ-robust if $\mathsf{cost}_Y(T) \leq \rho\,|T|$ for any $T \in \mathbb{R} \setminus \{0\}$, and χ-consistent if $\mathsf{cost}_Y(+1) \leq \chi$. A pair (ρ, χ) is *Pareto-optimal* if there exists a ρ-robust χ-consistent strategy, but no ρ'-robust χ-consistent strategy for any $\rho' < \rho$.

E.2 Deferred proofs from Section 4.1

Lemma 14. *The strategy driven by a (ρ, χ)-excursion profile $G_\pm$ is $(1 + 2\rho)$-robust and $(1 + 2\chi)$-consistent.*

Proof. Since each unsuccessful excursion incurs twice its length, while the successful excursion incurs the distance to the target, we will see that a (ρ, χ)-excursion profile drives a $(1 + 2\rho)$-robust, $(1 + 2\chi)$-consistent linear search strategy.

For $T > 0$, define

$$\tau_+(T) := \sup\{t \in \mathbb{R} : G_+(t) < T\}, \text{ and}$$
$$\tau_-(T) := \sup\{t \in \mathbb{R} : G_-(t) < T\}.$$

Note that $\tau_\pm(T)$ is well-defined and finite, from the robustness condition. The proof is analogous to the online bidding case, as

$$G_+(\rho T) \geq \frac{1}{\rho}\int_{-\infty}^{\rho T} G_+(t)\mathrm{d}t + \frac{1}{\rho}\int_{-\infty}^{\rho T} G_-(t)\mathrm{d}t \geq \frac{1}{\rho}\int_0^{\rho T} 1\mathrm{d}t = T,$$

implying $\tau_+(T) \leq \rho T < \infty$ for any $T > 0$. Also,

$$G_-(\rho T - 1) \geq \frac{1}{\rho}\int_{-\infty}^{\rho T} G_+(t)\mathrm{d}t + \frac{1}{\rho}\int_{-\infty}^{\rho T-1} G_-(t)\mathrm{d}t \geq \frac{1}{\rho}\int_0^{\rho T} \mathrm{d}t = T,$$

implying $\tau_-(T) \leq \rho T - 1$ for any $T > 0$.

For a target at $+x$ $(x > 0)$, observe that if $G_+(n + U) < x$, cost of $2G_+(n + U) + 2G_-(n + U)$ is incurred. Otherwise, the strategy terminates after paying x. Therefore, the expected cost is at most

$$\begin{aligned}
&x + 2\sum_{n\in\mathbb{Z}}\int_0^1 (G_+(n+u) + G_-(n+u))\mathbf{1}_{G_+(n+u)<x}\mathrm{d}u \\
&= x + 2\int_{-\infty}^{\tau_+(x)} (G_+(v) + G_-(v))\mathrm{d}v \\
&= x + 2C_+(\tau_+(x)) \\
&\leq x + 2\rho G_+(\tau_+(x)) \leq (1 + 2\rho)x.
\end{aligned}$$

Symmetrically for a target at $-x$ $(x > 0)$, cost of $2G_-(n + U) + 2G_+(n + 1 + U)$ is incurred if $G_-(n + U) < x$. Otherwise, the strategy terminates after paying x. Then the entire cost is at most

$$\begin{aligned}
&x + 2\sum_{n\in\mathbb{Z}}\int_0^1 (G_-(n+u) + G_+(n+1+u))\mathbf{1}_{G_-(n+u)<x}\mathrm{d}u \\
&= x + 2\int_{-\infty}^{\tau_-(x)} (G_+(v+1) + G_-(v))\mathrm{d}v \\
&= x + 2C_-(\tau_-(x)) \\
&\leq x + 2\rho G_-(\tau_-(x)) \leq (1 + 2\rho)x.
\end{aligned}$$

Hence the strategy is $(1+2\rho)$-robust. From the plus-offset condition, it is clear that the strategy is $(1+2\chi)$-consistent. □

E.3 Deferred proofs in Section 4.2

Lemma 16. *Fix $\rho > 1/(2W_0(e^{-1}))$. Given a bounded and integrable function $\psi : (0,1] \to [1,\infty)$, there exist functions $G_+ : (-\infty, 1] \to [0,\infty)$ and $G_- : (-\infty, 0] \to [0,\infty)$ such that*

(i) $G_+ \equiv \psi$ on $(0,1]$.

(ii) For all $x \le 0$, we have

$$\int_{-\infty}^{x} G_+(t)\mathrm{d}t + \int_{-\infty}^{x} G_-(t)\mathrm{d}t = \rho G_+(x),$$
$$\int_{-\infty}^{x+1} G_+(t)\mathrm{d}t + \int_{-\infty}^{x} G_-(t)\mathrm{d}t = \rho G_-(x).$$

(iii) For any pair $G'_\pm$ with the same domains and codomains satisfying (i) and the inequalities obtained by replacing $=$ with $\le$ in (ii), we have $G_+(x) \le G'_+(x)$ and $G_-(x) \le G'_-(x)$ for all $x \le 0$.

