The Pareto Frontier of Randomized Learning-Augmented Online Bidding

Mathis Degryse
Sorbonne University, CNRS, LIP6, Paris, France

Imrane Zaakour
Centrale Supélec, Paris, France
and ILLS, Montreal, Canada

Christoph Dürr
Sorbonne University, CNRS, LIP6, Paris, France

Spyros Angelopoulos
CNRS and ILLS, Montreal, Canada

Abstract

Online bidding is a classical problem in online decision-making, with applications in resource allocation, hierarchical clustering, and the analysis of approximation algorithms. We study its randomized learning-augmented variant, where an online algorithm generates a sequence of random bids while leveraging predictions from an oracle. We provide analytical upper and lower bounds on the optimal consistency C as a function of the robustness R, which match when $R \geq 2.885$, effectively closing the gap left by previous work. The key technical ingredient is the notion of a *bidding function*, a novel abstraction that provides a unified framework for the design and analysis of randomized bidding strategies. We complement our theoretical results with an experimental application of randomized bidding to the incremental median problem, demonstrating the applicability of our algorithm in practical clustering settings.

1 Introduction

In several sequential decision-making settings, the algorithm relies on increasing a quantity until it reaches a level that guarantees success, without knowing that threshold in advance. For instance, a scheduler may gradually increase the runtime allocated to a job until it finishes, or an algorithm may successively guess larger values of an unknown optimum until feasibility is reached. In such settings, all unsuccessful choices contribute to the total cost, hence the performance of the algorithm depends on how quickly it reaches the threshold without excessive overshooting.

Online bidding [21] provides a simple abstraction of this fundamental tradeoff. In this problem, a player submits an increasing sequence of *bids* until one of them reaches or exceeds an unknown threshold u, and pays the sum of all bids made up to that point. The objective is to design a bidding strategy whose total cost is as small as possible relative to the unknown threshold. Formally, in the deterministic setting, the player produces an increasing sequence[1] $X = (x_i)_{i \in \mathbb{Z}}$, and given a threshold u it incurs cost equal to

$$\mathrm{cost}(X, u) = \sum_{j \leq i} x_j \text{ for } i \text{ such that } x_{i-1} < u \leq x_i.$$

The *normalized* cost of X is then defined as $\mathrm{nc}(X, u) = \mathrm{cost}(X, u)/u$. For simplicity, we will use the notation $\mathrm{nc}(u)$ when X is implied from context. In the *randomized* setting, X is an increasing sequence of random bids, and the cost of X on a threshold u is measured by its expected value.

[1] As in previous work on online bidding and related incremental allocation problems, we consider *bi-infinite* sequences. This convention simplifies exposition and avoids additional technical assumptions, such as requiring that u be at least as large as the first bid.

Under the standard framework of *competitive analysis* [17], the objective is to minimize the so called *competitive ratio*, which is defined as the worst-case normalized cost, i.e., when the threshold is chosen adversarially. In practice, however, the algorithm may have access to additional information in the form of a *prediction*, obtained from historical data or learned models. This has motivated the study of *learning-augmented* algorithms [36], where a prediction oracle is incorporated into the design and analysis of the algorithm. The goal is to leverage such predictions to obtain significantly improved performance when they are accurate, while still retaining strong worst-case guarantees when they are not.

We consider a natural oracle which provides the algorithm with a prediction $\hat{u}$. In this context, the *consistency* and the *robustness* of a bidding strategy are defined as its worst-case normalized cost, assuming error-free or adversarial predictions, respectively. Formally

$$\mathrm{cons}(X) = \mathrm{nc}(X, \hat{u}), \quad \text{and} \quad \mathrm{rob}(X) = \sup_u \mathrm{nc}(X, u). \tag{1}$$

A natural goal is to design algorithms that achieve the optimal tradeoff between consistency and robustness; in the learning-augmented literature, such algorithms are often called *Pareto-optimal*. While this question has been answered for deterministic online bidding [5], it remains open in the randomized setting. Randomization is known to improve performance even in the classical setting without predictions: the optimal randomized competitive ratio is e [21], compared to the optimal deterministic ratio of 4 [14]. This naturally raises the question of how much randomization can improve learning-augmented bidding. In this work, we quantify this improvement by characterizing the best-possible tradeoff achievable by randomized algorithms.

1.1 Contributions

We present and analyze an algorithm that is R-robust for all $R \geq e$, and provably attains the optimal consistency[2] for all $R \geq 2/\ln 2 \approx 2.885$. For the remaining, and very narrow regime $R \in [e, 2.885]$, we derive analytical upper and lower bounds which significantly improve upon the state of the art of [7, 40]. In particular, for $R = e+\varepsilon$ and $\varepsilon \to 0$, the consistency of our algorithm is $C = e-\Omega(\varepsilon^{1/4})$, improving upon the previously known guarantee of $C = e - \Omega(\varepsilon)$. Our lower bound, which is tight for $R \geq 2.885$, also yields improved impossibility guarantees throughout the regime $R \in [e, 2.885]$. Figure 1 illustrates the obtained results and their comparison to previous work.

We complement our theoretical analysis with a numerical evaluation that explicitly incorporates prediction *error*. This study shows that our algorithm achieves improved *smoothness*, namely a more graceful degradation in performance as prediction quality deteriorates, compared to previous approaches. Finally, we present an experimental application to the classical *incremental median* problem, demonstrating that randomized online bidding can yield significant practical improvements in clustering settings, even under highly inaccurate predictions.

From a technical perspective, our design and analysis rely on a new abstraction that we call a *bidding function*. This framework shifts the focus from discrete, randomized sequences to the analysis of continuous functions, providing substantially greater flexibility in both algorithm design and lower-bound arguments. In particular, it enables a more systematic treatment of consistency–robustness tradeoffs than previous sequence-based approaches. A useful parallel can be drawn with online *conversion* problems, where *threshold functions* have played a central role in the development of improved algorithms, and in unifying the analysis of several online problems [18, 41].

1.2 Related Work

Learning-augmented algorithms Algorithms augmented with a machine-learned prediction on some unknown aspect of the input have been studied in a large variety of online problems, such as rent-and-buy problems [26], scheduling [29], caching [34], matching [8], packing [27], covering [13], sorting [12], graph problems [11], optimal stopping [24] and data structures [32]. These are only some representative works, and since the seminal works [34, 38] the field has experienced a remarkable growth; we refer to the online compendium [33] for a listing of several recent works.

[2] During the preparation of this manuscript, we became aware of concurrent independent work by C. Lee et al. [30] that obtains a similar result.

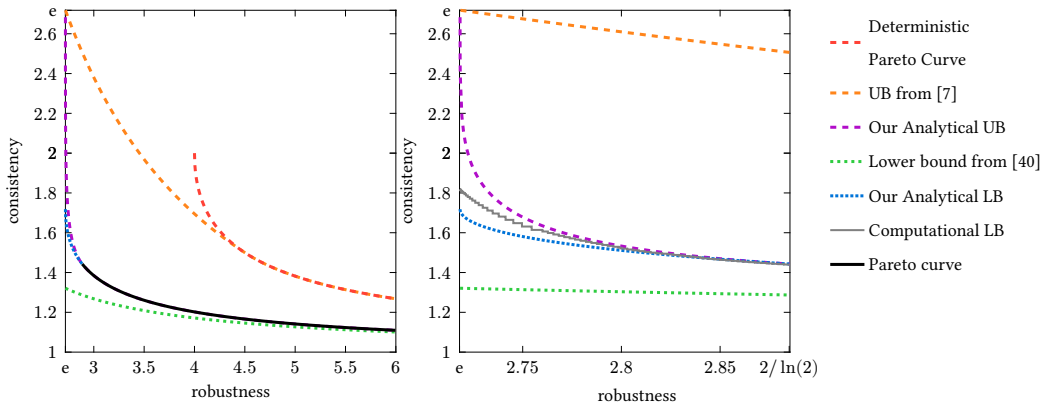


Figure 1: Consistency–robustness tradeoff for various bidding strategies and different ranges of the robustness R. The dashed orange curve corresponds to the bidding sequence from [7], while the dotted green curve corresponds to the lower bound given in [40]. UB and LB denote upper and lower bounds, respectively. The red dashed line corresponds to the deterministic Pareto curve of [5].

The analysis of learning-augmented algorithms is intricate. Several works have focused on tradeoffs between the consistency (i.e., the performance of the algorithm assuming error-free prediction) and the robustness (i.e., the performance under adversarial prediction). Tradeoffs for deterministic algorithms have been studied for a variety of problems, e.g. [41, 31, 4, 13, 2, 15], often yielding *Pareto*-optimal results; in contrast, randomized tradeoffs have been limited to problems such as ski rental [42], search games [6], k-server [20] and online bidding [7, 40], as we discuss below.

Online bidding and related problems For deterministic bidding, and in the absence of predictions, a folklore result that goes back to linear search [14] shows that the optimal robustness (i.e., the optimal *competitive* ratio) is equal to 4, and is achieved by a simple doubling strategy of the form $X = (2^i)_{i\in\mathbb{Z}}$. Randomization improves the competitive ratio to e, which is tight as proven in [21] using a complex approach based on dual fitting of a linear program. Online bidding has many applications in optimization problems such as incremental clustering and online k-median [19], load balancing [10], searching for a hidden target in the infinite line [14], and the design of algorithms with interruptible capabilities [39]; we refer to the survey [22] for a discussion of several applications.

Beyond worst-case competitive analysis, online bidding is one of the most studied problems under the learning-augmented lens. For deterministic algorithms, [5] gave a Pareto-optimal algorithm, with consistency $b/(b-1)$ and robustness $b^2/(b-1)$, for any $b > 1$, whereas [3, 28] gave performance guarantees as a function of the prediction error. Concerning randomized strategies, a heuristic and a non-tight analytical lower bound were given in [7]. In [16], randomization is leveraged implicitly by a perturbation of the prediction, though their algorithm is otherwise deterministic. The tradeoff between consistency and robustness was also studied in [40] under the name of the *button* problem, for which a non-tight analytic lower bound was shown; specifically, [40] showed that any $(1+\lambda)$-consistent algorithm must have robustness at least $(1+\lambda^2)/(2\lambda)$, for $\lambda > 0$.

2 Preliminaries

We say that a bidding algorithm is R-robust and C-consistent if its robustness is at most R and its consistency at most C. In our analysis, we will typically minimize the consistency, given a robustness requirement R. Note that R must be at least the optimal randomized competitive ratio, namely at least e.

Note that the normalized cost of a sequence is invariant under multiplication of the bids, a property that follows from the fact that the bidding sequences are bi-infinite. Formally, we have $\mathrm{nc}(\delta X, \delta u) = \mathrm{nc}(X, u)$ for any $\delta > 0$, where $\delta X := (\delta x_i)_{i\in\mathbb{Z}}$. Essentially, if the minimum of $\mathrm{nc}(X, u)$ is reached at u^*, then scaling X by $\hat{u}/u^*$ brings the minimum to the predicted threshold $\hat{u}$. This allows us to

work with a simpler expression of the consistency and robustness

$$\operatorname{cons}(X) = \inf_u \operatorname{nc}(X,u) \text{ and } \operatorname{rob}(X) = \sup_u \operatorname{nc}(X,u), \tag{2}$$

which provides a convenient expression that we will be using in the remainder of this work.

Given $R \geq e$, we define

$$\underline{w}(R) := -\frac{1}{W_{-1}(-1/R)}, \quad \overline{w}(R) := -\frac{1}{W_0(-1/R)}, \tag{3}$$

where W_0 and W_{-1} denote the principal and lower real branches of the Lambert function. We will use the shorthanded versions $\overline{w}, \underline{w}$ when R is clear from context.

The remainder of the paper is organized as follows. In Section 3, we introduce the framework of bidding functions, establish its equivalence with bidding sequences, and illustrate the analysis of several classes of algorithms. Section 4 presents and analyzes our main algorithm, while in Section 5 we prove the corresponding lower bounds using a linear programming approach. Section 6 concludes with a numerical comparison of algorithms that incorporates prediction error. Finally, in Section 7, we present an experimental application of randomized online bidding to the incremental median problem.

3 Bidding Functions : a Framework for Bidding Sequences

In this section, we introduce the concept of a *bidding function* that is instrumental in our analysis of the problem. Let $B : \mathbb{R} \to \mathbb{R}_+$ be a non-decreasing function such that $\lim_{t\to-\infty} B(t) = 0, \lim_{t\to+\infty} B(t) = +\infty$ and for all $t \in \mathbb{R}$, $\int_{-\infty}^{t} B(x)\,\mathrm{d}x$ is finite. We call the domain of B the set of *reference points*. We define the randomized bidding sequence *induced by* B as $X_B := (B(i+\lambda))_{i\in\mathbb{Z}}$ where λ is chosen uniformly at random in $[0,1]$.

In order to express the normalized cost of X as a function of B, we introduce the concept of *normalized mass* $\tilde{B}$, defined by

$$\tilde{B}(t) := \int_{-\infty}^{t+1} \frac{B(x)}{B(t)}\,\mathrm{d}x.$$

In addition, we define the *work function* w_B, which will be useful in the analysis of our algorithms. This function represents the expected sum over all bids up to the threshold $B(t)$, normalized by this threshold.

$$w_B(t) := \int_{-\infty}^{t} \frac{B(x)}{B(t)}\,\mathrm{d}x.$$

When B is clear from the context, we will sometimes use the shortcut $w_t := w_B(t)$. To be able to relate $\tilde{B}$ to $\operatorname{nc}(X_B)$, we define the *generalized inverse* of B. Formally $B^-(u)$ is the infimum over all t such that $B(t) \geq u$, see [25].

With the above definitions established, we can prove a strong connection between $\operatorname{nc}(X_B,u)$ and $\tilde{B}$.

Theorem 1. *For every $u \in \mathbb{R}_+$ we have*

$$\inf_u \operatorname{nc}(X_B,u) = \inf_t \tilde{B}(t), \quad \textit{and} \quad \sup_u \operatorname{nc}(X_B,u) = \sup_t \tilde{B}(t).$$

Theorem 1 allows us to use bidding functions explicitly in the analysis. Hence, we can restate the consistency and robustness expressions (2) in terms of $\tilde{B}(t)$ as

$$\operatorname{cons}(B) := \inf_t \tilde{B}(t), \qquad \operatorname{rob}(B) := \sup_t \tilde{B}(t). \tag{4}$$

As an example, the exponential bidding function $B(t) = e^t$ induces the sequence $(e^{i+\lambda})_{i\in\mathbb{Z}}$ where $\lambda \sim U[0,1]$ which has optimal robustness e as shown in [21]. As a different example, the randomized algorithm of [7] is induced by the exponential of a piecewise linear function, as we will discuss shortly. As a last example, a piecewise constant function such as $B(t) = 2^{\lfloor t \rfloor}$ induces the doubling sequence that has optimal deterministic robustness equal to 4 [14]. These examples illustrate that bidding functions capture a broad spectrum of classical bidding strategies, and in fact this connection is complete: *every* randomized bidding sequence is induced by some bidding function.

Theorem 2. *For every randomized sequence X, there is a bidding function B such that the consistency (respectively, robustness) of X is lower bounded by the consistency (respectively, robustness) of B.*

In Sections 4 and 5, we leverage Theorem 2 to obtain near-tight consistency–robustness tradeoffs. While these results require more intricate bidding functions, several interesting structural insights already arise from considering relatively simple ones. These examples help illustrate the main ideas and provide intuition for the more general techniques developed later. In particular, we consider the following classes of bidding functions:

- Class $\mathcal{E}$ is comprised of all exponential bidding functions. This class contains, for instance, the e-robust bidding function $B(t) = e^t$ of [21]. This warm-up class induces sequences of constant normalized cost, and reveals a connection between R-robustness and the expressions $\overline{w}, \underline{w}$, as defined in Section 2.

