PRIVATE PROPHETS

Anonymous authors
Paper under double-blind review

ABSTRACT

Prophet inequalities are a basic problem in decision making: values drawn independently from fixed distributions arrive sequentially, and a mechanism must irrevocably select one while competing with the expected maximum. In the learning version of the problem, the distributions are unknown, but we are given independently drawn samples from each distribution from which a mechanism is learned. Our focus is differentially private learning, requiring the learned decision rule to be insensitive to any single training sample.

We study two natural notions of privacy. First, we consider public threshold mechanisms, which release a collection of thresholds and accept the first value exceeding its corresponding decision threshold. Second, we consider index-private mechanisms in which only the identity of the selected element is released. For both privacy notions, we obtain essentially tight sample-privacy-utility tradeoffs.

1 INTRODUCTION

Prophet inequalities are a classic problem in optimal stopping. In this setting, we are given a set of n probability distributions supported on the non-negative reals. An online policy is presented, one-by-one, with a sample from each of the distributions and must decide whether to select and stop, or discard and move on to the next one immediately upon arrival. The goal is to determine a decision rule for selecting a sample from the distributions that approximates the maximum realization, known as the proverbial *prophet*, as closely as possible in expectation. A good decision rule must balance the two extreme ends of failure: (i) aim for a very high outcome and we might not select any sample at all or (ii) aim for a low outcome and the prophet's choice might drastically outperform our own. Classic algorithms in this line of research, often referred to as *prophet inequalities*, define a threshold τ and select the first sample exceeding τ. For an appropriate choice of τ, the expected value of the selected sample is competitive with the highest realized sample value.

Classic prophet inequalities require full information on the distributions, or at the very least, they require the mechanism to be able to evaluate the probability that a sample exceeds any given threshold τ. This is often an unrealistic assumption: In practice our knowledge of the underlying distributions is often inaccurate, incomplete, and noisy. Surprisingly, even in a very constrained setting, it is possible to obtain very strong results: Rubinstein et al. (2020) showed that if given a single sample from every probability distribution, there exists a mechanism that guarantees the same, optimal competitive performance with respect to the prophet's choice as with full knowledge of the distributions.

Nevertheless, prophet inequalities obtained from limited samples have many significant drawbacks. The learned decision rules are prone to noise and corruptions; even if the samples themselves are given truthfully, malicious agents corrupting the samples might drastically affect the performance of the mechanism. Indeed, the decision rule might be completely determined by very few samples, which makes it inherently non-robust. Thus, a key challenge in this line of work is to formulate robust prophet inequalities that retain the good performance of the idealized settings.

A particularly important field in designing robust estimators is differential privacy: Given a set of L samples, the goal is to formulate a mechanism for which the outputted decision rule is not strongly influenced by the value of any particular sample. For example, in the case where we output a threshold based on the samples $S = \{x_1, \ldots, x_L\}$, a differentially private mechanism satisfies the requirement that $\mathbb{P}[M(S) = \tau] \leq e^{\varepsilon}\mathbb{P}[M(S') = \tau]$ for any candidate output τ, where S' replaces a single sample of S with an arbitrary different value, see the preliminaries for a formal definition.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Differentially private mechanisms are inherently robust: manipulating a decision rule of a private mechanism requires manipulating and corrupting significantly more samples than the corresponding decision rule of a non-private mechanism. In addition, there are other benefits to designing private mechanisms. Indeed, privacy itself is an important concern: the samples might only be procured and admitted for data analysis if this data is guaranteed to not be leaked anywhere else. Thus, the main question in this paper is:

> When can a differentially private mechanism be competitive with an optimal prophet inequality?

1.1 Our Contribution

In this paper, we give differentially private mechanisms for prophet inequalities in two natural settings.

First, we study the classic setting where our goal is to select a candidate threshold $\tau \in [0, B]$, where B is an upper bound on the maximum value that any distribution can assume. As mentioned above, in the non-private setting, a single sample from each distribution is enough to ensure constant factor competitiveness. Surprisingly, even when requiring that threshold release guarantees differential privacy, we retain strong utility. Let $k = \log(B/B_0)$ and denote by OPT_q the q-th highest realized value.

Theorem 1.1 (Informal, see Theorem 2.1)**.** *Let α be a sufficiently small absolute constant. Using only one sample from each distribution supported on $[0, B]$, there exists an ε-differentially private threshold mechanism that achieves expected utility of at least $\alpha \cdot \mathsf{OPT}_q - B_0$, for*

$$q = O\left(\frac{\log k}{\varepsilon}\right).$$

Moreover, the bounds are sharp; no single sample mechanism releasing a threshold can do better either in terms of q or B_0.