If ρ is parameterized by $s \in (0, s_*)$ as $\rho = (1+e^s)/(2s)$, where $s_* := 1 + W_0(e^{-1})$, we have

$$\chi := \rho G_+(0) = \frac{1}{1+e^s} \int_0^1 \left(e^{2s(1-x)} - 1\right) \psi(x)\mathrm{d}x. \tag{4}$$

Proof. For a pair of nonnegative functions $H_\pm$ on $(-\infty, 0]$, let us define a functional operator $\mathcal{F}$ as follows:

$$(\mathcal{F}H)_+(x) := \frac{1}{\rho}\left(\int_{-\infty}^{x} H_+(t)\mathrm{d}t + \int_{-\infty}^{x} H_-(t)\mathrm{d}t\right),$$

and

$$(\mathcal{F}H)_-(x) := \frac{1}{\rho}\left(\int_{-\infty}^{\min\{x+1,0\}} H_+(t)\mathrm{d}t + \mathbf{1}_{\{x>-1\}} \int_0^{\max\{0,x+1\}} \psi(t)\mathrm{d}t + \int_{-\infty}^{x} H_-(t)\mathrm{d}t\right).$$

The operator $\mathcal{F}$ is order-preserving. Start from $G_+^{(0)} = G_-^{(0)} \equiv 0$ on $(-\infty, 0]$, and set $G^{(m+1)} = \mathcal{F}G^{(m)}$. Then

$$0 = G_\pm^{(0)} \le G_\pm^{(1)} \le G_\pm^{(2)} \le \cdots.$$

Here, $G_\pm^{(m)} \le G_\pm^{(m+1)}$ represents $G_+^{(m)} \le G_+^{(m+1)}$ and $G_-^{(m)} \le G_-^{(m+1)}$.

We next bound the iteration from above. Choose $c \in (s, s_*)$. Since $r \mapsto (1+e^r)/(2r)$ is strictly decreasing on $(0, s_*)$, we have

$$\frac{1+e^c}{2c} < \rho.$$

Equivalently, $2c\rho - 1 > e^c$. Put

$$\eta := \frac{e^{2c}}{2c\rho - 1}.$$

Then $0 < \eta \le 2c\rho - 1$. For a sufficiently large constant L, define

$$W_+(x) := Le^{2cx}, \qquad W_-(x) := L\eta e^{2cx}, \qquad x \le 0.$$

We claim that $\mathcal{F}W \le W$. For the plus coordinate,

$$(\mathcal{F}W)_+(x) = \frac{1+\eta}{2c\rho} Le^{2cx} \le W_+(x).$$

For the minus coordinate and $x \le -1$,

$$(\mathcal{F}W)_-(x) = \frac{e^{2c}+\eta}{2c\rho} Le^{2cx} = L\eta e^{2cx} = W_-(x).$$

For $-1 < x \le 0$, $(\mathcal{F}W)_-(x) \le W_-(x)$ is equivalent to

$$L\left(\rho\eta e^{2cx} - \frac{1}{2c} - \frac{\eta e^{2cx}}{2c}\right) \ge \int_0^{x+1} \psi(t)\mathrm{d}t.$$

Let $f(x)$ denote the expression in parenthesis. Since $f(-1) = 0$ and $f'(x) = e^{2c(x+1)} \ge 1$ for $-1 < x \le 0$,

$$f(x) = f(-1) + \int_{-1}^{x} f'(t)\mathrm{d}t \ge 0 + \int_{-1}^{x} 1\mathrm{d}t = x + 1,$$

on $[-1, 0]$. Taking $L \ge \sup_{0<t\le 1} \psi(t)$ proves the claim.

Thus $0 \le G_\pm^{(m)} \le W_\pm$ for all m. The monotone convergence theorem gives finite limits $G_\pm$, and the exponential bound gives finite lower-tail integrals. Passing to the limit in $G^{(m+1)} = \mathcal{F}G^{(m)}$ gives $G = \mathcal{F}G$, which is exactly (ii). Condition (iii) follows from the same induction. For any pair $G'_\pm$ as in (iii), we have $G_\pm^{(0)} \le G'_\pm$ and $(\mathcal{F}G')_\pm \le G'_\pm$. The order-preserving property of $\mathcal{F}$ then gives $G_\pm^{(m)} \le G'_\pm$ for every m. Taking the pointwise limit yields $G_\pm \le G'_\pm$. To obtain (i), simply set $G_+ \equiv \psi$ on $(0, 1]$.

It remains to prove (4). Define

$$A_\pm(x) := \int_{-\infty}^{x} G_\pm(t)\mathrm{d}t.$$

The tight equations give, almost everywhere on $(-\infty, 0]$,

$$\rho A'_+(x) = A_+(x) + A_-(x), \qquad \rho A'_-(x) = A_+(x+1) + A_-(x). \tag{21}$$

The exponential bound above implies when $T \to \infty$,

$$e^{2sT} A_\pm(-T) \to 0, \qquad \int_{-T}^{1-T} e^{-2su} A_+(u)\mathrm{d}u \to 0.$$

We multiply the two differential identities in (21) by e^{-2sx} and integrate over $[-T, 0]$. Define

$$X_\pm^T := \int_{-T}^{0} e^{-2sx} A_\pm(x)\mathrm{d}x, \quad Y_+^T := \int_{-T}^{0} e^{-2sx} A_+(x+1)\mathrm{d}x.$$