- Class $\mathcal{D}$ contains all bidding functions B that are exponential within each half-open interval of the form $[i, i+1)$ for $i \in \mathbb{Z}$, and satisfy, in addition, that $B(t+1)/B(t)$ is constant. This class captures the bidding sequences studied in [7]. Our analysis allows us to obtain exact, closed-form expressions of the consistency–robustness tradeoffs, which was left open in [7].

- Class $\mathcal{I}$ contains all *continuous* functions that are exponential within each half-open interval of the form $[i, i+1)$ for $i \in \mathbb{Z}$. Our analysis proves, in particular, that if the robustness requirement R is relatively small (i.e., smaller than 4.8), this class induces R-robust algorithms that have better consistency than the best known result of [7]. We also prove that the tradeoff improvement becomes quantifiably more pronounced as $R \to e$.

We refer to the Appendix A.1 for the detailed discussion, with statements and proofs of results. We note that the class $\mathcal{D}$ consists of functions that are *non-continuous*, whereas the class $\mathcal{I}$ is comprised by linear *interpolations*. Together, these classes illustrate that achieving Pareto-optimality requires bidding functions with genuinely nontrivial structure.

4 Upper Bounds

In this section, we present a bidding algorithm and analyze its performance. In Section 5, we will prove that our algorithm is Pareto-optimal for almost all robustness values R. Our algorithm is defined in terms of a bidding function, given in Definition 3. Lemma 5 will establish that this function is a valid bidding function and Theorems 6 and 7 will prove its theoretical guarantees. Last, Corollary 8 will prove that the algorithm has much better consistency than state of the art algorithms even when R is relatively small.

Let us first give some intuition behind our approach. The problem of finding the optimal bidding function B is equivalent to the following continuous optimization problem.

$$\min \int_{-\infty}^{1} B(x)\,dx \quad \text{s.t.} \int_{-\infty}^{t+1} B(x)\,dx \le RB(t) \quad \forall t \in \mathbb{R} \quad \text{and } B(0) = 1. \tag{5}$$

It is not obvious how to solve this optimisation problem. However, we can make some initial helpful observations about the structure of efficient solutions. Namely:

- In $(-\infty, 0)$, we make the robustness constraint of (5) tight, since this helps obtaining a solution. This yields the delay differential equation $B(t+1) = RB'(t)$, which gives a sequence of some polynomials we describe later in Definition 3.
- In $[0, 1]$, $B(t)$ must be at least 1, by monotonicity of B and the condition $B(0) = 1$. Hence, in order to minimize the objective in (5), we can take $B(t) = 1$ in this interval.
- In $[1, \infty)$, we choose $B(t) = e^{(t-1)/\overline{w}}$. Here, we are informally "after" the prediction, and we only need to guarantee robustness. Choosing a simple exponential function accomplishes this; the exponent coefficient is equal to $1/\overline{w}$ because this is the smallest coefficient that guarantees R-robustness (as we learned from the study of class $\mathcal{E}$ in Appendix A.1.1).

Solving experimentally discretized versions of (5) confirms the above intuition, as long as R is large enough, and specifically $R \ge 2/\ln 2$. However, when R is small, some modifications are needed. Specifically, B must be such that $B(t) = 1$ for a smaller interval than $[0, 1]$, as defined in Definition 21. The above process leads us to the following definition of the bidding function A.

Definition 3. *Given $R \geq 2/\ln 2$, define $x := 1/R$, $\mu := R - \overline{w}$. Define the function A*

$$A(t) = \begin{cases} e^{(t-1)/\overline{w}} & \text{if } 1 \leq t \\ 1 & \text{if } 0 \leq t \leq 1, \\ p_k(s) & \text{if } t \leq 0 \text{ for } k = \lceil -t \rceil, s = [-t]. \end{cases}$$

Here the polynomial p_k is defined inductively as follows:

$$p_0(s) = 1, \quad p_1(s) = x(\mu - 1 + s), \quad p_{k+1}(s) = p_k(0) - x \int_s^1 p_k(u)\,\mathrm{d}u.$$

Proposition 4. *Let $R \geq e$ and $S = (-\infty, 0) \backslash \mathbb{Z}_-$. Let B be a bidding function continuous in $(-\infty, 0)$ and differentiable in S that satisfies the following delay differential equation in S*

$$RB'(t) = B(t+1).$$

Then, for $t < 0$, we have $\tilde{B}(t) = R$.

The following lemma shows that A satisfies the properties outlined in Section 3, hence it is a valid bidding function.

Lemma 5. *A is a valid bidding function.*

The next theorem establishes our main upper bound.

Theorem 6. *For $R \geq 2/\ln 2$, it holds that $\mathrm{cons}(A) = R - \overline{w}$ and $\mathrm{rob}(A) = R$.*

The case $R \leq 2/\ln 2$ is technically more challenging. Leveraging the power of bidding functions, we solved a discretized LP so as to obtain intuition about the optimal bidding function. The solution reveals a similar structure to the case $R \geq 2/\ln(2)$ with one significant difference: the function does not induce a deterministic bid anymore. This complicates the task of finding a tight upper bound. Nevertheless, we prove the following upper bound, whose proof is along the lines of the proof of Theorem 6, using a similar, but technically more involved definition of the bidding function $A(t)$.

Theorem 7. *Let $e \leq R < 2/\ln 2$, let $\overline{w}$ be the largest solution of $we^{1/w} = R$, and set $\alpha := 1/\overline{w}$. Let $y \in (0, 1]$ be the unique solution of $\alpha = y + \ln(2 - y)$. Then*

$$\mathrm{cons}(A) = \frac{R}{2 - y} \text{ and } \mathrm{rob}(A) = R.$$

We also obtain the following corollary to Theorem 7 that bounds the consistency of A when R is relatively close to the optimal optimized competitive ratio e.

Corollary 8. *For the bidding function $A(t)$ used in the proof of Theorem 7, if $R = e + \varepsilon$ and $\varepsilon \to 0$, then*

$$\mathrm{cons}(A) \leq e - \Omega(\varepsilon^{1/4}).$$

We note that Corollary 8 significantly improves upon the upper bound of [7] which showed that $C = R - \Omega(\varepsilon)$.

5 Lower Bound

In this section, we derive a near-tight lower bound using a linear programming (LP) approach. We define a family of LPs parameterized by $R \geq e$ and integers $a, N, M \in \mathbb{Z}_+$. The variables, denoted by x_k for $k \in [-N, M]$, represent the values of the bidding function at discrete reference points. The objective is to minimize the normalized mass at the reference point $t = 0$, which corresponds to the consistency. At the same time, the function is required to have normalized mass at most R at every reference point that is an integer multiple of $1/a$ in the interval $[-N/a, M/a]$; these constraints encode the R-robustness requirement.

While LP-based approaches have previously been used to derive lower bounds for randomized online bidding [21, 40], our formulation based on bidding functions is fundamentally different. Our LPs directly capture a discretized bidding function using a mesh of granularity $1/a$, leading to a simple

and transparent representation of both consistency and robustness constraints. In contrast, the LPs of [21, 40] are based on discretized bidding sequences and use variables of the form $x_{i,j}$, representing the probability that bid b_i precedes bid b_j in the generated sequence. This sequence-based formulation leads to significantly more involved cost expressions, which limits the strength and tractability of the resulting bounds.

Formally, we define the linear program $P^R_{a,N,M}$ with constraints labeled $\lambda, \gamma_k, \beta_k, \theta$:

$$
\begin{aligned}
\min \quad & C \\
\text{s.t.} \quad & x_0 \geq 1, && && (\lambda) \\
& x_k \leq x_{k+1} && (-N \leq k \leq M-1) && (\gamma_k) \\
& \frac{1}{a} \sum_{j=-N}^{\min(M,k+a-1)} x_j \leq R x_k && (-N \leq k \leq M) && (\beta_k) \\
& \frac{1}{a} \sum_{j=-N}^{a-1} x_j \leq C, && && (\theta) \\
& x_k \geq 0 && (-N \leq k \leq M)
\end{aligned}
$$

The following lemma shows that $P^R_{a,N,M}$ is a lower-bound family of LPs.

Lemma 9. *For every $a, N, M \geq 1$ and $R \geq e$, the optimal objective value of $P^R_{a,N,M}$ is a lower bound on the consistency of every R-robust monotone bidding sequence.*

Let us denote by $D^R_{a,N,M}$ the dual of $P^R_{a,N,M}$, which is shown below. Note that the dual LP does not have a variable that corresponds to the primal constraint labeled (θ), since it can be assumed to be 1 by scaling. From weak duality, the objective of any feasible solution of the dual LP is a lower bound on the optimal objective of $P^R_{a,N,M}$, hence also on the optimal consistency of R-robust bidding functions.

$$
\begin{aligned}
\max \quad & \lambda \\
& \text{s.t. for all } -N \leq k \leq M \\
& \quad \frac{1}{a}\mathbf{1}_{\{k \leq a-1\}} + \frac{1}{a} \sum_{j=\max(k+1-a,-N)}^{M} \beta_j + \gamma_k - \gamma_{k-1} \geq \lambda \mathbf{1}_{\{k=0\}} + R\beta_k, && (x_k) \\
& \quad \lambda, \beta_k, \gamma_k \geq 0, \\
& \quad \gamma_M = \gamma_{-N-1} = 0.
\end{aligned}
$$

Proposition 10. *Let $R \geq e$ and $a, N, M \geq 1$. The objective value of every feasible solution to the dual LP $D^R_{a,N,M}$ is a lower bound on the consistency of an R-robust bidding sequences.*

We can now state the main result of this section.

Theorem 11. *Let $R > e$, and let $\overline{w} > 1$ be the largest solution of $\overline{w}e^{1/\overline{w}} = R$. Then for every $\epsilon > 0$, every R-robust bidding sequence has consistency at least*

$$R - \overline{w} - \varepsilon.$$

We conclude this section with a note that a computational solution to the lower-bound LP provides an improved lower bound in the robustness region $[e, 2/\ln(2)]$; refer to Fig 1.

6 Numerical Evaluation

We present a numerical evaluation of the performance of our algorithms. We consider the following setup. The oracle outputs a prediction $\hat{u}$, and the true threshold u is generated at random according to $u = \hat{u} \cdot \exp(\eta)$, where $\eta \sim \mathcal{N}(0, \sigma^2)$ represents a Gaussian noise. The variance σ^2 thus captures the prediction error of the oracle. Note that the median of u is equal to $\hat{u}$.

We compare the performance of the following algorithms: the best algorithms from the class $\mathcal{D}$ and $\mathcal{I}$, defined in Section 3, as well as the algorithm A, which we analyzed theoretically in Theorems 6

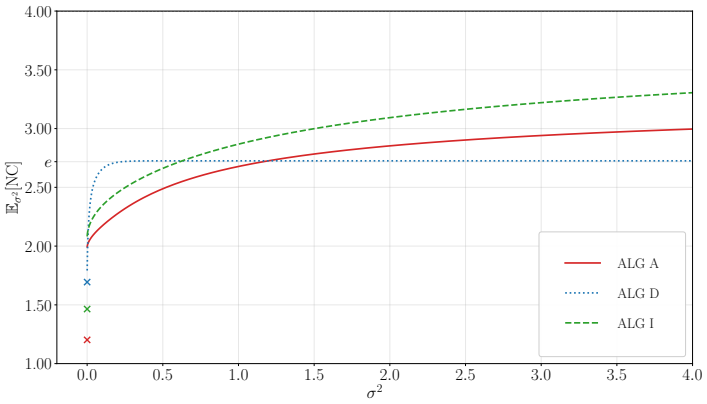


Figure 2: Expected normalized cost as function of $\sigma^2 \in [0, 4]$.

and 7. We denote these three algorithms by D, I, and A, respectively. Recall that algorithm D is the state-of-the-art given in [7].

Figure 2 depicts the expected normalized cost of the algorithms, as a function of the variance σ^2, for robustness $R = 4$. Additional plots for other robustness values can be found in Appendix D. The crossed points depict the consistency attained by the algorithms.

The plots show that both algorithms I and A achieve significantly better expected normalized cost than algorithm D, as long as the prediction error remains reasonably small. In particular, algorithm A attains the smallest normalized cost whenever $\sigma^2 \leq 1.19$, covering a broad range of prediction errors. Moreover, both A and I exhibit substantially better *smoothness*, that is, a more gradual performance degradation as the prediction error increases, compared to D, whose normalized cost approaches its worst-case value even for relatively small prediction errors.

7 Application to the Incremental Medians Problem

In this section, we conduct an experimental study of the application of randomized bidding to the incremental median problem, a classical online clustering problem. Previous work on this problem has been purely theoretical; we provide the first experimental results, in the context of a learning-augmented setting.

Problem statement Incremental medians is closely related to the well-studied k-median problem. Here, we are given a metric space V with distance function d. In its simplest variant, the goal is to choose a set $F \subseteq V$ of k *facilities*. Each point $v \in V$ is assigned to its closest facility, which incurs cost defined as

$$\operatorname{cost}(F) := \sum_{v \in V} \min_{f \in F} d(v, f).$$

The quality of a solution is measured by the *approximation ratio*, defined as $\operatorname{cost}(F)/\operatorname{cost}(F_k)$, where $F_k \subseteq V$ denotes an optimal set of k facilities minimizing the cost. For a randomized algorithm producing a random set F, we consider the expected value of $\operatorname{cost}(F)$ divided by the $\operatorname{cost}(F_k^*)$.

The *incremental median problem* introduces an additional constraint: the algorithm does not know in advance how many facilities it will be allowed to choose, and must therefore construct the solution sequentially, without the possibility of replacing previously selected facilities. Formally, a solution to the incremental median problem is a total order on V. We denote by F_k the first k facilities in that order, and by $\operatorname{cost}(F_k)/\operatorname{cost}(F_k^*)$ the approximation ratio of the algorithm for a specific value of k.

Algorithms Incremental medians was introduced under the name *online median problem* in [35], and its connection to the online bidding problem was studied in [21]. It was shown therein that, given a β-approximation algorithm for the k-median problem and a γ-competitive algorithm for the online bidding problem, one can obtain a $2\beta\gamma$-approximation algorithm for the incremental median problem. In Appendix E we extend this connection to the learning-augmented setting, with a prediction $\hat{k}$ on

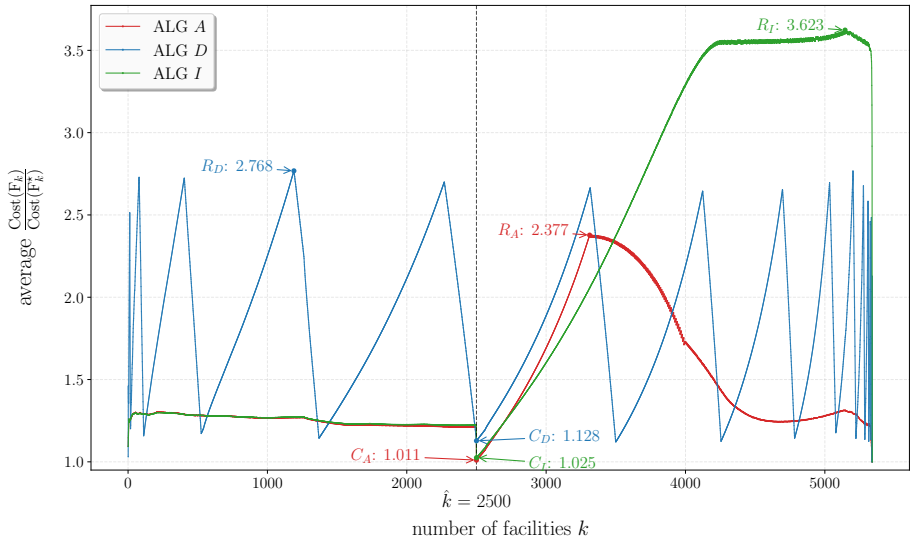


Figure 3: Average empirical approximation ratios for the incremental median problem applied to the Isle of Man road network, as a function of the actual number of facilities. C_X and R_X denote the empirical consistency and robustness of algorithm $X \in \{A, D, I\}$.

the allowable number of facilities to choose. We show that any R-robust and C-consistent bidding algorithm yields an algorithm for incremental median that is $2\beta C$-consistent, and $2\beta R$-robust. We also address technical details related to the implementation.