To be competitive with an optimal prophet inequality, we require multiple samples. Our next result provides a full characterization for a differentially privately released threshold with the all-knowing prophet's choice, given any number of samples:

Theorem 1.2 (Informal, see Theorem A.2)**.** *Let α be a sufficiently small absolute constant. Using L samples from each distribution supported on $[0, B]$, there exists an ε-differentially private threshold mechanism that achieves an expected utility of at least $\alpha \cdot \mathsf{OPT} - \gamma$, where $\gamma < B \exp\left(-\exp\left(\frac{\varepsilon L}{10^5}\right)\right)$.*

The additive bound is now substantially larger. Unfortunately, this is also necessary; any mechanism competitive with an optimal prophet must incur an additive error with a similar order of magnitude.

For both of these results, we incurred an additive error. While the additive error is necessary, as is typical for private algorithms, non-private prophet inequalities typically aim for a purely multiplicative error. Thus, to obtain a purely multiplicative approximation, we must aim for a weaker privacy notion.

Instead of releasing the decision thresholds, we only release the index and realized value of the distribution selected when running a private prophet inequality. Any thresholds the algorithm uses are not publicly released. We call this guarantee *index differential privacy*. In this setting, we show that there exists a private mechanism that is competitive with the prophet's choice if given sufficiently many samples from the distribution:

Theorem 1.3 (Informal, see Theorem 3.1)**.** *Let α be a sufficiently small absolute constant. Using $L \geq 512$ samples from each distribution supported on $[0, \infty]$, there is an online mechanism that is $512/L$-index differentially private and achieves expected utility of at least $\alpha \cdot \mathsf{OPT}$.*

Any mechanism achieving a purely multiplicative guarantee with ε-index differential privacy must use $\Theta(\varepsilon^{-1})$ many samples, thus the bound is tight up to constant factors.

In summary, we give a complete characterization of the sample complexity of differentially private prophet inequalities, both when the mechanisms output thresholds and when they only reveal the identity and value of the *winner*.

1.2 Related Work

Differentially Private Mechanism Design There is an important connection between differentially private mechanisms and approximately truthful mechanisms, as both require stability against changes in a single agent's private type. McSherry & Talwar (2007) formalizes this relation by showing that an ϵ-differentially private mechanism is 2ϵ-approximately dominant strategy truthful. Nissim et al. (2012) showed that differential privacy can enable exact dominant strategy truthfulness in certain settings.

Rather than just facilitating other nice mechanism design properties, differential privacy is often also the main topic of study. Huang & Kannan (2012) and Chen et al. (2016) provide private generalizations of the VCG mechanism, offering a smooth trade-off between the privacy budget and utility while guaranteeing exact dominant strategy truthfulness. The Sensitive Surveyor's Problem Roth (2012) explores conducting compensating agents for the cost of privacy loss. Finally, Gregorio et al. (2026) studied learning bilateral trade while guaranteeing differential privacy constraints, achieving nearly optimal sample complexity bounds.

Prophet Inequalities Since the seminal work of Krengel & Sucheston (1978), prophet inequalities have become a central model in online stochastic optimization, with manifold extensions to the original problem of selecting the highest value in the online sequence. Most relevant to us is a strong line of research that considers learning prophet inequalities from *limited data*: here, the strong but often unrealistic assumption of having full distributional information is replaced by access to a certain number of *samples* from each distribution. Even under such assumptions and often with only one sample, constant competitive ratios similar or equal to those from the classic, full-information setting can be recovered, see e.g. Azar et al. (2014); Rubinstein et al. (2020) for single-sample methods or Feldman et al. (2015); Dütting et al. (2024) for estimating prices via samples. The only work directly related to considering differential privacy in prophet inequalities is, to the best of our knowledge, Azize et al. (2026). Here, the authors consider local differential privacy for the online participants, resulting in each online value being available to the algorithm only in *perturbed* form, i.e. the output from an ε-LDP mechanism. They establish tight competitive ratios w.r.t. both a prophet who will observe the whole, original online sequence and an LDP prophet who will observe instead the sequence of perturbed inputs. This model differs from ours substantially: they assume full information on the underlying distributions, and assume that participants are unwilling to reveal their private data to the mechanism. We focus instead on the problem of *learning an online mechanism from limited, collected data*, without revealing any private information from the dataset.