Then integrating by parts,

$$e^s X_+^T - X_-^T + \rho A_+(0) - \rho e^{2sT} A_+(-T) = 0, \tag{22}$$

and

$$e^s X_-^T - Y_+^T + \rho A_-(0) - \rho e^{2sT} A_-(-T) = 0. \tag{23}$$

Note that

$$Y_+^T = e^{2s} \int_{1-T}^{1} e^{-2sx} A_+(x)\mathrm{d}x = e^{2s}\left(X_+^T + \int_0^1 e^{-2sx} A_+(x)\mathrm{d}x - \int_{-T}^{1-T} e^{-2sx} A_+(x)\mathrm{d}x\right).$$

Evaluating e^s(22) + (23) at $T \to \infty$, the boundary terms vanish and we get

$$e^{2s} \int_0^1 e^{-2sx} A_+(x)\mathrm{d}x = \rho\left(e^s A_+(0) + A_-(0)\right).$$

Since $A_+(x) = A_+(0) + \int_0^x \psi(t)\mathrm{d}t$ on $[0, 1]$, one integration by parts gives

$$\int_0^1 e^{-2sx}\psi(x)\mathrm{d}x = e^{-2s} A_+(1) - A_+(0) + 2s \int_0^1 e^{-2sx} A_+(x)\mathrm{d}x.$$

Using $2s\rho = 1 + e^s$ and $A_+(1) = A_+(0) + \int_0^1 \psi(t)\mathrm{d}t$, this becomes

$$(1 + e^s)\left(A_+(0) + A_-(0)\right) = \int_0^1 \left(e^{2s(1-x)} - 1\right)\psi(x)\mathrm{d}x.$$

Since $A_+(0) + A_-(0) = \rho G_+(0)$ by (ii), this proves (4). □

Theorem 17. *Let $s_K \approx 0.5878$ denote the unique positive solution of $e^s(e^{2s}-1+2se^s) = (1+e^s)^2$, and define $K : (0, s_*] \to (0, e^{2s_*}]$ as*

$$K(s) := \begin{cases} \frac{e^s(e^{2s}-1+2se^s)}{(1+e^s)^2}, & \text{for } s \in (0, s_K]; \\ \xi(s), & \text{for } s \in (s_K, s_*], \end{cases}$$

where $\xi(s)$ is the unique solution $\xi \in [1, e^{2s}]$ of $(e^s - \xi)\ln\xi + \xi(3 + e^{-s}) = e^s(e^s + 2s - 1)$. Then, for every trade-off parameter $s \in (0, s_*]$, there exists a $(\rho(s), \chi(s))$-excursion profile with

$$\rho(s) = \frac{1+e^s}{2s}, \quad \chi(s) = \begin{cases} \frac{e^{2s}-1-2s}{2s(1+e^s)} = \rho K(s)e^{-s} - 1, & \text{for } s \in (0, s_K], \\ \frac{e^{2s}+1+(1+K(s))\log K(s) - 2K(s) - 2s}{2s(1+e^s)}, & \text{for } s \in (s_K, s_*]. \end{cases}$$

Before the proof, we first show that $K(s)$ is well-defined and increases with s. If $s \le s_K$, we have

$$K'(s) = \frac{e^s(e^s(4s+3) + 5e^{2s} + e^{3s} - 1)}{(1+e^s)^3} > 0. \tag{24}$$

This implies that $K(s) \le K(s_K) = 1$ and $K(s)$ strictly increases when $s > 0$.

Remark 24. For $s_K < s \le s_*$, the defining equation

$$(e^s - \xi)\ln\xi + \xi(3 + e^{-s}) = e^s(e^s + 2s - 1)$$

has a unique solution $\xi \in [1, e^{2s}]$.

Proof. Define

$$F_s(\xi) := (e^s - \xi)\ln\xi + \xi(3 + e^{-s}) - e^s(e^s + 2s - 1).$$

Then

$$e^s F_s(1) = 3e^s + 1 - e^{2s}(e^s + 2s - 1) = (1+e^s)^2 - e^s(e^{2s} + 2se^s - 1),$$

which is negative for $s > s_K$, considering (24).

Moreover,

$$F_s(e^{2s}) = 2s(e^s - e^{2s}) + 3e^{2s} + e^s - e^s(e^s + 2s - 1) = 2e^s(1 + (1-s)e^s).$$

Since $1 + (1 - s_*)e^{s_*} = 0$ from the definition of $s_* = 1 + W_0(e^{-1})$, and $x \mapsto 1 + (1-x)e^x$ decreases for $x > 0$, it implies that $F_s(e^{2s}) \ge 0$.