Datasets For our experimental setup, we use the road network of the Isle of Man from the *OpenStreetMap* project [37]. Restricting to its largest connected component yields a graph on a vertex set V of size 5340, where vertices correspond to geographic locations on the island. In this setting, facilities may represent, for example, mailboxes or electric charging stations, while the cost reflects the quality of service provided across the island. Edges correspond to road segments, and their weights represent traversal times. For the purpose of this study, we make all edges bidirectional by assigning to each undirected edge the minimum arc weight among both orientations. This connected edge-weighted graph induces a metric on V, where distances are defined by shortest-path lengths.

Experimental evaluation We choose $\hat{k} = 2500$ as the predicted number of facilities, and use bidding sequences optimized for low consistency at this value. Appendix E provides additional plots for other values of $\hat{k}$. Figure 3 depicts the average expected approximation ratio of various algorithms, based on different randomized bidding algorithms, as a function of the actual number of facilities k. We set the robustness requirement to $R = 4$. The figure depicts the empirical ratios based on three different bidding algorithms: The best bidding function from class $\mathcal{D}$ introduced in [7], the best function from class $\mathcal{I}$ in Section 3, and the Pareto-optimal function A analyzed in Theorem 6. Following the same notation as in Section 6, we denote by D, I and A the three algorithms.

We draw two main conclusions from Figure 3. First, the theoretical consistency guarantees of these bidding functions are reflected in the same ordering of the consistency for incremental median: A achieves the best consistency, followed by I and then D. More importantly, algorithm A exhibits a pronounced improvement on the approximation ratio even when the prediction is significantly inaccurate. This indicates that the benefits of A extend well beyond improved consistency–robustness tradeoffs, yielding practical gains even under large prediction errors.

The overall shape of the plots can be interpreted in light of the definitions and the structural properties of the bidding functions. For D, the plot is saw-like, similar to a deterministic plot, because the algorithm uses a relatively small amount of randomness. In contrast, the bidding functions A and I are similar for $k \in [1, \hat{k}]$, due to the presence of the same slope $1/\overline{w}$ in their exponent and the same

deterministic bid at $k = \hat{k}$. For $k \in [\hat{k}, 5000]$, A is markedly better than I, because in this interval the latter uses few bids, due to the slope $1/\underline{w}$ in its exponent.

In Appendix E we present additional experimental results, including a sensitivity analysis of the algorithm.

8 Conclusion

We presented algorithms that provably achieve a near-optimal consistency–robustness tradeoff for randomized learning-augmented bidding, effectively closing the gap in previous work. A very small gap for $R \leq 2/\ln 2$ remains, which we believe can be further narrowed by allowing more resources to the computational lower bound. Central to our analysis is a novel function-based framework that maps bidding sequences to bidding functions and enables both the positive and impossibility results. Beyond online bidding, this methodology extends naturally to closely related problems such as searching for a hidden target on the infinite line [14], as shown in [7]. We also aim to extend our approach to more challenging sequencing problems such as *contract scheduling* [39], where the optimal randomized competitive ratio remains unknown even in the classical setting without predictions, making it a natural candidate for function-based analysis techniques.

Acknowledgments and Disclosure of Funding

This work was partially funded by the project PREDICTIONS, grant ANR-23-CE48-0010 from the French National Research Agency (ANR), by the Paris Île-de-France Région in the framework of DIM AI4IDF, and by the FMJH Program PGMO.

References

[1] scikit-learn-extra documentation : Clustering with kmedoids, clara and common-nearest-neighbors, 2019. `https://scikit-learn-extra.readthedocs.io/en/stable/modules/cluster.html#k-medoids]`.

[2] Matteo Almanza, Flavio Chierichetti, Silvio Lattanzi, Alessandro Panconesi, and Giuseppe Re. Online facility location with multiple advice. *Advances in Neural Information Processing Systems*, 34:4661–4673, 2021.

[3] 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.

[4] Spyros Angelopoulos. Online search with a hint. *Inf. Comput.*, 295(Part B):105091, 2023.

[5] Spyros Angelopoulos, Christoph Dürr, Shendan Jin, Shahin Kamali, and Marc Renault. Online Computation with Untrusted Advice. In *11th Innovations in Theoretical Computer Science Conference (ITCS 2020)*, volume 151 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 52:1–52:15. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020.

[6] Spyros Angelopoulos, Thomas Lidbetter, and Konstantinos Panagiotou. Search games with predictions. *Operations Research*, 2026.

[7] Spyros Angelopoulos and Bertrand Simon. Learning-augmented online bidding in stochastic settings. In *Neurips*, The Thirty-Ninth Annual Conference on Neural Information Processing Systems, 2025.

[8] Antonios Antoniadis, Themis Gouleakis, Pieter Kleer, and Pavel Kolev. Secretary and online matching problems with machine learned advice. *Advances in Neural Information Processing Systems*, 33:7933–7944, 2020.

[9] David Arthur and Sergei Vassilvitskii. k-means++: the advantages of careful seeding. In Nikhil Bansal, Kirk Pruhs, and Clifford Stein, editors, *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2007, New Orleans, Louisiana, USA, January 7-9, 2007*, pages 1027–1035. SIAM, 2007.

[10] Yossi Azar. On-line load balancing. *Online algorithms: the state of the art*, pages 178–195, 2005.

[11] 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.

[12] Xingjian Bai and Christian Coester. Sorting with predictions. In *NeurIPS*, 2023.

[13] 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.

[14] Anatole Beck and Donald J. Newman. Yet more on the linear search problem. *Israel Journal of Mathematics*, 8:419–429, 1970.

[15] Ziyad Benomar, Lorenzo Croissant, Vianney Perchet, and Spyros Angelopoulos. Pareto-optimality, smoothness, and stochasticity in learning-augmented one-max-search. *arXiv preprint arXiv:2502.05720*, 2025.

[16] Ziyad Benomar and Vianney Perchet. On tradeoffs in learning-augmented algorithms. In *AISTATS*, volume 258 of *Proceedings of Machine Learning Research*, pages 802–810. PMLR, 2025.

[17] Allan Borodin and Ran El-Yaniv. *Online computation and competitive analysis*. Cambridge University Press, 1998.

[18] Ying Cao, Bo Sun, and Danny HK Tsang. Optimal online algorithms for one-way trading and online knapsack problems: A unified competitive analysis. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 1064–1069. IEEE, 2020.

[19] Moses Charikar, Chandra Chekuri, Tomás Feder, and Rajeev Motwani. Incremental clustering and dynamic information retrieval. In *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing*, pages 626–635, 1997.

[20] Nicolas Christianson, Junxuan Shen, and Adam Wierman. Optimal robustness-consistency tradeoffs for learning-augmented metrical task systems. In *AISTATS*, volume 206 of *Proceedings of Machine Learning Research*, pages 9377–9399. PMLR, 2023.

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

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

[23] Robert Corless, Gaston Gonnet, D. Hare, David Jeffrey, and Donald Knuth. On the Lambert W function. *Advances in Computational Mathematics*, 5:329–359, 01 1996.

[24] Paul Dütting, Silvio Lattanzi, Renato Paes Leme, and Sergei Vassilvitskii. Secretaries with advice. *Mathematics of Operations Research*, 49(2):856–879, 2024.

[25] Paul Embrechts and Marius Hofert. A note on generalized inverses. *Mathematical Methods of Operations Research*, 77(3):423–432, 2013.

[26] Sreenivas Gollapudi and Debmalya Panigrahi. Online algorithms for rent-or-buy with expert advice. In *Proceedings of the 36th International Conference on Machine Learning, ICML*, volume 97 of *Proceedings of Machine Learning Research*, pages 2319–2327. PMLR, 2019.

[27] Sungjin Im, Ravi Kumar, Mahshid Montazer Qaem, and Manish Purohit. Online knapsack with frequency predictions. *Advances in Neural Information Processing Systems*, 34:2733–2743, 2021.

[28] 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.

[29] Silvio Lattanzi, Thomas Lavastida, Benjamin Moseley, and Sergei Vassilvitskii. Online scheduling via learned weights. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1859–1877. SIAM, 2020.

[30] Changyeol Lee, Dahoon Lee, Jongseo Lee, Yongho Shin, and Changki Yun. Optimal learning-augmented algorithm for online bidding. *arXiv preprint arXiv:2605.07349*, April 2026.

[31] Russell Lee, Bo Sun, Mohammad Hajiesmaili, and John C. S. Lui. Online search with predictions: Pareto-optimal algorithm and its applications in energy markets. In *The 15th ACM International Conference on Future and Sustainable Energy Systems, e-Energy 2024, Singapore, June 4-7, 2024*, pages 50–71. ACM, 2024.

[32] Honghao Lin, Tian Luo, and David Woodruff. Learning augmented binary search trees. In *International Conference on Machine Learning*, pages 13431–13440. PMLR, 2022.

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

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

[35] Ramgopal R. Mettu and C. Greg Plaxton. The online median problem. *SIAM Journal on Computing*, 32(3):816–832, 2003.

[36] Michael Mitzenmacher and Sergei Vassilvitskii. Algorithms with predictions. In *Beyond the Worst-Case Analysis of Algorithms*, pages 646–662. Cambridge University Press, 2020.

[37] OpenStreetMap contributors. OpenStreetMap. `https://www.openstreetmap.org`, 2026. Data accessed 2026. Licensed under the Open Data Commons Open Database License (ODbL) v1.0.

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

[39] Stuart J. Russell and Shlomo Zilberstein. Composing real-time systems. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 212–217, 1991.

[40] 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 Transactions on Algorithms*, 22(1):1–30, 2025.

[41] Bo Sun, Russell Lee, Mohammad Hajiesmaili, Adam Wierman, and Danny Tsang. Pareto-Optimal Learning-Augmented Algorithms for Online Conversion Problems. In *Advances in Neural Information Processing Systems*, volume 34, pages 10339–10350. Curran Associates, Inc., 2021.

[42] Alexander Wei and Fred Zhang. Optimal robustness-consistency trade-offs for learning-augmented online algorithms. In *Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS)*, 2020.

Appendix

A Omitted material from Section 3

Proof of Theorem 1. For $u \in \mathbb{R}_+$ we have

$$\begin{aligned}
\operatorname{cost}(X_B, u) &= \int_0^1 \sum_{i\in\mathbb{Z}} B(i+\lambda) \cdot \mathbb{1}\left[B(i+\lambda-1) < u\right] \mathrm{d}\lambda \\
&= \sum_{i\in\mathbb{Z}} \int_0^1 B(i+\lambda) \cdot \mathbb{1}\left[B(i+\lambda-1) < u\right] \mathrm{d}\lambda \\
&= \sum_{i\le\lfloor B^-(u)\rfloor} \int_0^1 B(i+\lambda)\,\mathrm{d}\lambda + \int_0^1 B(\lfloor B^-(u)\rfloor + 1 + \lambda)\mathbb{1}\left[B(\lfloor B^-(u)\rfloor + \lambda) < u\right] \mathrm{d}\lambda \\
&= \int_{-\infty}^{\lfloor B^-(u)\rfloor+1} B(x)\,\mathrm{d}x + \int_0^{\lceil B^-(u)\rceil} B(\lfloor B^-(u)\rfloor + 1 + \lambda)\,\mathrm{d}\lambda \\
&= \int_{-\infty}^{\lfloor B^-(u)\rfloor+1} B(x)\,\mathrm{d}x + \int_{\lfloor B^-(u)\rfloor+1}^{\lfloor B^-(u)\rfloor+1+\lceil B^-(u)\rceil} B(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{B^-(u)+1} B(x)\,\mathrm{d}x.
\end{aligned}$$

Hence we obtain that

$$\operatorname{nc}(X_B, u) = \int_{-\infty}^{B^-(u)+1} \frac{B(x)}{u}\,\mathrm{d}x. \tag{6}$$

Let $u \in \mathbb{R}_+$, it holds that

$$\lim_{\substack{t\to B^-(u)\\ t>B^-(u)}} \tilde{B}(t) \le \operatorname{nc}(X_B, u) \le \lim_{\substack{t\to B^-(u)\\ t<B^-(u)}} \tilde{B}(t). \tag{7}$$

Indeed, both inequalities come from (6) and the continuity of $\int_{-\infty}^{t+1} B(x)\,\mathrm{d}x$ in t. For the left inequality, we also use that for all $t > B^-(u)$ we have $B(t) \ge u$. Similarly, the right inequality holds because for all $t < B^-(u)$, $B(t) \le u$.

Let $t \in \mathbb{R}$, we have

$$\operatorname{nc}(X_B, B(t)) \le \tilde{B}(t) \le \lim_{\substack{u\to B(t)\\ u>B(t)}} \operatorname{nc}(X_B, u). \tag{8}$$

Similarly, this follows using (6). For the left inequality, we also use that if $u = B(t)$, then $B^-(u) \le t$, and for the right inequality that $\lim_{\substack{u\to B(t)\\ u>B(t)}} B^-(u) \ge t$.

Combining the inequalities (7) and (8) completes the proof. □

Proof of Theorem 2. Let X be a randomized bidding sequence. For $v > 0$ let the random variable $N(v)$ be the number of bids in X between 1 and v. If v is smaller than 1, it counts negatively the number of bids between v and 1. Formally

$$N(v) := \sum_{k\in\mathbb{Z}} (\mathbb{1}\left[X_k < v\right] - \mathbb{1}\left[X_k < 1\right]).$$

We denote the expected number of bids in $[1, v)$, and its generalized inverse by $F(v)$ and $B(t)$ respectively:

$$\begin{aligned}
F(v) &:= \mathbb{E}[N(v)], && v > 0, \\
B(t) &:= \inf\{v > 0 | F(v) > t\}, && t \in \mathbb{R}
\end{aligned}$$

The cost of X on threshold u is

$$\mathrm{cost}(X, u) = \sum_{X_k < u} X_k + Y_u,$$

where Y_u is the first bid greater than or equal to u.