1.3 Preliminaries

Prophet Inequalities There are n independent nonnegative distributions $\mathcal{D} := (\mathcal{D}_i)_{i\in[n]}$[1]. An online policy then is presented with online elements x_i, for each $i \in [n]$, drawn from $\mathcal{D}_i$ one at a time. They must, immediately and irrevocably on an element's arrival, decide to accept x_i, or reject it and see x_{i+1}. We use

$$\mathsf{OPT} := \mathbb{E}_{\mathbf{x}\sim\mathcal{D}}\left[\max_{i\in[n]} x_i\right],$$

to denote the expected value of the optimal choice in hindsight (i.e. the one made by an omniscient *prophet*). Similarly, we let $\mathsf{OPT}_q := \mathbb{E}_{\mathbf{x}\sim\mathcal{D}}[x_{(q)}]$ denote the expected value of the q-th order statistic of the online sequence $\mathbf{x}$. Furthermore, we use F_M for the CDF of the maximum value of an online sequence and $M(\mathbf{x}) := \max_{i\in[n]} x_i$.

We shall assume that all value distributions have bounded support and are all bounded by a constant $B > 0$. More precisely, we assume for all $i \in [n]$ that $\operatorname{supp}\mathcal{D}_i \subseteq [0, B]$.

The online policy maker only has access to $L \in \mathbb{Z}^+$ samples from each distribution and implements some online algorithm to maximize the value of the selected element from the online sequence. Let $\mathsf{ALG}(\mathbf{s}, \mathbf{x})$, denote the expected value of the selected element from the online sequence $\mathbf{x}$ given the samples $\mathbf{s} = (\mathbf{s}_\ell)_{\ell\in[L]}$, where $\mathbf{s}_\ell \sim \mathcal{D}$. A (possibly randomized) online algorithm is called *c-competitive* against OPT if, for all distributions $\mathcal{D}$ whose supports lie in $[0, B]$, the following

[1]Here, $[n] := \{1, 2, .., n\}$, the set of the first n positive integers.

inequality holds:

$$\mathbb{E}_{\mathbf{s}\sim\mathcal{D}^L,\mathbf{x}\sim\mathcal{D}}[\mathsf{ALG}(\mathbf{s},\mathbf{x})] \geq c \cdot \mathsf{OPT}.^2$$

Similarly, one can define competitiveness against OPT_q.

An important class of online mechanisms are threshold mechanisms. The policy maker determines from their samples a threshold $T(\mathbf{s})$ (possibly at random) and accepts the first value from the online sequence $\mathbf{x}$ satisfying $x_i \geq T(\mathbf{s})$. In that case, we let $T(\cdot)$ denote the threshold mechanism. Sometimes, with a slight abuse of notation, we write $\mathsf{ALG}(T(\mathbf{s}),\mathbf{x})$ to denote the value selected from the online sequence $\mathbf{x}$ given the threshold $T(\mathbf{s})$ to underscore that we analyze a threshold mechanism.

Differential Privacy

Definition 1.4 ((Neighboring) Dataset)**.** *The dataset for our setting is the set of samples $\mathbf{s} \in \mathbb{R}^{L\times n}$ initially available to the online policy. Two datasets $\mathbf{s}$ and $\mathbf{s}'$ are said to be* neighboring *if at most one entry is different. More precisely, $s_{\ell,i} = s'_{\ell,i}$ for all $\ell, i \in [L]\times[n]$ except for at most one tuple (ℓ^*, i^*).*

Following the traditional definition and assuming that all parts of the policy are externally observable, we use the following definition of differential privacy.

Definition 1.5 (Differential Privacy)**.** *Any threshold mechanism is (ε,δ)-differentially private if for any neighboring datasets $\mathbf{s},\mathbf{s}' \in \mathbb{R}^{L\times n}_{\geq 0}$ and every measurable event E, we have $\mathbb{P}(T(\mathbf{s}) \in E) \leq e^{\varepsilon}\mathbb{P}(T(\mathbf{s}') \in E) + \delta$.*

We will primarily consider the case $\delta = 0$, commonly referred to as pure *differential privacy. In this case, we simply say that T is ε-differentially private.*

Alternatively, it is reasonable to assume that not the threshold itself, but only the winner (chosen online element) is observable. This gives rise to the following notion of index-privacy, which naturally extends beyond threshold mechanisms.