Now, we show that F_s is strictly increasing.

$$F_s'(\xi) = \frac{e^s}{\xi} - \ln\xi + 2 + e^{-s},$$

and

$$F_s''(\xi) = -\frac{1}{\xi} - \frac{e^s}{\xi^2} < 0.$$

Since F_s' strictly decreases, it suffices to show that $F_s'(e^{2s}) \ge 0$.

$$F_s'(e^{2s}) = e^{-s} - 2s + 2 + e^{-s} = 2e^{-s}(1 + (1-s)e^s) \ge 0,$$

as we proved before. □

Proof of Theorem 17. The endpoint $s = s_*$ is the classical exponential profile, so assume $s < s_*$. Write $\rho = \rho(s)$, $K = K(s)$ and $M = \max\{1, K\}$. On $x > 0$, define

$$G_+(x) := \max\{1, Me^{2s(x-1)}\},$$

and

$$G_-(x) := Me^s e^{2s(x-1)} + (K - M)e^{-s}e^{x/\rho}.$$

The correction term in G_- appears only when $K < 1$, and it ensures that $G_-(0+) = Ke^{-s}$.

Let $\psi = G_+|_{(0,1]}$, which is bounded and integrable. Since $\rho > 1/(2W_0(e^{-1}))$, we define $G_\pm$ on $x \le 0$ using the pair of functions given by Lemma 16 for this ψ. As in the proof of Lemma 16, define $A_\pm(x) := \int_{-\infty}^{x} G_\pm(t)\mathrm{d}t$ for $x \in \mathbb{R}$. By (4), we have

$$\chi_s := \rho G_+(0) = \frac{1}{1+e^s}\int_0^1 \left(e^{2s(1-x)} - 1\right) G_+(x)\mathrm{d}x.$$

If $K < 1$, then $G_+(x) = 1$ on $(0, 1]$, so this is the first expression in the statement. If $K \ge 1$, set $x_K = 1 - \log K/(2s)$. Splitting the integral at x_K gives the second expression.

We next record the boundary identity

$$\chi_s + \int_0^1 G_+(x)\mathrm{d}x = \rho K e^{-s}. \tag{25}$$

When $K < 1$, this is exactly the definition of $K(s)$. When $K \ge 1$, we use

$$\int_0^1 G_+(x)\mathrm{d}x = 1 + \frac{K - 1 - \log K}{2s},$$

and the formula for χ_s; after multiplying by $2s(1 + e^s)$, the identity is equivalent to

$$(e^s - K)\log K + K(3 + e^{-s}) - e^s(e^s + 2s - 1) = 0.$$

Thus (25) holds in both cases. Consequently, by (ii) of Lemma 16 at $x = 0$, we have

$$\begin{aligned}\rho G_-(0) &= A_+(1) + A_-(0)\\ &= \chi_s + A_+(1) - A_+(0) = \rho K e^{-s} = \rho G_-(0+).\end{aligned}$$

We now verify robustness for $x > 0$. The negative constraint is tight. Indeed, the equation above gives $A_+(1) + A_-(0) = \rho G_-(0+)$, and direct differentiation gives

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(A_+(x+1) + A_-(x)\right) = G_+(x+1) + G_-(x) = \rho G'_-(x).$$

Hence $A_+(x + 1) + A_-(x) = \rho G_-(x)$ for every $x > 0$.

It remains to check $A_+(x) + A_-(x) \le \rho G_+(x)$ for $x > 0$. First suppose $K \ge 1$. Then the correction term in G_- vanishes. At $x = 1$, the boundary identity gives

$$\begin{aligned}A_+(1) + A_-(1) &= A_+(1) + A_-(0) + A_-(1) - A_-(0)\\ &= \rho K e^{-s} + \int_0^1 G_-(x)\mathrm{d}x\\ &= \rho K e^{-s} + \int_0^1 e^s K e^{2s(x-1)}\mathrm{d}x\\ &= \rho K = \rho G_+(1).\end{aligned}$$

When $G_+(x) > 1$, we have $A'_+(x) + A'_-(x) = G_+(x) + G_-(x) = \rho G'_+(x)$. Together with equality at $x = 1$, this gives equality throughout the exponential part and, by continuity, at the right endpoint of the plateau. Otherwise, the function $A_+(x) + A_-(x)$ is nondecreasing and reaches value ρ at the right endpoint of the plateau. Therefore $A_+(x) + A_-(x) \le \rho G_+(x)$ for all $x > 0$.

Now suppose $K < 1$. Then $G_+(x) = 1$ on $(0, 1]$ and $G_+(x) = e^{2s(x-1)}$ for $x \ge 1$. Also

$$G_-(x) = e^s e^{2s(x-1)} + (K - 1)e^{-s}e^{x/\rho}.$$

Since $1/\rho < 2s$, this function is positive and nondecreasing on $(0, \infty)$. Moreover,

$$A_+(1) + A_-(1) = \rho K e^{-s} + \int_0^1 G_-(x)\mathrm{d}x \le \rho e^{-s} + \int_0^1 e^s e^{2s(x-1)}\mathrm{d}x = \rho.$$

Thus $A_+(x) + A_-(x) \le \rho = \rho G_+(x)$ on $0 < x \le 1$. For $x > 1$,

$$\rho G'_+(x) - A'_+(x) - A'_-(x) = \rho G'_+(x) - G_+(x) - G_-(x) = (1 - K)e^{-s}e^{x/\rho} \ge 0.$$

Since the inequality holds at $x = 1$, it propagates to all $x \geq 1$.