For all $u \geq 1$, we have:

$$\begin{aligned}
\int_1^u F(x)\,\mathrm{d}x &= \int_1^u \mathbb{E}\left[N(x)\right]\mathrm{d}x \\
&= \mathbb{E}\left[\int_1^u N(x)\,\mathrm{d}x\right] \\
&= \mathbb{E}\left[\int_1^u \sum_{k\in\mathbb{Z}} (\mathbb{1}\left[X_k < x\right] - \mathbb{1}\left[X_k < 1\right])\,\mathrm{d}x\right] \\
&= \mathbb{E}\left[\sum_{k\in\mathbb{Z}} \int_1^u (\mathbb{1}\left[X_k < x\right] - \mathbb{1}\left[X_k < 1\right])\,\mathrm{d}x\right] \\
&= \mathbb{E}\left[\sum_{k\in\mathbb{Z}} \mathbb{1}\left[X_k \in [1, u)\right] \cdot (u - X_k)\right] \\
&= \mathbb{E}\left[\sum_{u \leq X_k < 1} u - X_k\right] \\
&= u\mathbb{E}\left[N(u)\right] - \mathbb{E}\left[\sum_{1 \leq X_k < u} X_k\right]. \qquad (9)
\end{aligned}$$

Similarly, for all $u \leq 1$, we have:

$$\begin{aligned}
\int_1^u F(x)\,\mathrm{d}x &= -\int_u^1 \mathbb{E}\left[N(x)\right]\mathrm{d}x \\
&= u\mathbb{E}\left[N(u)\right] + \mathbb{E}\left[\sum_{u \leq X_k < 1} X_k\right]. \qquad (10)
\end{aligned}$$

We obtain from (9) and (10) that for all $u \geq v$:

$$\int_u^v F(x)\,\mathrm{d}x = v\mathbb{E}\left[N(v)\right] - u\mathbb{E}\left[N(u)\right] - \mathbb{E}\left[\sum_{u \leq X_k < v} X_k\right].$$

By the area-inverse formula:

$$\begin{aligned}
\int_{F(u)}^{F(v)} B(x)\,\mathrm{d}x &= vF(v) - uF(u) - \int_u^v F(x)\,\mathrm{d}x \\
&= \mathbb{E}\left[\sum_{u \leq X_k < v} X_k\right]. \qquad (11)
\end{aligned}$$

Noticing that $\lim_{u\to 0} F(u) = -\infty$, it directly follows from (11) that

$$\int_{-\infty}^{F(v)} B(x)\,\mathrm{d}x = \mathbb{E}\left[\sum_{0 \leq X_k < v} X_k\right]. \qquad (12)$$

Next, we bound the expectation of Y_u. For convenience, we denote $s := F(u)$, and introduce

$$G_u(y) := \mathbb{P}(Y_u < y).$$

Note that $\mathbb{1}[Y_u < y] \leq N(y) - N(u)$. Taking expectations, we obtain

$$G_u(y) \leq F(y) - F(u) = F(y) - s.$$

We define the quantile of Y_u, i.e. for $0 \leq t \leq 1$ let

$$Q_u(t) := \inf\{y \geq u | G_u(y) > t\},$$

be the generalized inverse of G_u. If $G_u(y) > t$, then $F(y) > s + t$, hence $B(s + t) \leq y$. Taking the infimum on both sides, we obtain that $B(s + t) \leq Q_u(t)$. By integrating this inequality,

$$\int_s^{s+1} B(x)\,\mathrm{d}x \leq \int_0^1 Q_u(x)\,\mathrm{d}x = \mathbb{E}[Y_u]. \tag{13}$$

We can now complete the proof of the theorem. For an arbitrary $u \in \mathbb{R}_+$, let $t = F(u)$, we have from (12) and (13) that:

$$\begin{aligned}
\tilde{B}(t) &= \int_{-\infty}^{t+1} B(x)\,\mathrm{d}x / B(t) \\
&= \frac{1}{B(t)}\Big(\int_{-\infty}^{F(u)} B(x)\,\mathrm{d}x + \int_{F(u)}^{F(u)+1} B(x)\,\mathrm{d}x\Big) \\
&\leq \frac{1}{u}\Big(\mathbb{E}\left[\sum_{0 \leq X_k < u} X_k\right] + \mathbb{E}[Y_u]\Big) \\
&= \frac{\mathbb{E}(\mathrm{cost}(X, u))}{u}.
\end{aligned} \tag{14}$$

For $t \in \mathbb{R}$, let $u := B(t)$. For every $v > u$, we have $F(v) > t$, hence

$$\int_{-\infty}^{t+1} \frac{B(x)}{v}\,\mathrm{d}x \leq \int_{-\infty}^{F(v)+1} \frac{B(x)}{v}\,\mathrm{d}x \leq \frac{\mathbb{E}[\mathrm{cost}(X, v)]}{v}.$$

Letting v tend to u, it follows that

$$\tilde{B}(t) \leq \lim_{v \to u} \frac{\mathbb{E}[\mathrm{cost}(X, v)]}{v}. \tag{15}$$

Combining both inequalities (14) and (15), we obtain the wanted result, i.e.

$$\inf_{t \in \mathbb{R}} \tilde{B}(t) \leq \inf_{u \in \mathbb{R}_+} \mathrm{nc}(X_B, u), \quad \sup_{t \in \mathbb{R}} \tilde{B}(t) \leq \sup_{u \in \mathbb{R}_+} \mathrm{nc}(X_B, u).$$

□

The following Lemma will be used in several proofs.

Lemma 12. *Let $T \in \mathbb{R}$ and B a bidding function on $(-\infty, T)$. If the work of B at T is at most w, i.e.*

$$\int_{-\infty}^{T} B(x)\,\mathrm{d}x \leq wB(T),$$

then one can extend the bidding function B on $\mathbb{R}$ by $B(t) = B(T)\exp((t-T)/w)$ so the normalized mass of B is bounded on $(T, +\infty)$:

$$\tilde{B}(t) \leq we^{\frac{1}{w}}.$$

Proof. This extension is positive and increasing, hence the obtained B is a valid bidding function on $\mathbb{R}$. Let $t \geq T$, we have

$$\begin{aligned}
\tilde{B}(t) &= \frac{\int_{-\infty}^{t+1} B(x)\,\mathrm{d}x}{B(t)} \\
&\leq \frac{1}{B(T)} \exp(\frac{T-t}{w})(\int_{-\infty}^{T} B(x)\,\mathrm{d}x + \int_{T}^{t+1} B(x)\,\mathrm{d}x) \\
&\leq \exp(\frac{T-t}{w})(w + w(\exp(\frac{t+1-T}{w}) - 1)) \\
&= w\exp(1/w).
\end{aligned}$$

□

A.1 Analysis of Bidding Function Classes $\mathcal{E}$, $\mathcal{D}$, $\mathcal{I}$

A.1.1 Class $\mathcal{E}$

An exponential bidding function B is defined by $B(t) = e^{t/w}$, for some parameter $w > 0$ and for all $t \in \mathbb{R}$. The normalized cost of a linear function is invariant with respect to the reference point. In particular, for any $t \in \mathbb{R}$, we have

$$\tilde{B}(t) = \tilde{B}(0) = \frac{\int_{-\infty}^{1} e^{x/w}\,\mathrm{d}x}{e^0} = we^{1/w}.$$

It is interesting to know for which values of w the bidding function is R-robust. For this purpose, we observe that the ratio $we^{1/w}$ is convex in w, which holds since its second derivative $e^{1/w}w^{-3}$ is positive for $w > 0$. In addition its limit at 0 from above is ∞, as is its limit at ∞. As a result, we have $we^{1/w} \leq R$, for all $w \in [\underline{w}, \overline{w}]$ where $\underline{w}, \overline{w}$ are the two solutions of the equation $we^{1/w} = R$. Note that for the extreme case $R = e$, we have $\underline{w} = \overline{w} = 1$.

A.1.2 Class $\mathcal{D}$

Bidding functions in class $\mathcal{D}$ are defined by two values $\ell, h \geq 0$ with $\ell + h > 0$, and have the form

$$B_{\ell,h}(t) = \exp\left(\lfloor t \rfloor \cdot (\ell + h) + [t] \cdot \ell\right),$$

where $[t] := t - \lfloor t \rfloor$ denotes the fractional part of t. When it is clear from context, we will omit the subscripts ℓ, h in the notation. We note that this class induces sequences equivalent to those studied in [7]. In this section we give analytical results about the robustness–consistency tradeoff of these functions.

Note that the parameter ℓ allows to tune the amount of randomness in the sequence. In particular, of $\ell = 0$, then the resulting sequence is deterministic.

Lemma 13. *Let $\ell, h \geq 0$ such that $\ell + h > 0$. We have*

$$\begin{aligned}
\mathrm{rob}(B_{\ell,h}) &= \lim_{t \to 1^-} \tilde{B}_{\ell,h}(t) = \frac{(e^\ell - 1)}{\ell} \frac{e^{\ell+h}}{(e^{\ell+h} - 1)} e^h \\
\mathrm{cons}(B_{\ell,h}) &= \tilde{B}_{\ell,h}(0) = \frac{(e^\ell - 1)}{\ell} \frac{e^{\ell+h}}{(e^{\ell+h} - 1)}.
\end{aligned}$$

Proof. Since bidding functions of class $\mathcal{D}$ satisfy that $B(t+1)/B(t)$ is constant, the normalized mass is 1-periodic. For this reason, we can restrict the analysis to the interval $[0, 1)$. We first compute $w_B(0)$.

By periodicity of the normalized mass, $w_B(0) = w_B(1)$, furthermore, if $\ell > 0$, then $w_B(1) = \frac{w_B(0)B(0) + \frac{e^\ell - 1}{\ell}}{B(1)}$. Thus we have $w_B(0) = \frac{e^\ell - 1}{\ell(e^{\ell+h} - 1)}$, which for convenience we denote w.

Next we compute

$$\begin{aligned}\tilde{B}(t) &= \frac{wB(1) + \int_1^{t+1} B(x)\,\mathrm{d}x}{B(t)} \\ &= \frac{we^{\ell+h} + e^{\ell+h}(\frac{e^{t\ell}-1}{\ell})}{e^{t\ell}} \\ &= e^{\ell+h}\left(\frac{1}{\ell} + e^{-t\ell}\left(w - \frac{1}{\ell}\right)\right).\end{aligned}$$

Since $\ell + h > \ell$, we have $w < 1/\ell$ so this function is increasing in $[0, 1)$. We finally obtain the equalities

$$\mathrm{rob}(B) = \lim_{t\to 1^-} \tilde{B}(t) = e^h(w + \frac{e^\ell - 1}{\ell}) = we^{\ell+2h}, \quad \mathrm{cons}(B) = we^{\ell+h}$$

.

We note that the limits for $\ell \to 0$ of these expressions give $\mathrm{rob}(B) = \lim_{t\to 1^-} \tilde{B}(t) = \frac{e^{2h}}{e^h-1}$ and $C = \frac{e^h}{e^h-1}$, as in the deterministic case. □

The obtained formulae for the robustness and consistency of bidding functions from class $\mathcal{D}$ make it possible to analyse the tradeoff they induce.

Theorem 14. *Let $1 < C \leq e$. Let $R = \min_{\ell,h} \{\mathrm{rob}(B_{\ell,h}) | \, \mathrm{cons}(B_{\ell,h}) \leq C\}$ be the optimal robustness among C-consistent bidding functions from class $\mathcal{D}$.*

- *If $1 < C \leq \frac{3}{2}$, then $R = \frac{C^2}{C-1}$, which is achieved by a deterministic bidding sequence.*
- *If $\frac{3}{2} \leq C \leq e$, then $R = \frac{\ell C^2 e^{-\ell}}{\ell C+1-e^\ell}$ for ℓ being the unique non negative root of*

$$1 - C\ell^2 - \ell - e^\ell(1 - 2\ell).$$

Proof. From the proof of Theorem 13 we have $R = we^{\ell+2h}$ and $C = we^{\ell+h}$ for $w = \frac{e^\ell-1}{\ell(e^{\ell+h}-1)}$. First we express R in terms of C and ℓ only. We have

$$\begin{aligned} C = \frac{e^\ell - 1}{\ell}\frac{e^{\ell+h}}{e^{\ell+h}-1} &\quad\Leftrightarrow \\ Ce^{\ell+h} = \frac{e^\ell - 1}{\ell}e^{\ell+h} + C &\quad\Leftrightarrow \\ e^h\left(e^\ell C - \frac{e^\ell - 1}{\ell}e^\ell\right) = C &\quad\Leftrightarrow \\ e^h(\ell C - (e^\ell - 1)) = \ell C e^{-\ell} &\quad\Leftrightarrow \\ e^h = \frac{\ell C e^{-\ell}}{\ell C + 1 - e^\ell}. & \end{aligned}$$

Hence $R = Ce^h = \frac{\ell C^2 e^{-\ell}}{\ell C+1-e^\ell}$. For fixed C, we want to find the value of ℓ which minimizes R. We have

$$\partial_\ell R = \frac{C^2 e^{-\ell}\ell^2}{(C\ell - e^\ell + 1)^2}\frac{1 - C\ell^2 - \ell - e^\ell(1-2\ell)}{\ell^2}.$$

For this product of two fractions, observe that the first fraction is always positive. Hence the sign of $\partial_\ell R$ equals the sign of the second fraction, which we denote by $F(C, \ell)$. Using a series expansion, we obtain

$$F(C, \ell) = \sum_{i=0}^{+\infty} \frac{3 + 2i}{(i + 2)!}\ell^i - C.$$

We immediately obtain the following :

- F is strictly increasing along ℓ.
- $F(C, 0) \geq 0$ if and only if $C \leq \frac{3}{2}$.
- If $C > 3/2$, then F is negative on $[0, \ell^*)$ and positive on $(\ell^*, +\infty)$, where ℓ^* is the unique root of F.

Hence we have that, if $C > 3/2$, then the optimal robustness is reached for the unique ℓ satisfying $F(C, \ell) = 0$. Else, it is reached by the deterministic case $\ell = 0$. □

A.1.3 Class $\mathcal{I}$

A bidding function B in class $\mathcal{I}$ is defined by its value at integer reference points $(B(t))_{t\in\mathbb{Z}}$ for fractional reference points t, it is defined as the exponential interpolation between $B(\lfloor t \rfloor)$ and B($\lceil t \rceil$). In this section we use an equivalent representation based on a bi-infinite sequence of exponential *slopes* $(\ell_i)_{i\in\mathbb{Z}}$, where $e^{\ell_i} = B(i+1)/B(i)$. Without loss of generality we assume $B(0) = 1$, which allows us to express B as

$$B(t) = \exp\left([t] \cdot \ell_{\lfloor t \rfloor} + \begin{cases} \sum_{i=0}^{\lfloor t \rfloor - 1} \ell_i & \text{if } t \geq 0 \\ -\sum_{i=\lfloor t \rfloor}^{-1} \ell_i & \text{if } t < 0 \end{cases} \right).$$

The following theorem is the consequence of the lower bound in Theorem 18 and the matching upper bound in Theorem 19.

Theorem 15. *Let B be an R-robust bidding function from class $\mathcal{I}$ which minimizes its consistency. Let $(\ell_i)_{i\in\mathbb{Z}}$ be its associated sequence of slopes.*

- *If $R \geq R_0$, then* $\mathrm{cons}(B) = \underline{w} + 1$.
- *If $R \leq R_0$, let ℓ^* such that $\underline{w} + e^{\ell^* - 1} = \overline{w} e^{\ell^*}$, then* $\mathrm{cons}(B) = \underline{w} + \frac{e^{\ell^*} - 1}{\ell^*}$.

For the analysis it will be important to bound the work $w_B(i)$ at every integer reference point $i \in \mathbb{Z}$.

Lemma 16. *If a linear continuous bidding function B is R-robust, then for every $i \in \mathbb{Z}$, $w_B(i) \in [\underline{w}, \overline{w}]$. Moreover, $\ell_i \leq 1/\underline{w}$ for every $i \in \mathbb{Z}$.*

Proof. Let $(\ell_i)_{i\in\mathbb{Z}}$ be a sequence of slopes, B the associated bidding function and $(w_B(i))_{i\in\mathbb{Z}}$ be the induced sequence of works. We define $w_i := w_B(i)$ for the ease of notation. Let us suppose that B is R-robust.

For convenience, we define a function f, which computes the increase of the cost in one step, and g normalizes f to compute the change in the work:

$$f(w, \ell) = w + \frac{e^\ell - 1}{\ell}, \qquad g(w, \ell) = f(w, \ell) e^{-\ell}.$$

Note that f is increasing in its two arguments, and g is increasing in w but decreasing in ℓ, both function are continuous on $[0, +\infty)$, by extending f with $f(w, 0) = w + 1$.

The normalized mass at integer points $\tilde{B}(i)$ is equal to $w_i + \frac{e^{\ell_i} - 1}{\ell_i}$, so R-robustness translates for every $i \in \mathbb{Z}$ into

$$w_i + \frac{e^{\ell_i} - 1}{\ell_i} \leq R.$$

We make the following claims:

1. If the sequence $(w_i)_{i\in\mathbb{Z}}$ converges (either when $i \to -\infty$, or $i \to +\infty$) towards w^*, then $w^* \in [\underline{w}, \overline{w}]$

2. If for some $k \in \mathbb{Z}$, $w_k > \overline{w}$ or $w_k < \underline{w}$, then $w_{k+1} > w_k$.

3. If for some $k \in \mathbb{Z}$, $w_k \geq \underline{w}$, then $w_{k+1} \geq \underline{w}$.

Next, we prove these claims:

1. By continuity of g, w^* must satisfy that there exists x such that $g(w^*, x) = w^*$, which gives $x = 1/w^*$. Thus the normalized cost at integer points must converge to $f(w^*, 1/w^*) = w^* e^{1/w^*} \leq R$. Which gives $w^* \in [\underline{w}, \overline{w}]$ by definition of these two values.