Definition 1.6 (Output Index Privacy)**.** *Let $\mathcal{M}(\mathbf{s},\mathbf{x}) \in [n]\cup\{\perp\}$ denote the online winner, given dataset $\mathbf{s}$ and the online sequence $\mathbf{x}$, where $\perp$ means no element was chosen. A mechanism $\mathcal{M}$ is then (ε,δ)-index private if and only if for any pair of neighboring datasets $\mathbf{s},\mathbf{s}'$ and any $E \subseteq [n]\cup\{\perp\}$ the following inequality holds:*

$$\mathbb{P}(\mathcal{M}(\mathbf{s},\mathbf{x}) \in E) \leq e^{\varepsilon}\mathbb{P}(\mathcal{M}(\mathbf{s}',\mathbf{x}) \in E) + \delta.$$

Note that for threshold mechanisms specifically, Definition 1.5 implies Definition 1.6.

Laplace Mechanism A standard way to obtain pure differential privacy is the Laplace mechanism. Let $f : \mathbb{R}^{L\times n} \to \mathbb{R}$ be a deterministic function of the dataset, and define its *sensitivity* by

$$\Delta_f := \sup_{\mathbf{s}\sim\mathbf{s}'} |f(\mathbf{s}) - f(\mathbf{s}')|,$$

where $\mathbf{s} \sim \mathbf{s}'$ denotes that $\mathbf{s}$ and $\mathbf{s}'$ are neighboring datasets. The Laplace mechanism, see Dwork et al. (2016) outputs

$$\mathcal{M}(\mathbf{s}) := f(\mathbf{s}) + Z, \qquad Z \sim \mathrm{Lap}\left(\frac{\Delta_f}{\varepsilon}\right),$$

where $\mathrm{Lap}(b)$ denotes the Laplace distribution with density $\frac{1}{2b}e^{-|z|/b}$. The resulting mechanism is ε-differentially private. Thus, the amount of noise required is proportional to how much the statistic f can change when a single sample in the dataset is changed.

1.4 Our Techniques

Threshold Privacy A commonly used technique when designing private learning algorithms is to discretize the space of available candidate solutions and run an exponential mechanism, that is, we sample a candidate threshold proportionate to its value score, while ensuring that the sampling distribution itself guarantees privacy. We use this approach as well. In the single-sample setting,

[2]The notation $\mathbf{s} \sim \mathcal{D}^L$ denotes L independent samples from $\mathcal{D}$.

the main challenge is that privacy prevents us from locating the desirable threshold too accurately from the samples. Our upper bound shows that it is enough to identify the threshold up to rank error $O((\log k)/\varepsilon)$; a symmetrization argument then converts this approximate sample-rank guarantee into a constant-factor approximation to OPT_q for $q = O((\log k)/\varepsilon)$. For the matching lower bound, we construct single-sample instances at many well-separated value scales and show that good utility forces the mechanism to place constant probability on a different threshold interval at each scale. Differential privacy transfers these probabilities to a common dataset, and the disjointness of the intervals yields $q = \Omega((\log k)/\varepsilon)$.

Index Privacy Our mechanism uses Laplace noise to perturb a rank cutoff centered at $L/4$, counting from the largest of all $L \cdot n$ sample elements. It selects the first arriving value whose number of larger samples is at most the noisy cutoff. Privacy follows from a standard noise-shifting argument. The main challenge is the reward analysis: premature selections must remain unlikely across many arrivals, while sufficiently high values must be selected with constant probability. We analyze counts through the aggregate tail mass $\lambda(t) = \sum_i \mathbb{P}(x_i > t)$. For low values, acceptance probabilities decay exponentially in their aggregate tail mass. We choose lower and upper analysis levels with aggregate tail masses $1/2$ and $1/8$. Selections below the lower level are rare, while values above the upper level satisfy the acceptance inequality with high probability. Since the probability of any preceding value exceeding the lower level is at most $1/2$, the mechanism reaches each arrival with constant probability. Subtracting the small probability that a high value violates the acceptance inequality then gives a constant probability of selecting it. These estimates capture constant fractions of a base reward and the expected excess contributed by high values, which together bound OPT up to a constant factor.

2 Threshold Privacy

We study threshold algorithms when the threshold must be differentially private with respect to the samples used to construct it. We focus on the single-sample setting, where the algorithm receives one sample from each distribution, uses these samples to choose a single threshold, and then accepts the first online value that exceeds this threshold. Our benchmark is OPT_q, the expected q-th largest online value. Without privacy, one can obtain a constant approximation to the maximum, corresponding to $q = 1$. Privacy changes the picture: changing one sample should not significantly change the distribution of the chosen threshold, so the mechanism cannot simply follow the largest observed sample.

We prove the single-sample upper bound in this section and give a sketch of the matching lower bound at the end; the full lower-bound proof appears in Theorem A.1 in the appendix. We also study the multiple-sample setting in the appendix, where sufficiently many samples allow us to compete directly with OPT; see Theorem A.2.