The left half-line is tight by construction, and the positive half-line was checked above. Hence the robustness condition holds everywhere. Finally, by (ii) of Lemma 16 at $x = 0$, we have $A_+(0) + A_-(0) = \rho G_+(0) = \chi_s$, which gives the consistency condition. Also $\chi_s \leq \rho$: in the case $K < 1$ this follows from the boundary identity, while in the case $K \geq 1$ it follows from $K \leq e^{2s}$, which gives $\log K + Ke^{-s} \leq 2s + e^s$. Therefore $G_+(0) = \chi_s/\rho \leq 1$.

On $(-\infty, 0]$, the integral equations in (ii) of Lemma 16 make both functions continuous and non-decreasing. Since $\psi \geq 1$, the second equation of (ii) gives $G_- > 0$ on $(-1, 0]$, and the first equation of (ii) then gives $G_+ > 0$ on the same interval. Repeating this argument on successive unit intervals to the left shows that both functions are positive on $(-\infty, 0]$. The first equation of (ii) then makes G_+ strictly increasing there, so $G_+(x) < 1$ for $x < 0$. Since $G_+(x) \geq 1$ for $x > 0$, the plus-offset condition holds. The functions are non-decreasing across 0 because $G_+(0) \leq 1 \leq G_+(0+)$ and $G_-(0) = G_-(0+)$. Thus $G_\pm$ is a (ρ, χ_s)-excursion profile. □

Remark 18. The profile of Theorem 17 yields a ρ_{LS}-robust χ_{LS}-consistent strategy with $\rho_{\mathrm{LS}} = \frac{2}{\chi_{\mathrm{LS}}-1} + \frac{7}{3} + O(\chi_{\mathrm{LS}} - 1)$. In Appendix E.4, we present a lower bound of $\rho_{\mathrm{LS}} = \frac{2}{\chi_{\mathrm{LS}}-1} + \frac{3}{2} + \Omega(\chi_{\mathrm{LS}} - 1)$, demonstrating our strategy is near-optimal up to sub-leading constant. Both trade-off bounds are depicted in Figure 1b.

Proof. We only provide the asymptotics of the upper bound from Theorem 17. When $s < s_K$,

$$\chi_{\mathrm{LS}} - 1 = 2\chi(s) = \frac{e^{2s} - 1 - 2s}{s(1 + e^s)} = s + \frac{s^2}{6} + O(s^4).$$

On the other hand,

$$\rho_{\mathrm{LS}} = 1 + \frac{1 + e^s}{s} = \frac{2}{s} + 2 + \frac{s}{2} + \frac{s^2}{6} + O(s^3),$$

implying that

$$\rho_{\mathrm{LS}} - \frac{2}{\chi_{\mathrm{LS}} - 1} = \frac{7}{3} + O(s) = \frac{7}{3} + O(\chi_{\mathrm{LS}} - 1).$$

□

E.4 Lower bound on robustness-consistency trade-off for learning-augmented linear search

In this section, we provide a lower bound on the robustness-consistency trade-off of algorithms for the linear search problem. The proof employs Yao's lemma, which provides a lower bound of the performance for every randomized algorithm by an optimal performance of the deterministic algorithm for the selected input distribution.

Analogous to the lower bound construction for the online bidding problem from [SLLA25], we select our input distribution to be proportional to the distance from the origin, but scaled differently by its direction and truncated by a closed interval $[-\alpha_2, +\alpha_1]$. By choosing the scaling factor and the interval appropriately, we prove the following theorem:

Theorem 25. *For $t \in (0, 1]$, any (randomized) $\left(1 + 2\frac{t(t+2)^2}{-t^3+3t+4}\right)$-consistent linear search algorithm has the robustness ratio of at least*

$$1 + 4\frac{(t^2 + t + 1)}{t(-t^3 + 3t + 4)}.$$

Asymptotically, any χ_{LS}-consistent ρ_{LS}-robust algorithm must satisfy

$$\rho_{\mathrm{LS}} \geq \frac{2}{\chi_{\mathrm{LS}} - 1} + \frac{3}{2} + \Theta(\chi_{\mathrm{LS}} - 1)$$

as $\chi_{\mathrm{LS}} \to 1^+$.

Proof. Given a deterministic algorithm, we encode its excursions by a sign-alternating sequence $\{y_i\}_{i=1}^{\infty}$. Recall a randomized algorithm is a probability distribution over deterministic algorithms.

Again, we assume without loss of generality that the subsequence of the same parity has a strictly increasing property in magnitude.

Having these properties of excursions in mind, we capture the *state* of the algorithm at a moment by a pair $\mathbf{u} \in \mathbb{R}_{\geq 0}^{\{+,-\}}$ of non-negative numbers, where each entry $\mathbf{u}_\pm$ indicates the maximum distance traveled along the half-line of corresponding sign so far.

The state is updated every excursion starting from $\mathbf{u}^0 := (0,0)$. For each $j \in \mathbb{Z}_{\geq 1}$, denote the state after j-th excursion as $\mathbf{u}^j$. If the excursion point is $+y_j > 0$, $\mathbf{u}^j = (y_j, \mathbf{u}_-^{j-1})$. Otherwise, $\mathbf{u}^j = (\mathbf{u}_+^{j-1}, -y_j)$. Exceptionally, when the algorithm terminates at the target T at m-th excursion, the state stops updating with $\mathbf{u}^m_{\mathrm{sgn}(T)} = |T|$.