2. Let w_k be such a value. Note that $f(w_k, 1/w_k) > R$. We have $f(w_k, \ell_k) \leq R$, so $\ell_k < 1/w_k$ by monotonicity of f. Hence $g(w_k, \ell_k) = w_{k+1} > g(w_k, 1/w_k) = w_k$, by monotonicity of g.

3. Let w_k be such a value. Note that $f(\underline{w}, 1/\underline{w}) = R \geq f(w_k, \ell_k) \geq f(\underline{w}, \ell_k)$, hence $\ell_k \leq 1/\underline{w}$ by monotonicity of f. So $g(w_k, \ell_k) \geq g(\underline{w}, 1/\underline{w}) = \underline{w}$, by monotonicity of g.

Next, we prove that for all $i \in \mathbb{Z}$, $w_i \leq \overline{w}$: Suppose, by way of contradiction, that for some $k \in \mathbb{Z}$, $w_k > \overline{w}$. Then by (2), we have that for every $j \geq k$, $w_j < w_{j+1}$, hence the sequence $(w_i)_{i \in \mathbb{Z}}$ is increasing. It is also upper bounded because $w_j \leq f(w_j, x) \leq R$, and therefore it must converges to some $w^* > w_k > \overline{w}$, which is not possible by (1).

We apply the same process to prove that for every $i \in \mathbb{Z}$, $w_i \geq \underline{w}$: Let us suppose towards contradiction that some $w_k < \underline{w}$. By (2) and (3), it follows that the sequence $(w_i)_{i \leq k}$ is increasing, we also have it is lower bounded by 0, so it must converge when i goes to $-\infty$. Its limit w^* must satisfy $w^* < \underline{w}$, which is not possible by (1). □

The above Lemma formalizes an intuitive idea: if the work is too high at some point, then the robustness requirement cannot be met in the immediate future. Conversely, if the work is too low, it means the sequence is not dense in the past, hence the robustness requirement must have been broken at some point.

From now on, it will be important to study the normalized mass of a given bidding function in class $\mathcal{I}$, hence the next proposition.

Proposition 17. *Let B be a bidding function from class $\mathcal{I}$, let $(\ell_i)_{i \in \mathbb{Z}}$ be its sequence of slopes and $(w_i)_{i \in \mathbb{Z}}$ be its sequence of work at reference point $i \in \mathbb{Z}$. We have*

$$\tilde{B}(t) = e^{(1-[t])\ell_{\lfloor t \rfloor}} \left(w_{\lceil t \rceil} + \frac{e^{[t]\ell_{\lceil t \rceil}} - 1}{\ell_{\lceil t \rceil}} \right).$$

Proof. For any $t \in \mathbb{R}$, we have

$$\begin{aligned}
\tilde{B}(t) &= \int_{-\infty}^{t+1} B(x)/B(t) \, \mathrm{d}x \\
&= \int_{-\infty}^{\lceil t \rceil} B(x)/B(t) \, \mathrm{d}x + \int_{\lceil t \rceil}^{t+1} B(x)/B(t) \, \mathrm{d}x \\
&= \frac{1}{B(t)} \left(w_{\lceil t \rceil} B(\lceil t \rceil) + \int_{\lceil t \rceil}^{t+1} B(\lceil t \rceil) e^{[x]\ell_{\lceil t \rceil}} \, \mathrm{d}x \right) \\
&= \frac{B(\lceil t \rceil)}{B(t)} \left(w_{\lceil t \rceil} + \int_{0}^{[t]} e^{x\ell_{\lceil t \rceil}} \, \mathrm{d}x \right) \\
&= e^{(1-[t])\ell_{\lfloor t \rfloor}} \left(w_{\lceil t \rceil} + \frac{e^{[t]\ell_{\lceil t \rceil}} - 1}{\ell_{\lceil t \rceil}} \right).
\end{aligned}$$

□

Lemma 18. *Let B be an R-robust bidding function from class $\mathcal{I}$, $(\ell_i)_{i\in\mathbb{Z}}$ its associated sequence of slopes, and $(w_B(i))_{i\in\mathbb{Z}}$ its sequence of works. Then* $\mathrm{cons}(B)$ *satisfies the following :*

- *If $R \geq R_0$, then* $\mathrm{cons}(B) \geq \underline{w} + 1$.
- *If $R \leq R_0$, let ℓ^* such that $g(\underline{w}, \ell^*) = \overline{w}$, then* $\mathrm{cons}(B) \geq \underline{w} + \frac{e^{\ell^*}-1}{\ell^*}$.

Proof. We start by proving that we can modify the prefix of a slope sequence to ensure its work is $\underline{w}$ on the modified prefix, without violating robustness, nor making the consistency worse.

Let $(\ell_i)_{i\in\mathbb{Z}}$ be a sequence of slopes that, without loss of generality, reaches its minimal normalized mass at $t \in [1, 2]$. Let $(d_i)_{i\in\mathbb{Z}}$ be the same sequence of slopes except for $i \leq 0$ where $d_i = 1/\underline{w}$. We call μ and π their associated bidding functions, respectively. First, note that the normalized mass of π on $(-\infty, 0]$ is equal to R, and that for $t \in [1, +\infty)$, we have $\tilde{\mu}(t) \geq \tilde{\pi}(t)$.

On $[0, 1]$ we have using Proposition 17

$$\begin{aligned}\tilde{\pi}(t) &= e^{(1-t)/\underline{w}} \left(\underline{w} + \frac{e^{t\ell_1} - 1}{\ell_1}\right) \\ &= \left(R - \frac{e^{1/\underline{w}}}{\ell_1}\right) e^{-t/\underline{w}} + e^{t(\ell_1 - 1/\underline{w})} \frac{e^{1/\underline{w}}}{\ell_1}.\end{aligned}$$

We recall that by Lemma 16, $\underline{w} \leq 1/\ell_1$. Deriving this expression in t, we obtain

$$\begin{aligned}(\tilde{\pi})'(t) &= -\frac{1}{\underline{w}} e^{-t/\underline{w}} \left(\underline{w} e^{1/\underline{w}} + e^{1/\underline{w}} \frac{e^{t\ell_1} - 1}{\ell_1}\right) + e^{t(\ell_1 - 1/\underline{w})} e^{1/\underline{w}} \\ &= e^{t(\ell_1 - 1/\underline{w})} \frac{e^{1/\underline{w}}}{\underline{w}} \left(\underline{w} - \frac{1}{\ell_1}\right) + \frac{e^{-t/\underline{w}}}{\underline{w}} e^{1/\underline{w}} \left(\frac{1}{\ell_1} - \underline{w}\right) \\ &= \frac{e^{1/\underline{w}}}{\underline{w}} e^{-t/\underline{w}} \left(\underline{w} - \frac{1}{\ell_1}\right) (e^{t\ell_1} - 1) \leq 0.\end{aligned}$$

Since $\tilde{\pi}(0) = R$, we have that $\tilde{\pi}(t) \leq R$ for every $t \in [0, 1]$, thus the bidding function π is R-robust.

The consistency is reached for $t \in [1, 2]$, using Proposition 17 with $w_1 = g(\underline{w}, \ell_1) = e^{-\ell_1}(\underline{w} + \frac{e^{\ell_1}-1}{\ell_1})$:

$$\tilde{\pi}(t) = \left((\underline{w} + \frac{e^{\ell_1} - 1}{\ell_1}) e^{-\ell_1} + \frac{e^{(t-1)\ell_2} - 1}{\ell_2}\right) \cdot e^{(2-t)\ell_1}.$$

Equivalently, for $t \in [0, 1]$,

$$\begin{aligned}\tilde{\pi}(t+1) &= \left(\underline{w} + \frac{e^{\ell_1} - 1}{\ell_1} + e^{\ell_1} \frac{e^{t\ell_2} - 1}{\ell_2}\right) \cdot e^{-t\ell_1} \\ &= e^{\ell_1(1-t)} \left(w_2 + \frac{e^{t\ell_2} - 1}{\ell_2}\right).\end{aligned}$$

Deriving this expression in t, we obtain

$$\begin{aligned}(\tilde{\pi})'(t+1) &= \frac{e^{\ell_1(1-t)}}{\ell_2} \left(\ell_2 e^{\ell_2 t} - \ell_1(e^{\ell_2 t} + \ell_2 w_2 - 1)\right) \\ (\tilde{\pi})'(1) &= \frac{e^{\ell_1}}{\ell_2}(\ell_2 - \ell_1\ell_2 w_2) = e^{\ell_1}(1 - \ell_1 w_2) = 1 - \ell_1\underline{w} \geq 0\end{aligned}$$

Hence the function $\tilde{\pi}$ is increasing at $t = 1$. Noticing that the function $\tilde{\pi}$ is bi-tonic on $[1, 2]$ as being the sum of two exponential functions, we obtain that its minimum is reached on the boundary. Hence $\mathrm{cons}(\pi) = \min(\tilde{\pi}(1), \tilde{\pi}(2))$.

We have $\tilde{\pi}(1) = f(\underline{w}, \ell_1)$ and $\tilde{\pi}(2) = f(w_2, \ell_2)$. Thus:

- $\text{cons}(\pi) = \min(\tilde{\pi}(1), \tilde{\pi}(2)) \geq f(\underline{w}, 0) = \underline{w} + 1$, by monotonicity of f in its two arguments.
- If $\overline{w} - \underline{w} \leq 1$, then let ℓ_1^* be such that $g(\underline{w}, \ell_1^*) = \overline{w}$. We have $g(\underline{w}, \ell_1) = w_2 \leq \overline{w} = g(\underline{w}, \ell_1^*$. Therefore $\ell_1 \geq \ell_1^*$, so we have $\tilde{\pi}(1) = f(\underline{w}, \ell_1) \geq f(\underline{w}, \ell_1^*)$. Moreover, introducing ℓ_2^* with a similar definition, $g(w_2, \ell_2^*) = \overline{w} = g(\underline{w}, \ell_1^*)$, and since $w_2 \geq \underline{w}$, we must have $\ell_2^* \geq \ell_1^*$, so $\ell_2 \geq \ell_2^* \geq \ell_1^*$. We have $\tilde{\pi}(2) = f(w_2, \ell_2) \geq f(\underline{w}, \ell_1^*)$.

□

It remains to show the upper bound on the consistency of class $\mathcal{I}$ bidding functions. The following lemma will be useful, as it shows that a sequence of slopes can always be extended by the slopes $1/\overline{w}$.

Lemma 19. *Let $R \geq e$.*

- *If $R \geq R_0$, there exists a R-robust class $\mathcal{I}$ bidding function that achieves consistency $\underline{w} + 1$.*
- *Otherwise, there exists an R-robust class $\mathcal{I}$ bidding function that achieves consistency $\underline{w} + \frac{e^{\ell^*} - 1}{\ell^*}$, where ℓ^* is defined as in Theorem 15.*

Proof. For the case $R \geq R_0$, we consider the bidding function b with slopes $(\ell_i)_{i \in \mathbb{Z}}$ defined by:

$$\ell_i = \begin{cases} 1/\underline{w} & \text{if } i < 0 \\ 0 & \text{if } i = 0 \\ 1/\overline{w} & \text{if } i > 0. \end{cases}$$

We want to show that $\tilde{B}(t) \leq R$ for all $t \in \mathbb{R}$. We consider the four following cases:

For all $t \in (-\infty, -1]$, we have $\tilde{B}(t) = R$, due to the constant slope $1/\underline{w}$.

For all $t \in (-1, 0]$, we have $\tilde{B}(t) = (\underline{w} + t + 1)e^{-t/\underline{w}}$, so $\tilde{B}(-1) = R$, $\tilde{B}(0) = \underline{w} + 1$, and the derivative on this interval is $(\tilde{B})'(t) = -\frac{t+1}{\underline{w}} e^{-t/\underline{w}} \leq 0$, so the maximum normalized mass on this interval is R.

For all $t \in (0, 1]$, $\tilde{B}(t) = (\underline{w} + 1 + \overline{w}(e^{t/\overline{w}} - 1))$. This is clearly increasing, so the maximum on this interval is $\tilde{B}(1) = \underline{w} + 1 + R - \underline{w} \leq R$, with equality when $R = R_0$.

For all $t \in (1, +\infty)$, we use Theorem 12. For this purpose we need to check $w_B(2) < \overline{w}$, which holds by $w_B(2) = g(w_B(1), \ell_1) = g(\underline{w} + 1, 1/\overline{w}) \leq g(\overline{w}, 1/\overline{w}) = \overline{w}$.

As a result, this bidding function has robustness R and consistency $\underline{w} + 1$.

For the case $R < R_0$, we consider the bidding function B with sequence of slopes $(\ell_i)_{i \in \mathbb{Z}}$ defined by:

$$\ell_i = \begin{cases} 1/\underline{w} & \text{if } i < 0 \\ \ell^* & \text{if } i = 0 \\ 1/\overline{w} & \text{if } i > 0. \end{cases}$$

We recall that by definition, ℓ^* satisfies $g(\underline{w}, \ell^*) = \overline{w}$. The normalized mass for $t \in (-\infty, -1) \cup (1, +\infty)$ can be treated exactly as in the case $R \geq R_0$.

We first show that $\ell^* \leq 1/\overline{w}$, this comes from monotonicity of g and the inequality

$$g(\overline{w}, \ell^*) \geq g(\underline{w}, \ell^*) = \overline{w} = g(\overline{w}, 1/\overline{w})$$

If $t \in [-1, 0]$, then $\tilde{B}(t) = e^{-t/\underline{w}}(\underline{w} + \frac{e^{(t+1)\ell^*} - 1}{\ell^*})$, so the derivative gives

$$(\tilde{B})'(t) = e^{-t/\underline{w}}(1 - \frac{1}{\underline{w}\ell^*})(e^{(t+1)\ell^*} - 1) \leq 0.$$

Hence the maximum normalized mass on this interval is reached at -1, for a value of $\tilde{B}(-1) = R$.

If $t \in [0,1]$, then $\tilde{B}(t) = e^{-t\ell^*}(\overline{w} + \overline{w}(e^{t/\overline{w}} - 1)) = \overline{w}e^{t(1/\overline{w}-\ell^*)}$. We showed earlier that $\ell^* \leq 1/\overline{w}$, hence $\tilde{B}$ is increasing, therefore, the maximum normalized mass on the interval $[0,1]$ is $\tilde{B}(1) = Re^{-\ell^*} \leq R$.

We showed that in both cases, the given bidding function B of class $\mathcal{I}$ is R-robust and has the stated consistency. □

Theorem 20. *If $R = e + \varepsilon$ and $\varepsilon \to 0$, then the optimal consistency C reached by R-robust class $\mathcal{I}$ bidding functions satisfies*

$$C = e - O(\sqrt{\varepsilon}) + o(\varepsilon).$$

Proof. We first recall that $\underline{w}(R), \overline{w}(R)$ values are $-1/W_{-1}(-1/R)$ and $-1/W_0(-1/R)$.

Let $\varepsilon = R/e - 1$ so that $R = (1+\varepsilon)e$. We have $-1/R = -\frac{1}{e}(1 - \varepsilon + o(\varepsilon))$. Using the series expansion of the Lambert function from [23], we have

$$W_0\left(-\frac{1}{e}(1-u)\right) \underset{u\to 0}{=} -1 + \sqrt{2u} + o(\sqrt{u}), \quad W_{-1}\left(-\frac{1}{e}(1-u)\right) \underset{u\to 0}{=} -1 - \sqrt{2u} + o(\sqrt{u}).$$

Hence, we obtain $\underline{w} = 1 - \sqrt{2\varepsilon} + O(\varepsilon))$ and $\overline{w} = 1 + \sqrt{2\varepsilon} + O(\varepsilon)$. We know that ℓ^* is solution of the equation

$$\underline{w} + \frac{e^\ell - 1}{\ell} = e^\ell \overline{w}.$$

Let us define $\psi(\ell) = \frac{(1-\ell)(e^\ell - 1)}{\ell} = \frac{e^\ell - 1}{\ell} - e^\ell + 1$, we have $\psi(\ell^*) = e^\ell(\overline{w} - 1) - (\underline{w} - 1)$. Since $\lim_{\varepsilon\to 0} \overline{w} = 1$ and $\lim_{\varepsilon\to 0} \underline{w} = 1$, we have that $\psi(\lim_{\varepsilon\to 0} \ell^*) = 0$, hence $\lim_{\varepsilon\to 0} \ell^* = 1$.