Assume that every distribution is supported on $[0, B]$, where B is known, and fix an additive-error parameter $B_0 \in (0, B)$. Since we allow additive loss B_0, values below this scale do not need to be distinguished more finely. Let $k := \left\lceil \log_2 \frac{B}{B_0} \right\rceil$ denote the number of relevant dyadic scales. Throughout this section we assume $0 < \varepsilon \leq 1$, sufficiently many value scales, and that n is large enough to accommodate the stated rank; these assumptions let us suppress additive constants in q. Our main result is a matching upper and lower bound: for $q = O\left(\frac{\log k}{\varepsilon}\right)$, there is a purely ε-differentially private threshold satisfying $\mathbb{E}[\mathsf{ALG}] \geq \gamma(\mathsf{OPT}_q - B_0), \qquad \gamma := \frac{1-e^{-1}}{2}$, while every private threshold satisfying the same guarantee must have $q = \Omega\left(\frac{\log k}{\varepsilon}\right)$. Thus the dependence on both the number of relevant value scales and the privacy parameter is tight up to constant factors.

2.1 Upper Bound

We obtain the upper bound by privately choosing one of the k dyadic levels and comparing the sample and online values directly. The private choice only needs to identify the threshold to within rank error $O((\log k)/\varepsilon)$, and the analysis shows that this is enough with constant probability. As a result, the mechanism competes with an order statistic of rank $O((\log k)/\varepsilon)$.

Algorithm 1 PRIVATETHRESHOLD

Require: Samples s, dyadic grid $b_0, \dots, b_k$, target rank r

1: $C_j \leftarrow |\{i : s_i > b_j\}|$ for $j = 0, \dots, k$
2: $C_{-1} \leftarrow n$, $C_k \leftarrow 0$
3: Sample $J \in \{-1, \dots, k-1\}$ with

$$\Pr(J = j) \propto \exp\left(-\frac{\varepsilon}{2} \operatorname{dist}(r, [C_{j+1}, C_j])\right)$$

4: **return** 0 if $J = -1$, and b_J otherwise

The mechanism. Let $b_j := \min\{2^j B_0, B\}$, $j = 0, \dots, k$, be the dyadic grid. For a sample vector s, define

$$C_j(s) := |\{i : s_i > b_j\}|, \qquad C_{-1}(s) := n, \qquad C_k(s) := 0.$$

Thus $C_j(s)$ is the number of samples above level b_j. Fix a target sample rank r. For each $j = -1, \dots, k-1$, define

$$u_j(s) := -\operatorname{dist}(r, [C_{j+1}(s), C_j(s)]),$$

where $\operatorname{dist}(r, [a, b])$ is the distance from r to $[a, b]$. We then sample J using the exponential mechanism and return the corresponding dyadic threshold. In words, the mechanism privately chooses the dyadic level whose position among the samples is closest to the desired rank.

Theorem 2.1 (Upper bound)**.** *There is an ε-differentially private threshold mechanism such that, for $q = O((\log k)/\varepsilon)$,*

$$\mathbb{E}[\mathsf{ALG}] \geq \gamma(\mathsf{OPT}_q - B_0), \qquad \gamma = \frac{1 - e^{-1}}{2}.$$

Proof sketch. We give a sketch of the proof here and the complete proof is deferred to Appendix A. Changing one sample changes every C_j by at most one, and hence every score u_j by at most one. Algorithm 1 is therefore ε-differentially private.

Let

$$\rho := \left\lceil \frac{2}{\varepsilon} \log(16(k+1)) \right\rceil = O\left(\frac{\log k}{\varepsilon}\right).$$

Here ρ is the rank error we allow for the private choice. Fix a constant $m \geq 8$, and set $r := m + \rho$ and $h := m + 2\rho + 1$. The target r is chosen so that even after an error of ρ, the selected threshold remains below the m-th sample; h is the corresponding lower sample rank on the other side.

The exponential mechanism guarantees, with probability at least $15/16$, that

$$\operatorname{dist}(r, [C_{J+1}, C_J]) \leq \rho.$$

On this event, if $s^{(h)} > B_0$, then $C_J \geq m$ and $C_{J+1} \leq h - 1$, which implies

$$\frac{s^{(h)}}{2} \leq T < s^{(m)}. \tag{1}$$

Indeed, $C_J \geq m$ gives $T < s^{(m)}$, while $C_{J+1} \leq h - 1$ gives $s^{(h)} \leq b_{J+1} \leq 2b_J = 2T$.