We apply Yao's lemma on robustness with consistency terms:

$$\begin{aligned}
&\sup_{\mu\in\mathcal{M}} \inf_{y\in\mathcal{D}} \left(\mathbb{E}_{T\sim\mu}\left[\frac{\mathsf{cost}_y(T)}{|T|}\right] + \lambda \cdot \mathsf{cost}_y(+1)\right) && (26)\\
&= \sup_{\mu\in\mathcal{M}} \inf_{Y\in\mathcal{A}} \left(\mathbb{E}_{T\sim\mu}\left[\frac{\mathsf{cost}_Y(T)}{|T|}\right] + \lambda \cdot \mathsf{cost}_Y(+1)\right)\\
&\leq \inf_{Y\in\mathcal{A}} \sup_{\mu\in\mathcal{M}} \left(\mathbb{E}_{T\sim\mu}\left[\frac{\mathsf{cost}_Y(T)}{|T|}\right] + \lambda \cdot \mathsf{cost}_Y(+1)\right)\\
&\leq \rho_{\mathrm{LS}} + \lambda \cdot \chi_{\mathrm{LS}}, && (27)
\end{aligned}$$

where $\lambda \geq 0$ is any given multiplier, $\mathcal{M}$ is a set of all target distributions supported on $(-\infty, -c] \cup [c, \infty)$, $\mathcal{D}$ is a set of all deterministic algorithms while $\mathcal{A}$ is a set of all randomized algorithms. (Recall c is the lower bound on the target magnitude.) By dividing the whole inequality by $1+\lambda$, note that the inequality can be fully understood as a variant of Yao's lemma, whose distribution set is conditioned to have an atom at $+1$ with size of at least $\frac{\lambda}{1+\lambda}$.

As a lower bound of (26), we use a distribution μ whose density is defined as follows:

$$\frac{\mathrm{d}\mu}{\mathrm{d}m}(T) = \frac{\mathbf{1}\left[-\alpha_2 \leq T \leq -c\right] + w\mathbf{1}\left[c \leq T \leq \alpha_1\right]}{\frac{\alpha_2^2 - c^2}{2} + \frac{w(\alpha_1^2 - c^2)}{2}} |T|, \tag{28}$$

where

$$w = \frac{2t+1}{t(t+2)}, \quad \alpha_1 = \frac{t^2(t+2)}{t^2+t+1}, \quad \alpha_2 = \frac{t(t+2)}{t^2+t+1} \tag{29}$$

for $t \in (0, 1]$. Note that $\alpha_1 \in [0, 1]$ and $\alpha_2 > 0$ if $t \in (0, 1]$, since $t^2(t+2), t(t+2), t^2+t+1 > 0$ and

$$t^2 + t + 1 - t^2(t+2) = (t+1)(1-t^2) \geq 0.$$

Note that we may further assume that $c < \min\{\alpha_1, \alpha_2\}$ by considering instances in which the magnitude of the (original) prediction is sufficiently large. We also use the following multiplier:

$$\lambda = \frac{1-t^2}{t^2(t+2)}. \tag{30}$$

(29) and (30) are selected appropriately to provide the best lower bound curve conditioned on the distribution μ having the form of (28).

Let any deterministic algorithm y be given, and assume without loss of generality that every excursion has magnitude greater than c. Since the support of the distribution of a target and the prediction point is contained in $[-\alpha_2, +1]$, we can also assume that the final state of y is $\mathbf{f} := (1, \alpha_2) \in \mathbb{R}_{\geq 0}^{\{+,-\}}$. Regarding y as a single state sequence $\{\mathbf{u}^j\}_{j=0}^m$, which defines a non-decreasing curve from $\mathbf{u}^0 = (0,0)$ to $\mathbf{u}^m = \mathbf{f}$ whose directions should be parallel to either of two axes, the cost of reaching any target $T \in [-\alpha_2, +1]$ can be written as follows:

$$\mathsf{cost}_y(T) = 2\sum_{j=1}^{m} \mathbf{u}^j_{\mathrm{dir}(j)} \cdot \mathbf{1}_{\mathbf{u}^j_{\mathrm{sgn}(T)} < |T|} + |T|, \tag{31}$$

where $\mathrm{dir}(j) \in \{+,-\}$ is the direction taken from j-th excursion. Therefore, the expected (competitive) ratio under μ is as follows.