We have, on a neighborhood of 0 for ε, that

$$\begin{aligned}
\psi(\ell^*) &= \psi(1 + (\ell^* - 1)) \\
&= \psi(1) + \psi(1)(\ell^* - 1) + o(|\ell^* - 1|) \\
&= (1-e)(\ell^* - 1) + o(|\ell^* - 1|) \\
\text{and} \quad \psi(\ell^*) &= e^{\ell^*}(\overline{w} - 1) - (\underline{w} - 1) \\
&= e(1 + o(1))(\sqrt{2\varepsilon} + O(\varepsilon)) + (\sqrt{2\varepsilon} + O(\varepsilon)) \\
&= (e+1)\sqrt{2\varepsilon} + o(\sqrt{\varepsilon}).
\end{aligned}$$

By equating both expressions, we have that $\ell^* = 1 - \frac{e+1}{e-1}\sqrt{2\varepsilon} + o(\sqrt{\varepsilon})$.

Finally,

$$\begin{aligned}
C &= \overline{w}e^{\ell^*} \\
&= e(1 + \sqrt{2\varepsilon} + O(\varepsilon)) \left(1 - \frac{e+1}{e-1}\sqrt{2\varepsilon} + o(\sqrt{\varepsilon})\right) \\
&= e - \frac{2e\sqrt{2\varepsilon}}{e-1} + o(\sqrt{\varepsilon})) \\
&= e - \frac{2\sqrt{2e}}{e-1}\sqrt{R-e} + o(\sqrt{R-e}).
\end{aligned}$$

□

B Omitted Material from Section 4

Proof of Proposition 4. Let $F(t) := \int_{-\infty}^{t} B(x)\,\mathrm{d}x$. The function $F(t+1) - RB(t)$ is constant on $(-\infty, 0)$. Indeed, this function is continuous and its derivative on S is $B(t+1) - RB'(t) = 0$. Now for $t \to -\infty$, $F(t+1)$ and $B(t)$ both tend to 0, hence the said constant is 0 and we obtain $F(t+1) = RB(t)$, i.e. $\tilde{B}(t) = R$. □

Proof of Lemma 5 ($R \geq 2/\ln(2)$***).*** By definition of the polynomials, the function A is continuous, except at $t = 0$, and is monotone increasing in $[1, +\infty)$. It remains to show that it is monotone increasing in $(-\infty, 0]$ as well as positive.

For this purpose we define $q_k := p_k(0)$ for $k \geq 1$. In order to obtain a recurrence relation for the sequence (q_k). We define the following formal power series for $s \in [0, 1)$:

$$G(z, s) := \sum_{k \geq 0} p_k(s) z^k.$$

By solving for G, we obtain:

$$\begin{aligned}
\frac{\partial G}{\partial s}(z, s) &= \sum_{k \geq 0} p_k'(s) z^k \\
&= \sum_{k \geq 1} x p_{k-1}(s) z^k \\
&= xz \sum_{k \geq 0} p_k(s) z^k \\
&= xz G(z, s).
\end{aligned}$$

For convenience, we denote $\lambda := xz$; multiplying the previous equality by the power series $e^{-\lambda s}$, we obtain

$$e^{-\lambda s} \frac{\partial G}{\partial s}(z, s) - \lambda e^{-\lambda s} G(z, s) = 0,$$

which can be rewritten as:

$$\frac{\partial e^{-\lambda s} G}{\partial s}(z, s) = 0.$$

Integrating this equality from 0 to 1, it follows that

$$G(z, 1) = e^{\lambda} G(z, 0). \tag{16}$$

We also have

$$\begin{aligned}
G(z, 1) &= 1 + \mu x z + \sum_{k \geq 2} p_k(1) z^k \\
&= 1 + \mu x z + \sum_{k \geq 2} p_{k-1}(0) z^k \\
&= 1 + \mu x z + z(G(z, 0) - 1).
\end{aligned} \tag{17}$$

From (16) and (17) we retrieve the desired value, using $a := 1 - \mu x = \overline{w}/R$

$$G(z, 0) = \frac{1 - az}{e^{xz} - z}.$$

Proving that $q_k > 0$ for all $k \geq 0$ is equivalent to showing that the power series $H(z) := G(ze^{\alpha}, 0)$ has positive coefficients. Introduce $\alpha := 1/\overline{w}$ so that $R = e^{\alpha}/\alpha$. Thus we have

$$H(z) = \frac{1 - z}{e^{\alpha z} - z e^{\alpha}}.$$

Next, we compute the coefficients of the power series $1/H(z)$. The power series H can be inverted since its constant coefficient is 1.

$$\frac{1}{H(z)} = \frac{e^{\alpha z} - z e^{\alpha}}{1 - z} = 1 - \sum_{n \geq 1} \tau_n z^n,$$

with $\tau_n = e^{\alpha} - \sum_{j=0}^{n} \frac{\alpha^j}{j!} > 0$. Hence

$$H(z) = \frac{1}{1 - \sum_{n \geq 1} \tau_n z^n} = \sum_{m \geq 0} \left(\sum_{n \geq 1} \tau_n z^n \right)^m = \sum_{m \geq 0} e^{\alpha m} q_m z^m$$

has only positive coefficients, which gives us that $q_k = p_k(0)$ is positive for all $k \geq 0$. From here, a quick induction shows that A is positive and increasing on $(-\infty, 0)$, using the fact that $A(t+1) = RA'(t)$ on this interval. It only remains to prove that A has a finite integral from $-\infty$ to 0. Let us write $F(t) := \int_{-\infty}^{t} B(u)\,\mathrm{d}u$. Noticing that $\int_0^1 p_k(u)\,\mathrm{d}u = R(q_k - q_{k+1})$ (consequence from $p_k' = p_{k-1}$) we obtain:

$$F(0) = \sum_{k \geq 1} R(q_k - q_{k+1}) \leq Rq_1.$$

Hence, A is positive, increasing and has a finite integral. Thus it is a valid bidding function. □

Proof of Theorem 6. Recall the notation $x := 1/R$ and $\mu := R - \overline{w}$. We will show that $\tilde{B}(t)$ is at most R.

Recall that A is solution of the delay differential equation $A(t+1) - RA'(t) = 0$ on $(-\infty, 0)$, is continuous on the same interval, and differentiable on this interval except at integer points. Hence, by Theorem 4, we have

$$\forall t \in (-\infty, 0), \tilde{A}(t) = \frac{F(t+1)}{A(t)} = R.$$

For the consistency, we want it to be reached at $t = 0$. Taking $t \to 0^-$, we obtain $F(1) = RA(0^-) = \mu$, hence, using $A(0) = 1$:

$$\tilde{A}(0) = F(1) = \mu = R - \overline{w}.$$

On the interval $[0, 1]$, the worst normalized ratio is reached for $t = 1$. since F is increasing,

$$\begin{aligned}
\forall t \in [0,1], \tilde{A}(t) &\leq \tilde{A}(1) \\
&= F(2)/A(1) \\
&= F(1) + \overline{w}(e^{1/\overline{w}} - 1) \\
&= R - \overline{w} + R - \overline{w} \\
&= 2(R - \overline{w}) \\
&\leq R.
\end{aligned}$$

The last inequality holds by the assumption $R \geq 2/\ln(2)$. For $t \geq 1$, it holds that $\tilde{A}(t) \leq R$ directly from Theorem 12. □

The following is the definition of the bidding function A for $R < 2/\ln(2)$.

Definition 21. *Given $R < 2/\ln 2$, define $\alpha := 1/\overline{w}$ and consider $y \in (0, 1]$ solution to the equation $\alpha = y + \ln(2 - y)$. Let $L := 1 - y/\alpha$, and $\mu := e^y/\alpha$. Define the function A*

$$A(t) = \begin{cases} e^{(t-L)/\overline{w}} & \text{if } L \leq t \\ 1 & \text{if } 0 \leq t \leq L, \\ p_k(s) & \text{if } t \leq 0 \text{ for } k = \lceil -t \rceil, s = \lfloor -t \rfloor. \end{cases}$$

Here the polynomial p_k is defined inductively as follows:

$$p_0(s) = \begin{cases} 1 & \text{if } 0 \leq s \leq L, \\ e^{\alpha(s-L)} & \text{if } L \leq s \leq 1 \end{cases}$$

$$p_1(s) = x\left(\mu - \int_s^1 p_0(u)\,\mathrm{d}u\right)$$

$$p_{k+1}(s) = p_k(0) - x\int_s^1 p_k(u)\,\mathrm{d}u.$$

Proof of Lemma 5 ($R < 2/\ln(2)$). The case $R = e$ is immediate since it gives $A(t) = e^t$. We can thus assume $e < R \leq 2/\ln 2$. We set $x := 1/R$, $\alpha := 1/\overline{w}$, $L := 1 - y/\alpha$, and $\mu := e^y/\alpha$. Since $\overline{w}e^{1/\overline{w}} = R$, we have $R = e^\alpha/\alpha$. Moreover, the equation defining y gives $e^\alpha = e^y(2 - y)$. Hence

$$R = \frac{e^\alpha}{\alpha} = \frac{e^y(2-y)}{\alpha} = (2-y)\mu.$$

We start with the positivity of the function, as in the proof of Theorem Theorem 6. It suffices to prove positivity at the left endpoints, hence we define $q_k := p_k(0)$ for $k \geq 1$, and we will show that $q_k > 0$ for every $k \geq 1$. we introduce the formal power series

$$G(z, s) := \sum_{k \geq 1} p_k(s) z^k.$$

Note that the sum starts at $k = 1$ instead of 0, because p_0 is not constant anymore as in the proof of Theorem Theorem 6. The computation is formal in z, and the derivative is only with respect to s. Using $p_1' = xp_0$ and $p_k' = xp_{k-1}$ for $k \geq 2$, it follows that

$$\frac{\partial G}{\partial s}(z, s) = xz(p_0(s) + G(z, s)).$$

Let $\lambda := xz$. Multiplying the above by $e^{-\lambda s}$, we obtain

$$\frac{\partial}{\partial s}(e^{-\lambda s} G(z, s)) = \lambda e^{-\lambda s} p_0(s).$$

Integrating from 0 to 1, this gives

$$e^{-\lambda} G(z, 1) - G(z, 0) = \lambda \int_0^1 e^{-\lambda u} p_0(u) \, \mathrm{d}u.$$

We now express $G(z, 1)$ in terms of $G(z, 0)$. Since $p_1(1) = x\mu$ and $p_k(1) = p_{k-1}(0)$ for $k \geq 2$, we have

$$G(z, 1) = x\mu z + zG(z, 0) = \lambda\mu + zG(z, 0).$$

Substituting this in the above equation and solving for $G(z, 0)$, we obtain

$$G(z, 0) = \lambda \frac{\mu e^{-\lambda} - \int_0^1 e^{-\lambda u} p_0(u) \, \mathrm{d}u}{1 - ze^{-\lambda}}.$$

By rescaling z, define $H(z) := G(e^\alpha z, 0)$. Since $R = e^\alpha/\alpha$, the parameter $\lambda = xz$ becomes αz after this rescaling. We compute the integral in the numerator using the two pieces of p_0:

$$\int_0^1 e^{-\alpha z u} p_0(u) \, \mathrm{d}u = \int_0^L e^{-\alpha z u} \, \mathrm{d}u + \int_L^1 e^{-\alpha z u} e^{\alpha(u-L)} \, \mathrm{d}u.$$

The first integral is equal to $(1 - e^{-\alpha z L})/(\alpha z)$. For the second one, using $\alpha(1 - L) = y$, we obtain

$$\int_L^1 e^{-\alpha z u} e^{\alpha(u-L)} \, \mathrm{d}u = \frac{e^{y - \alpha z} - e^{-\alpha z L}}{\alpha(1 - z)}.$$

Using also $\mu = e^y/\alpha$, a direct simplification gives $H(z) = \frac{J(z)}{D(z)}$, where

$$D(z) := e^{\alpha z} - ze^\alpha \qquad J(z) := \frac{e^{yz} - (1 - z)e^{\alpha z} - z^2 e^y}{1 - z}.$$

Since $H(z) = \sum_{k \geq 1} e^{\alpha k} q_k z^k$, it remains to prove that all coefficients of H are positive. For the denominator $D(z)$ we have

$$\frac{D(z)}{1 - z} = 1 - \sum_{n \geq 1} \tau_n z^n,$$

where $\tau_n = e^\alpha - \sum_{j=0}^n \frac{\alpha^j}{j!} = \sum_{j \geq n+1} \frac{\alpha^j}{j!} > 0$. Therefore, since the above series has non zero constant coefficient, we can reverse it and obtain that

$$\frac{1 - z}{D(z)} = \frac{1}{1 - \sum_{n \geq 1} \tau_n z^n} = \sum_{m \geq 0} \left(\sum_{n \geq 1} \tau_n z^n \right)^m$$

has non negative coefficients. Hence, if

$$\frac{1}{D(z)} = \sum_{n\geq 0} i_n z^n,$$

then the sequence (i_n) is positive and non decreasing, because $1/D(z)$ is obtained from $(1-z)/D(z)$ by multiplying by $1/(1-z)$. Next we write the numerator $J(z)$ as

$$J(z) = \theta z - \sum_{n\geq 2} \sigma_n z^n,$$

where $\theta = 1 + y - \alpha = 1 - \ln(2-y) > 0$, and, for $n \geq 2$,

$$\sigma_n = \sum_{m\geq n+1} \frac{y^m}{m!} + \frac{\alpha^n}{n!} > 0.$$

We will use the identity $\sum_{n\geq 2} A_n = \theta$. Indeed:

$$\sum_{n\geq 2} \sigma_n = \sum_{m\geq 3} (m-2)\frac{y^m}{m!} + \sum_{n\geq 2} \frac{\alpha^n}{n!}.$$

In the above expression the first sum is equal to $(y-2)e^y + 2 + y$, and the second one is equal to $e^\alpha - 1 - \alpha$. Thus, using that $e^\alpha = e^y(2-y)$,

$$\sum_{n\geq 2} \sigma_n = (y-2)e^y + 2 + y + e^\alpha - 1 - \alpha = 1 + y - \alpha = \theta. \tag{18}$$

We can now prove the positivity of the coefficients of H. Let $H(z) = \sum_{k\geq 1} h_k z^k$. Since $H = J/D$, we have

$$h_k = \theta i_{k-1} - \sum_{n=2}^{k} \sigma_n i_{k-n} = \sum_{n=2}^{k} \sigma_n \left(i_{k-1} - i_{k-n}\right) + i_{k-1} \sum_{n>k} \sigma_n$$

using (18). The first term is non negative because (i_n) is non decreasing. The second term is strictly positive because $i_{k-1} > 0$ and $\sigma_n > 0$ for every $n \geq 2$. Hence $h_k > 0$ for every $k \geq 1$. Since $h_k = e^{\alpha k} q_k$, we have $q_k > 0$ for every $k \geq 1$. From here, positivity of A on $(-\infty, 0)$ follows by induction. Indeed, if p_k is positive on $[0,1]$, then $p'_{k+1} = x p_k > 0$, so p_{k+1} is increasing. Its minimum is reached at 0 and is equal to $q_{k+1} > 0$. Hence, every p_k is positive.