The symmetrization step. The key step is a simple symmetrization that lets us compare the sample and online values in one shot. For each i, draw two independent values from D_i, and only afterwards choose uniformly which becomes the sample s_i and which becomes the online value x_i. Condition on the resulting $2n$ numerical values, breaking ties by a fixed rule independent of these random assignments, and let v be their $4h$-th largest value. The point of choosing $4h$ is that a constant fraction of these large values will fall on the sample side with high probability, while the same random assignment also leaves a large value on the online side.

With probability at least $7/8$, at least h of the largest $4h$ combined values become samples, so $s^{(h)} \geq v$. To see this, complete pairs among these values contribute one sample automatically, while the remaining values are independently assigned to the sample with probability $1/2$; a Chernoff bound gives the claim. Similarly, with probability at least $15/16$, at least one of the largest m combined values becomes online: if these m values contain a complete pair this is automatic, and otherwise the probability that all m become samples is $2^{-m} \leq 1/16$. Hence on this event $\max_i x_i \geq s^{(m)}$.

Together with the private rank event, a union bound shows that all three properties hold simultaneously with probability at least $1 - \frac{1}{8} - \frac{1}{16} - \frac{1}{16} = \frac{3}{4}$.

If $v > B_0$, they imply $\frac{v}{2} \leq \frac{s^{(h)}}{2} \leq T < s^{(m)} \leq \max_i x_i$.

Thus the online algorithm crosses the threshold and earns at least $v/2$. Conditioning on the fixed $2n$ values therefore gives

$$\mathbb{E}[\mathsf{ALG}] \geq \frac{3}{8}(v - B_0)_+.$$

Finally, the $4h$-th largest online value is always at most the $4h$-th largest value among all $2n$ combined values, namely v. Averaging gives

$$\mathbb{E}[\mathsf{ALG}] \geq \frac{3}{8}\,\mathbb{E}[(x^{(4h)} - B_0)_+] \geq \frac{3}{8}(\mathsf{OPT}_{4h} - B_0).$$

We therefore set $q := 4h$, matching the order statistic controlled by the coupling. Since $h = m + 2\rho + 1$, we have $q = O((\log k)/\varepsilon)$. If $\mathsf{OPT}_q \leq B_0$, the desired guarantee follows from $\mathsf{ALG} \geq 0$; otherwise $3/8 > \gamma = (1 - e^{-1})/2$, and hence

$$\mathbb{E}[\mathsf{ALG}] \geq \gamma(\mathsf{OPT}_q - B_0).$$

□

2.2 Lower Bound

We complement the upper bound with a matching lower bound; the full proof appears in Theorem A.1 in the appendix. The idea is to construct instances at many different value scales, each of which requires the threshold to lie in a different range. Differential privacy then forces noticeable probability on all of these ranges even on one common dataset, which is impossible unless q is sufficiently large.

Let $K := \lfloor \log_8(B/B_0) \rfloor = \Theta(k)$. For each $j = 1, \ldots, K$, set $v_j := 8^j B_0$ and $b_j := v_j/8$, and consider the deterministic instance

$$\left(b_j, \underbrace{v_j, \ldots, v_j}_{q \text{ copies}}, 0, \ldots, 0 \right).$$

Here $\mathsf{OPT}_q = v_j$. To obtain a constant fraction of this benchmark, the mechanism must skip b_j while still accepting one of the subsequent v_j's, so with constant probability its threshold must lie in $I_j := (b_j, v_j]$.

The corresponding sample vector differs from the all-zero vector in only $q + 1$ entries. Differential privacy therefore forces probability $\Omega(e^{-\varepsilon(q+1)})$ on I_j even on the all-zero dataset. Since the $K = \Theta(k)$ intervals I_j are pairwise disjoint, their probabilities on this common dataset must sum to at most one. Hence $1 \geq \Omega!\left(Ke^{-\varepsilon(q+1)}\right)$, so $\varepsilon(q+1) = \Omega(\log K)$. Since $K = \Theta(k)$, $q = \Omega\left(\frac{\log k}{\varepsilon}\right)$, matching the upper bound up to constant factors.

3 Index Privacy

The weaker requirement of index privacy allows us to obtain a constant-factor approximation to OPT under pure differential privacy, without the additive error terms present in the threshold-private guarantees of the preceding section. For continuous distributions, we achieve this guarantee using L samples with $\varepsilon = O(1/L)$. Conversely, any fixed positive competitive ratio guaranteed uniformly over all sequence lengths requires $\varepsilon = \Omega(1/L)$.