$$\begin{aligned}
&\mathbb{E}_{T\sim\mu}\left[\frac{\mathsf{cost}_y(T)}{|T|}\right] \\
&= \int_{|T|\geq c} \frac{\mathsf{cost}_y(T)}{|T|}\frac{\mathrm{d}\mu}{\mathrm{d}m}(T)\mathrm{d}T \\
&= \frac{w\int_c^{\alpha_1}\mathsf{cost}_y(T)\mathrm{d}T + \int_c^{\alpha_2}\mathsf{cost}_y(-T)\mathrm{d}T}{\frac{\alpha_2^2-c^2}{2}+\frac{w(\alpha_1^2-c^2)}{2}} \\
&= 1 + \frac{2\left(w\int_c^{\alpha_1}\sum_{j=1}^m \mathbf{u}^j_{\mathrm{dir}(j)}\cdot\mathbf{1}\left[\mathbf{u}^j_+ < T\right]\mathrm{d}T + \int_c^{\alpha_2}\sum_{j=1}^m \mathbf{u}^j_{\mathrm{dir}(j)}\cdot\mathbf{1}\left[\mathbf{u}^j_- < T\right]\mathrm{d}T\right)}{\frac{\alpha_2^2-c^2}{2}+\frac{w(\alpha_1^2-c^2)}{2}}. \qquad (32)
\end{aligned}$$

The expression in parentheses in the numerator in (32) is equal to

$$\begin{aligned}
&\sum_{j=1}^m \mathbf{u}^j_{\mathrm{dir}(j)}\left(w\int_c^{\alpha_1}\mathbf{1}\left[\mathbf{u}^j_+ < T\right]\mathrm{d}T + \int_c^{\alpha_2}\mathbf{1}\left[\mathbf{u}^j_- < T\right]\mathrm{d}T\right) \\
&= \sum_{j=1}^m \mathbf{u}^j_{\mathrm{dir}(j)}\left(w\int_{\min\{\alpha_1,\max\{c,\mathbf{u}^j_+\}\}}^{\alpha_1}\mathrm{d}T + \int_{\min\{\alpha_2,\max\{c,\mathbf{u}^j_-\}\}}^{\alpha_2}\mathrm{d}T\right) \\
&= \sum_{j=1}^m \mathbf{u}^j_{\mathrm{dir}(j)}\left(w\max\left\{\alpha_1-\max\{c,\mathbf{u}^j_+\},0\right\} + \max\{\alpha_2-\max\{c,\mathbf{u}^j_-\},0\}\right) \\
&= \sum_{j=1}^m \mathbf{u}^j_{\mathrm{dir}(j)}\left(w\max\left\{\alpha_1-\max\{c,\mathbf{u}^j_+\},0\right\} + \alpha_2-\max\{c,\mathbf{u}^j_-\}\right),
\end{aligned}$$

where the last equality follows from the final state $(1,\alpha_2)$. Substituting into (32),

$$\begin{aligned}
&\mathbb{E}_{T\sim\mu}\left[\frac{\mathsf{cost}_y(T)}{|T|}\right] \\
&= 1 + \frac{2\left(\sum_{j=1}^m \mathbf{u}^j_{\mathrm{dir}(j)}\left(w\max\left\{\alpha_1-\max\{c,\mathbf{u}^j_+\},0\right\} + \alpha_2-\max\{c,\mathbf{u}^j_-\}\right)\right)}{\frac{\alpha_2^2-c^2}{2}+\frac{w(\alpha_1^2-c^2)}{2}}. \qquad (33)
\end{aligned}$$

On the other hand, the cost of reaching $+1$ is as follows.

$$\mathsf{cost}_y(+1) = 1 + 2\sum_{j=1}^m \mathbf{u}^j_{\mathrm{dir}(j)}\cdot\mathbf{1}_{\mathbf{u}^j_+<1}. \qquad (34)$$

Next, we derive the conditions for y that minimize the inner part of (26). Suppose $x := y_j$ for some $j \in \mathbb{Z}_{\geq 1}$. Note that $x \in [-\alpha_2,-c)\cup(c,+1]$. We show that x is degenerated to specific points to minimize the cost. For this purpose, assume that $x \in (-\alpha_2,-c)\cup(c,\alpha_1)\cup(\alpha_1,1)$.

If $x \in (-\alpha_2,-c)$, the consistency term (34) is linear in $|x|$, while the robustness term (33) is negative quadratic in $|x|$. Therefore, the total cost is concave in $|x|$ on $\left[\max\{c,\mathbf{u}^{j-1}_-\},\mathbf{u}^{j+2}_-\right]$. Hence, we can change x to one of $-\max\{c,\mathbf{u}^{j-1}_-\}$ and $-\mathbf{u}^{j+2}_-$ without increasing the total cost. If $x \in (+\alpha_1,+1)$, then changing every such x to $+\alpha_1$ only reduces both the robustness and consistency. Otherwise, $x \in (+c,+\alpha_1)$. Analogously the concavity of the cost with respect to $|x|$ on $\left[\max\{c,\mathbf{u}^{j-1}_+\},\min\{\alpha_1,\mathbf{u}^{j+2}_+\}\right]$ implies that we can change x to one of the endpoints without

increasing the total cost. Finally, removing an excursion of magnitude c or a repeated excursion to the same endpoint does not increase the objective.

Removing degeneracies, the excursion points of the optimal deterministic algorithm must belong to $\{+1, +\alpha_1, -\alpha_2\}$. This limits the candidate sequences to

- $0 \to +1 \to -\alpha_2$,
- $0 \to -\alpha_2 \to +1$,
- $0 \to +\alpha_1 \to -\alpha_2 \to +1$.