It remains to verify that $\int_{-\infty}^{0} A(x)\,\mathrm{d}x$ is finite. From the recurrence of Definition 21, evaluated at $s = 0$, we have

$$q_{k+1} = q_k - x \int_0^1 p_k(u)\,\mathrm{d}u,$$

so

$$\int_0^1 p_k(u)\,\mathrm{d}u = R(q_k - q_{k+1}).$$

Since all q_k are positive, summing both parts of the above equation gives

$$\int_{-\infty}^{0} A(u)\,\mathrm{d}u = \sum_{k\geq 1} \int_0^1 p_k(u)\,\mathrm{d}u \leq R q_1 < +\infty.$$

Thus A is a valid bidding function. □

Proof of Theorem 7. As for Theorem 6, we can use Theorem 4, which gives

$$\tilde{A}(t) = R, \quad \forall t \in (-\infty, 0).$$

Taking $t \to 0^-$ in the equality $F(t+1) = RA(t)$, we have $F(1) = RA(0^-) = \mu$. Since $A(0) = 1$, the consistency is

$$\tilde{A}(0) = F(1) = \mu = \frac{R}{2-y}.$$

It remains to verify that $\tilde{A}(t) \leq R$ for $t \geq 0$. If $0 \leq t \leq L$, then $A(t) = 1$, and

$$F(t+1) = F(1) + \int_1^{t+1} e^{\alpha(u-L)}\,\mathrm{d}u = \mu + \frac{e^{\alpha(t+1-L)} - e^{\alpha(1-L)}}{\alpha}.$$

Using $\mu = e^y/\alpha$ and $\alpha(1-L) = y$, this becomes $F(t+1) = \frac{e^{y+\alpha t}}{\alpha}$. Thus

$$\tilde{A}(t) = \frac{e^{y+\alpha t}}{\alpha}.$$

Since $t \leq L$, we have $y + \alpha t \leq y + \alpha L = \alpha$, and therefore $\tilde{A}(t) \leq e^\alpha/\alpha = R$. Finally, if $t \geq L$, then $A(t) = e^{\alpha(t-L)}$, and the same computation gives $F(t+1) = e^{\alpha(t+1-L)}/\alpha$. Hence

$$\tilde{A}(t) = \frac{e^{\alpha(t+1-L)}}{\alpha e^{\alpha(t-L)}} = \frac{e^\alpha}{\alpha} = R.$$

We have proved that the robustness is at most R, while the consistency is exactly $R/(2-y)$. This proves the theorem. □

Proof of Corollary 8. We first recall that if R is close enough to e, by Theorem 7, the consistency of the bidding function A is $R/(2-y)$ where y is solution of the equation $1/\overline{w} = y + \ln(2-y)$. Set $\varepsilon := R/e - 1$ and $s = \overline{w} - 1$. Since $R \to e$, we have $\varepsilon \to 0$ and $s \to 0$. From $\overline{w}e^{1/\overline{w}} = R$, it follows that

$$1 + \varepsilon = (1+s)\exp\left(\frac{1}{1+s} - 1\right).$$

Expanding the right-hand side at $s = 0$, we obtain

$$(1+s)\exp\left(\frac{1}{1+s} - 1\right) = 1 + \frac{s^2}{2} + O(s^3).$$

Therefore $\varepsilon = \frac{s^2}{2} + O(s^3)$ and hence $s = \sqrt{2\varepsilon} + O(\varepsilon)$. We have that y tends to 1 by below, hence we can write $y = 1 - q$ with $q \to 0^+$. From the definition of y we have

$$1 - q + \ln(1+q) = \frac{1}{1+s}.$$

Expanding both sides gives

$$1 - \frac{q^2}{2} + O(q^3) = 1 - s + O(s^2).$$

Thus $q^2/2 = s + o(s)$, so $q = \sqrt{2s} + O(s)$. Using the expansion of s, this gives

$$q = \sqrt{2\sqrt{2\varepsilon}} + O(\varepsilon^{1/2}) = 2^{3/4}\varepsilon^{1/4} + O(\varepsilon^{1/2}).$$

Finally, since $2 - y = 1 + q$, we have

$$\frac{R}{2-y} = \frac{e(1+\varepsilon)}{1+q} = e(1+\varepsilon)\left(1 - q + O(q^2)\right).$$

Since $q = O(\varepsilon^{1/4})$, we have $q^2 = O(\varepsilon^{1/2})$. Therefore

$$\frac{R}{2-y} = e - eq + O(\varepsilon^{1/2}).$$

Substituting the expansion of q, we obtain

$$\frac{R}{2-y} = e - 2^{3/4} e \varepsilon^{1/4} + O(\varepsilon^{1/2}).$$

Since $\varepsilon = (R-e)/e$, this is exactly

$$\frac{R}{2-y} = e - 2^{3/4} e^{3/4} (R-e)^{1/4} + O\left((R-e)^{1/2}\right).$$

□

C Omitted Materials from Section 5

Proof of Lemma 9. Let $a \geq 1$ and $R \geq e$. Let B be a bidding function of robustness R and consistency C. Since B is C consistent, we have that $\inf_{t \in \mathbb{R}} \tilde{B}(t) = C$. Then, for any $\varepsilon > 0$, we can shift the bidding function so that $\tilde{B}(0) \leq C + \varepsilon$.

For all $k \in \mathbb{Z}$, define $x_k := a \int_{k/a}^{(k+1)/a} B(u)\,\mathrm{d}u$. Note that we have $x_k \geq B(k/a) > 0$, and that we can suppose, without loss of generality, that $x_0 = 1$, up to dividing the bidding function B by the constant factor x_0. From the robustness conditions, we obtain

$$\begin{aligned}
\frac{1}{a} \sum_{j=-N}^{\min(M,k+a-1)} x_j &\leq \frac{1}{a} \sum_{j=-\infty}^{k+a-1} x_j \\
&= \int_{-\infty}^{k/a+1} B(u)\,\mathrm{d}u \\
&\leq RB(k/a) \\
&\leq Rx_k.
\end{aligned}$$

The same calculations, combined with the fact that $x_0 = 1$, give us that

$$\frac{1}{a} \sum_{j=-N}^{\min(M,a-1)} x_j \leq \frac{1}{a} \sum_{j=-\infty}^{a-1} x_j \leq C + \epsilon.$$

Thus the objective value of $P^R_{a,N,M}$ is at most $C + \varepsilon$. □

We first prove two intermediate lemmas that will be important in the proof of the Theorem.

Lemma 22. *Let $a \geq 1$ and $R \geq e$. The smallest root in modulus of $X^a - aR(X-1)$ is its unique real root $\rho \in (1, 1 + 1/a)$. Moreover, ρ is a root of multiplicity* 1.

Proof. We first prove that ρ exists. Let $f(x) = x^a - aR(x-1)$, we have $f(1) = 1$ and

$$f(1 + \frac{1}{a}) = (1 + \frac{1}{a})^a - R \leq e - R < 0.$$

Hence ρ exists; we now show uniqueness of ρ. Indeed, $f'(x) = a(x^{a-1} - R)$ is negative on $[1, 1 + 1/a]$. Since f' is non-zero on this interval, ρ is of multiplicity 1.

It remains to verify that there does not exist any root of smaller modulus. Suppose, by way of contradiction, that there exists a root $\eta = v\rho$ with $|v| \leq 1$ and $v \neq 1$. We have

$$v^a \rho^a = aR(v\rho - 1).$$

Using that $\rho^a = aR(\rho - 1)$, we obtain the following

$$
\begin{aligned}
& v^a = \frac{v\rho - 1}{\rho - 1} \\
\iff & v^a - 1 = \frac{\rho(v-1)}{\rho - 1} \\
\iff & \frac{v^a - 1}{v - 1} = \frac{\rho}{\rho - 1}.
\end{aligned}
$$

The above can be rewritten as $\sum_{i=0}^{a-1} v^i = \frac{\rho}{\rho-1}$. Note that left hand side has modulus at most a by the triangular inequality ($|v^i| \leq 1, \forall i$) , and the right hand side has modulus at least $1 + a$ since $\rho \in [1, 1 + 1/a]$, a contradiction. □

Lemma 23. *Let $a \geq 1$. The equation*

$$
aRb_n^{(N)} = 1 + \sum_{m=1}^{\min(N, n+a-1)} b_m^{(N)} \quad (1 \leq n \leq N). \tag{19}
$$

has a family of positive solutions indexed by $N \geq 1$ such that for all $n \geq 1$, we have

$$
b_n^{(N)} \to_{N \to +\infty} (\rho - 1)\rho^{n-1},
$$

where ρ is defined in Lemma 22.

Proof. Let us fix $a \geq 1$.

We start by justifying the existence of a positive solution $(b_n^{(N)})$. We first introduce an infinite sequence that is solution a to the unbounded version of (19). Let ρ be as defined in Lemma 22. The candidate infinite sequence is $(\sigma_n)_{n \geq 1}$ where

$$
\sigma_n := (\rho - 1)\rho^{n-1}.
$$

This sequence indeed satisfies, using the definition of ρ

$$
\begin{aligned}
1 + \sum_{j=1}^{n+a-1} \sigma_j &= 1 + (\rho - 1) \sum_{j=0}^{n+a-2} \rho^j \\
&= \rho^{n+a-1} \\
&= aR\sigma_n.
\end{aligned} \tag{20}
$$

Consider the map $F_N : \mathbb{R}_+^N \to \mathbb{R}_+^N$ defined by, $\forall 1 \leq n \leq N$

$$
(F_N(b))_n := \frac{1}{aR} \left(1 + \sum_{m=1}^{\min(N, n+a-1)} b_m \right).
$$

First note that this map is monotone, in the sense that if $(u_n)_{n \leq N} \leq (v_n)_{n \leq N}$, then $F_N((u_n)_{n \leq N}) \leq F_N((u_n)_{n \leq N})$, where $\leq$ corresponds to term by term comparison.

Next, recall that σ is solution of (20), hence we obtain, denoting by $\sigma^{(N)}$ the finite sequence $(\sigma_n)_{n \leq N}$ that

$$
F_N(\sigma^{(N)}) \leq \sigma^{(N)}.
$$

Moreover, $F_N(0) = (\frac{1}{aR}, ..., \frac{1}{aR}) > 0$, hence the sequence $0, F_N(0), F_N^{(2)}(0), \ldots$ is monotone. It is also bounded by $\sigma^{(N)}$ from above. Indeed, we have $0 \leq \sigma^{(N)}$, and using monotonicity of F_N, we obtain $F_N^{(k)}(0) \leq F_N^{(k)}(\sigma^{(N)}) \leq \sigma^{(N)}$.

Thus, $0, F_N(0), F_N^{(2)}(0), \ldots$ is monotone and bounded from above, so it converges to a positive fixed point that we denote $b^{(N)}$, and which is smaller than $\sigma^{(N)}$. It remains to compute the limit of $b_n^{(N)}$ as N tends to $+\infty$.

Let $n \geq 1$. For all $N \geq n + a - 1$, we have

$$aRb_n^{(N)} = 1 + \sum_{m=1}^{n+a-1} b_m^{(N)}.$$

We define b_n^* as the limit of $b_n^{(N)}$ as N tends to $+\infty$. Note that for every n and $N \geq n + a - 1$, $(F_N^{(k)}(0))_n \leq \sigma_n^{(N)}$. As N tends to $+\infty$, we obtain that

$$b_n^* \leq \sigma_n. \tag{21}$$

We also have

$$aRb_n^* = 1 + \sum_{m=1}^{n+a-1} b_m^*,$$

which is exactly (20). The goal is now to prove that (21) is tight. We define the difference of these two sequences, i.e. $d_n := \sigma_n - b_n^* \geq 0$, (d_n) is solution of the homogeneous version of (20):

$$\forall n \geq 1, \quad aRd_n = \sum_{m=1}^{n+a-1} d_m.$$

Subtracting this equation for n and $n + 1$ gives that (d_n) is solution of

$$\forall n \geq 1, \quad aR(d_{n+1} - d_n) = d_{n+a}.$$

This linear recurrent equation has the characteristic polynomial $P(X) = X^a - aR(X - 1)$. Since ρ is a root of multiplicity 1 of $P(X)$, we obtain that there exists a constant K and polynomials functions $(K_i)_{i \in I}$ for each other root $(\rho_i)_{i \in I}$ of

$$d_n = K\rho^n + K_i(n)\rho_i^n.$$

However, $d_n = \sigma_n - b_n^*$ is bounded by σ^n, so $d_n = O(\rho^n)$. ρ being the root of smallest modulus of P, we must have that every K_i is equal to the null polynomial. Otherwise, we would have $|d_n| \sim |\rho_i|^n$ for some $i \in I$. Thus $d_n = K\rho^n$ for some $K \geq 0$. Let us suppose towards a contradiction that $K > 0$. It follows that

$$aRK\rho^n = \sum_{m=1}^{n+a-1} K\rho^m.$$

Dividing by K, and using that $aR = \frac{\rho^a}{\rho - 1}$, we obtain on the left hand side $\frac{\rho^{a+n}}{\rho-1}$ and on the right hand side $\frac{\rho^{a+n} - \rho}{\rho - 1}$, which are different, a contradiction. Thus $K = 0$ and $b_n^* = \sigma_n$ for all n, hence $b^* = \sigma$, which completes the proof. □

Proof of Theorem 11. We construct a family of feasible solutions of the dual, indexed by a and N. Their objective values will converge to $R - \overline{w}$.

Fix $a \geq 1$ and $N \geq 1$, and choose $M = a - 1$. Note that the term $\mathbf{1}_{\{k \leq a-1\}}$ is equal to 1 for every $k \in \{-N, \ldots, a - 1\}$.

We first consider the sequence $b_1^{(N)}, \ldots, b_N^{(N)}$ from Lemma 23. We specify the values of the dual variables. Let

$$\begin{aligned} \beta_{-n} &:= b_n^{(N)} && (1 \leq n \leq N), \\ \beta_k &:= 0 && (0 \leq k \leq a - 1). \end{aligned}$$

For γ, let $\gamma_i = 0$ for $-N \leq i \leq -1$. For $0 \leq i \leq a - 2$, define

$$\gamma_i := \frac{1}{a} \sum_{m=i+1}^{a-1} \left(1 + \sum_{j=1}^{a-1-m} \beta_j^{(N)}\right), \tag{22}$$

keeping the boundary convention $\gamma_{a-1} = 0$. Finally, we specify

$$\lambda_{a,N} := 1 + \frac{1}{a} \sum_{m=1}^{a-1} (a-m) b_m^{(N)}.$$

We will prove that the above specify a feasible dual solution.

First consider a negative index $-N \leq k \leq -1$. Since $\gamma_k = \gamma_{k-1} = 0$, the dual constraint that corresponds to k becomes

$$\frac{1}{a} + \frac{1}{a} \sum_{j=\max(k+1-a,-N)}^{-1} \beta_j \geq R\beta_k^{(N)}.$$

Substituting $j = -m$ and $n = -k$, this becomes

$$\frac{1}{a} \left(1 + \sum_{m=1}^{\min(N,n+a-1)} b_m^{(N)} \right) \geq R b_n^{(N)}.$$

By (19) from Lemma 23, the above inequality holds with equality.

Next, consider $1 \leq k \leq a-1$. Then $\beta_k = 0$, and there is no λ-term. The left-hand side of the dual constraint for k is

$$\frac{1}{a} + \frac{1}{a} \sum_{j=k+1-a}^{-1} \beta_j + \gamma_k - \gamma_{k-1}.$$

In the above expression, the sum over β is equal to $\sum_{m=1}^{a-1-m} b_m^{(N)}$. By Equation (22), we have

$$\gamma_{k-1} - \gamma_k = \frac{1}{a} \left(1 + \sum_{m=1}^{a-1-m} b_n^{(N)} \right).$$

Thus the left-hand side of the constraint is equal to 0, so it holds with equality.

It remains to verify the case $k = 0$. Since $\beta_0 = 0$ and $\gamma_{-1} = 0$, the left-hand side of the constraint is

$$\frac{1}{a} + \frac{1}{a} \sum_{m=1}^{a-1} b_m^{(N)} + \gamma_0.$$

Using Equation (22), this is

$$\frac{1}{a} + \frac{1}{a} \sum_{m=1}^{a-1} b_m^{(N)} + \frac{1}{a} \sum_{m=1}^{a-1} \left(1 + \sum_{j=1}^{a-1-m} b_j^{(N)} \right).$$

The term $b_m^{(N)}$ appears $a-m$ times in total in the above expression. Therefore the left-hand side of the constraint for $k = 0$ becomes

$$1 + \frac{1}{a} \sum_{m=1}^{a-1} (a-m) b_m^{(N)} = \lambda_{a,N}.$$

Hence the dual constraint for $k = 0$ also holds with equality. We conclude that (β, γ, λ) is a feasible solution to the dual $D_{a,N,M}^R$.