3.1 UPPER BOUND

For each online value, Algorithm 2 counts how many samples exceed it. It accepts the first value whose count is at most a shared noisy cutoff centered at $\lfloor L/4 \rfloor$ plus fresh Laplace noise for that arrival. Its output $\mathcal{M}(\mathbf{s}, \mathbf{x}) \in [n] \cup \{\perp\}$ is the selected index, with $\perp$ denoting no selection and no reward.

Algorithm 2 INDEX-PRIVATE RANK SELECTION

Require: Sample matrix $\mathbf{s} = (s_{\ell,j})_{\ell \in [L], j \in [n]}$, with $L \geq 512$.
Require: Online sequence $\mathbf{x} = (x_i)_{i \in [n]}$, revealed one value at a time.
Ensure: Selected index $\mathcal{M}(\mathbf{s}, \mathbf{x}) \in [n] \cup \{\perp\}$.
All noises are mutually independent and independent of $\mathbf{s}, \mathbf{x}$.
Set $\varepsilon \leftarrow 512/L$ and $k \leftarrow \lfloor L/4 \rfloor$.
Draw the shared noise $\eta_0 \sim \mathrm{Lap}(2/\varepsilon)$.
for $i = 1, \ldots, n$ **do**
  Draw $\eta_i \sim \mathrm{Lap}(4/\varepsilon)$.
  Observe the next online value x_i.
  Compute
$$c(\mathbf{s}, x_i) \leftarrow \#\{(\ell, j) \in [L] \times [n] : s_{\ell,j} > x_i\}.$$
  if $c(\mathbf{s}, x_i) \leq k + \eta_i + \eta_0$ **then**
    return i. ▷ Select x_i and stop.
  end if
end for
return $\perp$. ▷ No element selected; reward zero.

Theorem 3.1 (Index privacy from pooled samples)**.** *For every $n \geq 1$ and $L \geq 512$, there is an online mechanism that is $(512/L, 0)$-index private and, for every tuple $\mathcal{D}$ of independent distributions on $[0, \infty)$ with continuous cumulative distribution functions, is $(1/10)$-competitive against* OPT*:*

$$\mathbb{E}_{\mathbf{s} \sim \mathcal{D}^L, \mathbf{x} \sim \mathcal{D}}[\mathsf{ALG}(\mathbf{s}, \mathbf{x})] \geq \frac{1}{10}\mathsf{OPT}.$$

In fact, its privacy inequality holds for every fixed online sequence $\mathbf{x} \in \mathbb{R}^n_{\geq 0}$ and all events in $[n] \cup \{\perp\}$, including the outcome of selecting no element.

Proof overview. We defer the full proof to Section B.1. Privacy follows from a standard translation of the Laplace noises. For the reward analysis, define $\lambda(t) = \sum_i \mathbb{P}(x_i > t)$ and choose analysis levels $a < b$ with $\lambda(a) = 1/2$ and $\lambda(b) = 1/8$. Write R for the reward and $H = \sum_i \mathbb{E}[(x_i - a)\mathbf{1}_{\{x_i \geq b\}}]$ for the total expected excess above a contributed by high values. Values in $[a, b)$ contribute at most $3(b-a)/8$ in expected excess, while $H \geq (b-a)/8$. Splitting the prophet's reward and the mechanism's reward into a base amount and excess therefore gives

$$\mathsf{OPT} \leq a + 4H, \qquad \mathbb{E}[R] \geq a\mathbb{P}(R \geq a) + \mathbb{E}[(R-a)\mathbf{1}_{\{R \geq b\}}]. \tag{2}$$

Controlling premature selections. For a fixed value $z < a$, the count has mean $L\lambda(z) \geq L/2$, well above the cutoff $\lfloor L/4 \rfloor$. Concentration and Laplace tails bound its probability of satisfying the acceptance inequality by $2e^{-16\lambda(z)}$. We may sum these bounds over all arrivals by integrating over the aggregate tail mass, giving a total probability of selecting below a at most $2\int_{1/2}^{n} e^{-16u}\,du < 1/1000$. Conversely, each fixed value at least b has expected count at most $L/8$, and its acceptance inequality fails with probability at most $1/12$.