For each of the three candidates, the objective value converges as $c \downarrow 0$ to the value obtained by setting $c = 0$ in (33) and (34). Since there are only three candidates, the minimum of their objective values converges to the minimum of their limits. Note that, since $\alpha_1 = t\alpha_2$, we have

$$\frac{w\alpha_1^2}{2} + \frac{\alpha_2^2}{2} = \left(wt^2 + 1\right)\frac{\alpha_2^2}{2} = \alpha_1.$$

The three candidates give the following values as $c \downarrow 0$.

Excursion sequence y	(33)	(34)
$0 \to +1 \to -\alpha_2$	$1 + \frac{2\alpha_2}{\alpha_1}$	1
$0 \to -\alpha_2 \to +1$	$1 + 2w\alpha_2$	$1 + 2\alpha_2$
$0 \to +\alpha_1 \to -\alpha_2 \to +1$	$1 + 2\alpha_2$	$1 + 2(\alpha_1 + \alpha_2) \cdot \mathbf{1}\left[\alpha_1 < 1\right]$

Note that $\lambda \cdot \mathbf{1}\left[\alpha_1 < 1\right] = \lambda$ since $\alpha_1 = 1$ implies $\lambda = 0$.

Combining (26), (27) with these candidate values as $c \downarrow 0$, we obtain

$$\rho_{\mathrm{LS}} + \lambda \cdot \chi_{\mathrm{LS}} \geq 1 + \lambda + 2\alpha_2 \min\left\{\frac{1}{\alpha_1}, w + \lambda, 1 + \left(\frac{\alpha_1}{\alpha_2} + 1\right)\lambda\right\}. \tag{35}$$

Plugging (29) and (30) to (35) yields

$$\rho_{\mathrm{LS}} + \lambda \cdot \chi_{\mathrm{LS}} \geq 1 + \lambda + 2\min\left\{\frac{1}{t}, \frac{1}{t}, \frac{1}{t}\right\} = 1 + \lambda + \frac{2}{t},$$

thus

$$t^2(t+2)\frac{\rho_{\mathrm{LS}} - 1}{2} + (1 - t^2)\frac{\chi_{\mathrm{LS}} - 1}{2} \geq t(t+2).$$

Therefore, if $\chi_{\mathrm{LS}} = 1 + 2\frac{t(t+2)^2}{-t^3+3t+4}$, then

$$t^2(t+2)\frac{\rho_{\mathrm{LS}} - 1}{2} + (1 - t^2) \cdot \frac{t(t+2)^2}{-t^3 + 3t + 4} \geq t(t+2),$$

thus

$$\frac{\rho_{\mathrm{LS}} - 1}{2} \geq \frac{1}{t} \cdot \left(1 - (1 - t^2) \cdot \frac{t+2}{-t^3 + 3t + 4}\right) = \frac{1}{t} \cdot \frac{2(t^2 + t + 1)}{-t^3 + 3t + 4},$$

which implies

$$\rho_{\mathrm{LS}} \geq 1 + 4\frac{(t^2 + t + 1)}{t(-t^3 + 3t + 4)}.$$

Now we verify the claimed asymptotics. We first subtract the lower bound obtained by $\frac{2}{\chi_{\mathrm{LS}} - 1}$, which is

$$\begin{aligned}
1 + 4\frac{(t^2 + t + 1)}{t(-t^3 + 3t + 4)} - \frac{-t^3 + 3t + 4}{t(t+2)^2} &= 1 + \frac{-t^6 + 10t^4 + 28t^3 + 27t^2 + 8t}{t(-t^3 + 3t + 4)(t+2)^2} \\
&= 1 + \frac{-t^5 + 10t^3 + 28t^2 + 27t + 8}{(-t^3 + 3t + 4)(t+2)^2} \\
&= \frac{3}{2} + \frac{13}{16}t + O(t^2) \qquad \text{as } t \downarrow 0.
\end{aligned}$$

Since $\chi_{\mathrm{LS}} - 1 = 2t + O(t^2)$, we obtain

$$\rho_{\mathrm{LS}} - \frac{2}{\chi_{\mathrm{LS}} - 1} \geq \frac{3}{2} + \frac{13}{32}(\chi_{\mathrm{LS}} - 1) + O\left((\chi_{\mathrm{LS}} - 1)^2\right),$$

which concludes the proof. □

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 abstract and introduction accurately reflect the paper's 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: We discuss the limitations and future directions of this work in the conclusion.

   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 the full details of our model and mathematically rigorous analysis in the paper.

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: The paper only includes theoretical results.

   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: The paper only includes theoretical results.

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: The paper only includes theoretical results.

   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: The paper only includes theoretical results.

   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: The paper only includes theoretical results.

   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: The paper conforms with the NeurIPS 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: As the paper studies the theory of learning-augmented online bidding, we do not anticipate any (negative) social impacts arising from this work.

    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: We do not anticipate any misuse of the results presented in this work.

    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: The paper does not use any existing assets.

    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: The paper does not release new assets.

    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: The paper does not involve crowdsourcing nor research with human subjects.

    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: The paper does not involve crowdsourcing nor research with human subjects.

    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: [Yes]

Justification: We clearly acknowledge the extent to which LLMs were used in this work.

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.