The best lower bound is obtained as the parameters a and N tend to $+\infty$. We now compute the limit of the objective $\lambda_{a,N}$. For fixed a, let ρ_a be the root from Lemma 22. We denote $\lambda_a = \lim_{N \to \infty} \lambda_{a,N}$

and obtain

$$\begin{aligned}\lambda_a &= 1 + \frac{1}{a}\sum_{n=1}^{a-1}(a-n)\rho_a^n - \rho_a^{n-1} \\ &= 1 + \frac{1}{a}\sum_{n=1}^{a-1}\rho_a^n - (a-1) \\ &= \frac{1}{a}\sum_{n=0}^{a-1}\rho_a^n,\end{aligned} \tag{23}$$

By defining θ_a so that $\rho_a = e^{\theta_a/a}$, the definition of ρ_a gives

$$e^{\theta_a} = aR\left(e^{\theta_a/a} - 1\right).$$

As $a \to \infty$, since ρ_a tends to 1, θ_a/a tends to 0 and this gives $e^\theta = R\theta$ where $\theta = \lim_{a\to\infty}\theta_a$. Since $\rho_a \in (1, a/(a-1))$, the relevant limit satisfies $\theta \in (0, 1]$, and the equation $e^\theta/\theta = R$ gives $\theta = 1/\overline{w}$.

Finally, (23) becomes a Riemann sum:

$$\lambda_a = \frac{1}{a}\sum_{m=0}^{a-1} e^{\theta_a m/a} \longrightarrow \int_0^1 e^{\theta x}\,\mathrm{d}x = \frac{e^\theta - 1}{\theta}.$$

From the fact that $e^\theta/\theta = R$ and $\theta = 1/\overline{w}$, we obtain

$$\frac{e^\theta - 1}{\theta} = \frac{e^\theta}{\theta} - \frac{1}{\theta} = R - \overline{w}.$$

Therefore every R-robust bidding sequence has consistency at least $R - \overline{w} - \varepsilon$. This proves the theorem. □

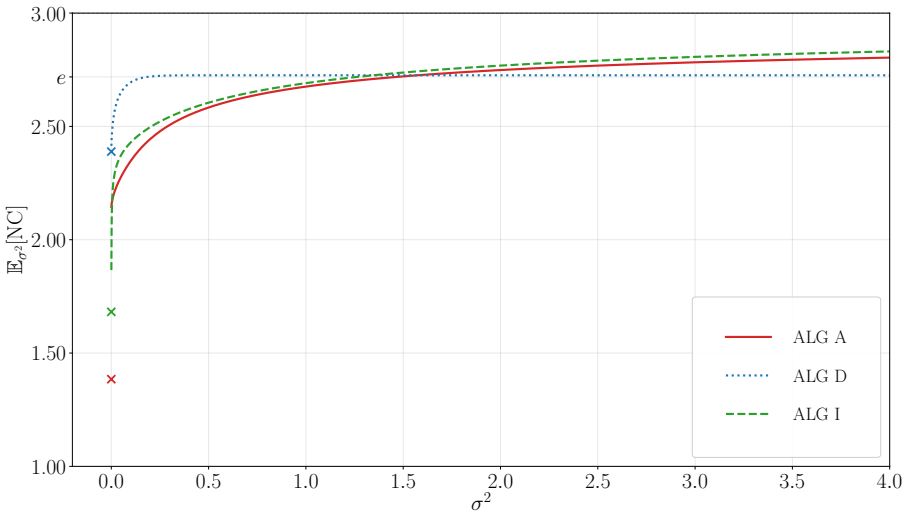
D Additional numerical analysis

Figure 4 depicts the expected normalized cost of the different bidding algorithms for $R = 3$ and $R = 2.73$, as a function of the variance σ^2, following the same experimental setting as in the main paper. As R gets closer to e, algorithm A improves over D on a wider interval of error, namely for $\sigma^2 \lesssim 40$. On the other hand, A improves over I only when the prediction oracle is fairly accurate.

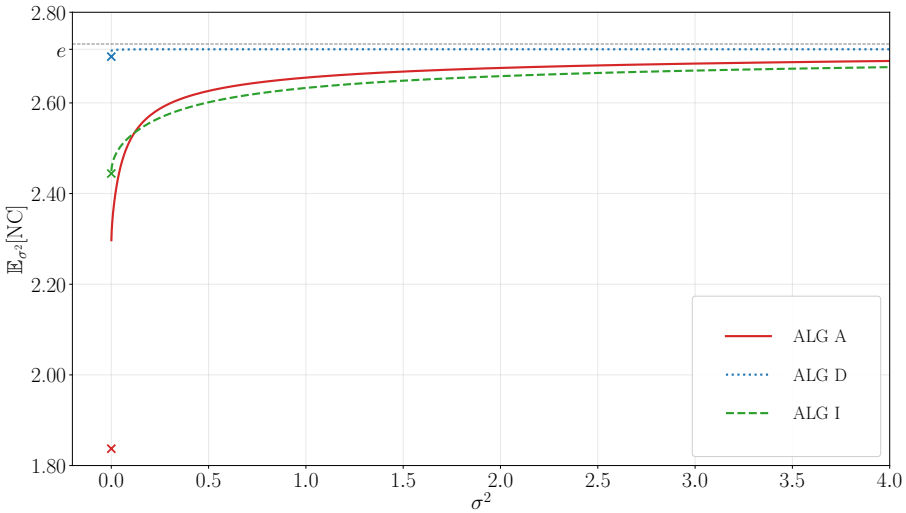
E Omitted material from Section 7

Our general approach is the following. First, we compute an approximation $\widetilde{\mathrm{opt}}_k$ of opt_k, using the Python class `KMedoids` from the package `sklearn_extra.cluster`[3], and their corresponding sets $\widetilde{F}_k$. These values $\widetilde{F}_n, \ldots, \widetilde{F}_1$ form a universe $\mathcal{U}$ of increasing values. A bidding sequence is projected on $\mathcal{U}$, in the following sense. It translates into a set of indices $\mathcal{K} \subseteq \{1, \ldots, n\}$ such that for every $i \in \mathcal{K}$, there is a bid x in this sequence with $\widetilde{F}_i \le x < \widetilde{F}_{i-1}$. The algorithm constructs sets F_k for every $k \in \mathcal{K}$. First, it initializes $F_n = \widetilde{F}_n = \mathcal{U}$. Then for every $k \in \mathcal{K} \setminus \{n\}$ in decreasing order, F_k is obtained from the previously built set F_ℓ, by projecting each point $p \in \widetilde{F}_k$ to the closest point in F_ℓ, hence approximating the near-optimal set $\widetilde{F}_k$ with the only available points due to the *incremental* constraint of the problem. Intermediate sets F_i for $k < i < \ell$ are simply set to F_k. Here, we adhere to the original description of the algorithm of [21], where an incremental solution $F_1 \subseteq F_2 \subseteq \ldots \subseteq F_n$ only needs to satisfy $|F_i| \le i$. Another option would have been to fill the sets greedily with points so to satisfy $|F_i| = i$ for all i.

[3] Released under the 3-Clause BSD license



$R = 3$



$R = 2.73$

Figure 4: Expected normalized cost as function of σ^2 for $R = 3$ and $R = 2.73$.

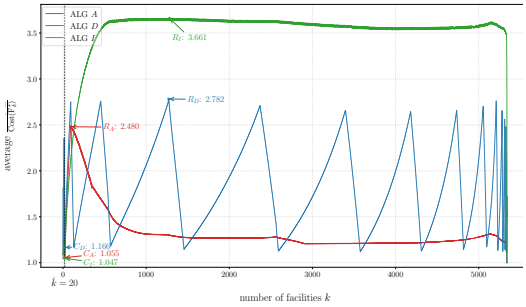
The theoretical performance guarantee of the bidding sequence translates to a similar guarantee for the incremental medians problem, in the following sense. Suppose that the approximation ratio $\widetilde{\mathrm{opt}}_k/\mathrm{opt}_k$ is upper bounded by some factor $c > 1$. Unfortunately the documentation does not specify it [1], but since we use the `kmeans++` initialization option, we can assume $c = O(\log n)$, see [9]. If the bidding sequence guarantees a normalized cost $\leq \beta$, then the produced incremental medians solution guarantees a normalized cost $2\beta \cdot c$, where c comes from the approximation of $\widetilde{\mathrm{opt}}$, and the factor 2 comes from the triangular inequality of the distance function and the use of the projection in the algorithm.

Lemma 24. *Let $B = (X_i)_{i\in\mathbb{Z}}$ be a (randomized) sequence for the online bidding problem, whose robustness is at most β and whose consistency is at most C. Then, for every k, there exists a projection onto the cost set $U = \{\mathrm{cost}(F_k^*), k \leq n\}$ of B such that there exists an incremental median algorithm that is 2β-robust and achieves a normalized cost of at most $2C$ at point k.*

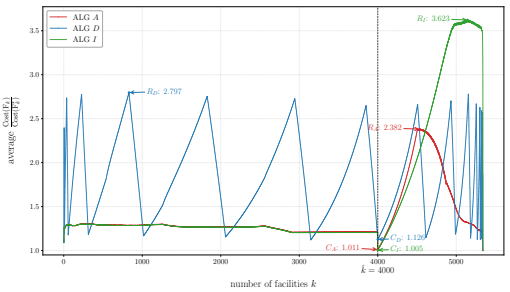
Proof. We denote by $t \in \mathbb{R}$ the value such that $\mathrm{nc}(\mathrm{B}, t) \leq C$. Let $k \leq n$, and define $\widetilde{\mathrm{B}} = (X_i \cdot \frac{\mathrm{cost}(F_k^*)}{t})_{i \in \mathbb{Z}}$. By scale invariance, this sequence remains β-competitive, and moreover $\mathrm{nc}(\widetilde{\mathrm{B}}, \mathrm{cost}(F_k^*)) \leq C$. It then suffices to consider the projection of $\widetilde{\mathrm{B}}$ onto the set U, as we defined. The conclusion then follows from [21, Lemma 5]. □

We conclude this appendix with a sensitivity analysis of the experiments presented in the main paper. We show that the observations made in Section 7 remain stable when varying a single parameter of the experiment presented in Figure 3. Each figure focuses on a different parameter: $\hat{k}$, R, and the dataset itself. Figure 5 depicts the empirical average approximation ratios of the various algorithms for two different predicted numbers of facilities, $\hat{k} = 20$ and $\hat{k} = 4000$, respectively. Figure 6 depicts the sensitivity of the results to different robustness requirements ($R = 3$ and $R = 5$). Figure 7 presents the empirical average approximation ratios on two different datasets. For two different random seeds, we generate 500 points uniformly at random in $[0, 10]^2$, with pairwise Euclidean distances. We run the same experiment on these random graphs. The plots exhibit the same relative performance trends as those reported in the main paper.

Sensitivity to the predicted parameter $\hat{k}$



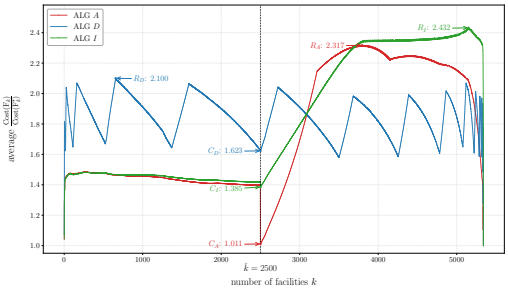
Average empirical approximation ratios for $\hat{k} = 20$.



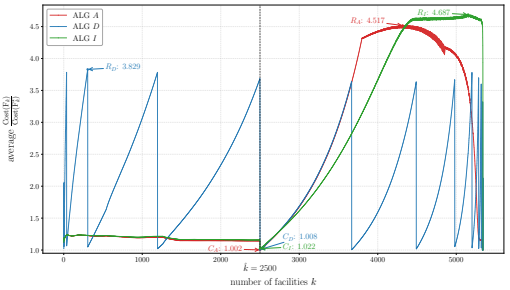
Average empirical approximation ratios for $\hat{k} = 4000$.

Figure 5: Average empirical approximation ratios of the incremental median algorithms for different values of $\hat{k}$.

Sensitivity to the robustness parameter R



Average empirical approximation ratios for $R = 3$.



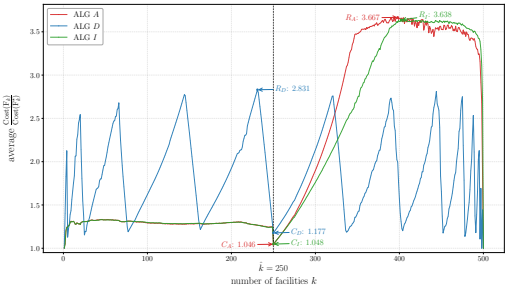
Average empirical approximation ratios for $R = 5$.

Figure 6: Average empirical approximation ratios of the incremental median algorithms for different values of R.

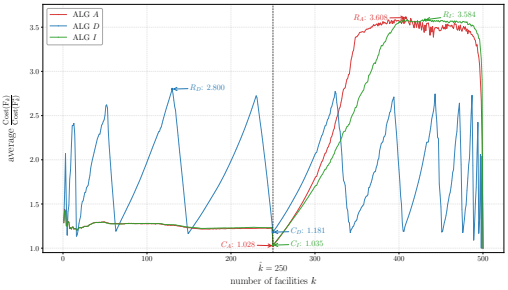
F Declaration of Computing Resources

All experiments were run on a standard laptop with processor 11th Gen Intel Core i5-1135G7 × 8 with 16 GB of memory. For the computational lower bound, the execution time took approximately 4 seconds per point in the curve. That is a total of 5 minutes approximately.

Sensitivity to the dataset



Average empirical approximation ratios for the random graph with seed = 101.



Average empirical approximation ratios for the random graph with seed = 202.

Figure 7: Average empirical approximation ratios of the incremental median algorithms for different random graphs.

For our experiments on the incremental-median problem, there is a necessary preprocessing offline phase that took approximately 2 days on the above equipment. After this preprocessing phase, each experiment required time around 2 minutes.

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: We provide theoretical results with full proofs given in the supplementary material, as well as an experimental evaluation of an application of our main contribution.

   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: [N/A]

   Justification: Our theoretical results hold in the most general setting, and do not rely on any hidden assumptions. Likewise, we do not rely on any assumptions in the experimental application to the incremental median problem.

   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: The proofs of all theorems and statements are provided in the technical appendix. All assumptions are stated in the statements.

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: [Yes]

Justification: The full experimental setup is described in the paper and the experiments are based on publicly available data. The code and data are provided in the supplemental material. We also provided a README file that explains how to execute the code and reproduce the 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: [Yes]

   Justification: The code on which the experimental evaluation is based is given in the supplementary material. It will be made available as open source for the broad public at a later stage.

   Guidelines:

   - The answer [Yes] 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: [Yes]

   Justification: We do not rely on any training data. We explain how predictions are generated in the discussion of the experimental setting.

   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: [No]

   Justification: For incremental median, we report only expected performance ratios. For online bidding, the numerical evaluation is exact, so there are no error bars.

   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: [Yes]

   Justification: This information is provided at the last section of the technical appendix.

   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 points in the NeurIPS Code of Ethics do not apply to the nature of results of this paper. We thus conform to the code of ethics, in every respect.

   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: Our work studies foundational issues in the analysis of learning-augmented algorithms. There are no negative societal impacts.

    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: The paper poses no such risks.

    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: [Yes]

    Justification: We cite all open-sourced datasets according to the guidelines.

    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: [Yes]

    Justification: All code and data is provided in the supplementary material, with a README file.

    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: No study participants involved.

    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: LLMs were used to help with some technical computations, which were subsequently verified.

    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.