Recovering the base and excess. Independence gives $\mathbb{P}(M(\mathbf{x}) \geq b) \geq 1 - e^{-1/8}$. Conditional on such an online sequence, the maximum satisfies its acceptance inequality with probability at least $11/12$; the mechanism then selects it or has already selected an earlier value. Subtracting the probability of a selection below a gives $\mathbb{P}(R \geq a) \geq (11/12)(1 - e^{-1/8}) - 1/1000 > 1/10$. For the excess, stopping before an arrival requires either a selection below a or a preceding value at least a. Thus each arrival is reached with probability at least $1/2 - 1/1000$, independently of its value. Subtracting the acceptance-failure probability shows that each fixed value at least b is selected with

probability at least $1/2 - 1/1000 - 1/12 > 2/5$. Integrating these selection probabilities against the excess $z - a$ for $z \geq b$ yields $\mathbb{E}[(R-a)\mathbf{1}_{\{R \geq b\}}] \geq (2/5)H$, so equation 2 gives

$$\mathbb{E}[R] \geq \frac{a}{10} + \frac{2}{5}H = \frac{1}{10}(a + 4H) \geq \frac{1}{10}\mathsf{OPT}.$$

3.2 Lower Bound

In this section, we present a general lower bound on the privacy guarantee ε of pure differentially private mechanisms (i.e. $\delta = 0$ in Definition 1.6).

Theorem 3.2. *Fix $n, L \geq 1$ and $0 < c < 1$, and suppose $n > 1/c$. Let $\mathcal{M}$ be a law-independent randomized online threshold mechanism that is universally c-competitive over all sequences of n independent continuous value distributions. If $\mathcal{M}$ is $(\varepsilon, 0)$-index private, then*

$$\varepsilon \geq \frac{a_{c,n}}{L}, \tag{3}$$

where $a_{c,n} > 0$ is such that $a_{c,n} \to \log(1+c)/(1-c)$ as $n \to \infty$.

The proof of Theorem 3.2 is based on the following lemma.

Lemma 3.3. *Under the same assumptions as Theorem 3.2, for any $m \in [n]$, we have*

$$\sum_{r=1}^{n} e^{-\varepsilon L|r-m|} \leq \frac{1}{c}. \tag{4}$$

The idea of the proof of the lemma is to build a sequence of instances for which we know the winner beforehand, and any algorithm satisfying the assumptions of Theorem 3.2 selects it with probability bounded below by a quantity approaching c. We construct n-dimensional product distributions whose components are either supported near 0 or on small intervals starting at exponentially increasing points, with the intervals determined by a scale parameter R and pairwise disjoint. The differential privacy guarantees along with the lower bounds on the probabilities of selecting the winner imply the final result.

We explain how Lemma 3.3 implies the theorem and defer the proof of the lemma to Section B.2.

Proof of Theorem 3.2. For $m \in [n]$ and $a \geq 0$, define $G_{n,m}(a) := \sum_{r=1}^{n} e^{-a|r-m|}$.

Given this definition, equation 4 becomes $\frac{1}{c} \geq G_{n,m}(\varepsilon L)$. Let $G_n := G_{n,m^\star}$, where $m^\star = \lfloor (n+1)/2 \rfloor$. Observe that $G_n(a) = \max_{m \in [n]} G_{n,m}(a)$ if $a > 0$. For $1 < \frac{1}{c} < n$, $G_n(a)$ is continuous and strictly decreasing from n to 1, so there exists a unique positive solution to the equation $G_n(a) = \frac{1}{c}$, call it $a_{c,n}$. Consequently Lemma 3.3 implies that $\varepsilon \geq \frac{a_{c,n}}{L}$.

Writing $n = 2d+1$ or $n = 2d$, respectively,

$$G_n(a) = \begin{cases} 1 + 2\displaystyle\sum_{j=1}^{d} e^{-aj}, & n = 2d+1, \\ 1 + 2\displaystyle\sum_{j=1}^{d-1} e^{-aj} + e^{-ad}, & n = 2d. \end{cases} \tag{5}$$

For each fixed $a > 0$, both expressions in equation 5 converge to

$$1 + 2\sum_{j=1}^{\infty} e^{-aj} = 1 + 2\left(\frac{1}{1-e^{-a}} - 1\right) = \frac{1+e^{-a}}{1-e^{-a}}.$$

By continuity and strict monotonicity, as $n \to \infty$, $a_{n,c}$ converges to the unique solution of

$$\frac{1+e^{-a}}{1-e^{-a}} = \frac{1}{c},$$

which is $a = \log((c+1)/(1-c))$ (given $0 < c < 1$). □

Acknowledgment of AI assistance. We used GPT 5.6 Sol to explore improvements to our initial, non-optimal bounds and to refine the proofs and exposition. These interactions helped us identify arguments that closed the gap between our upper and lower bounds. The authors take responsibility for all mathematical claims and the final manuscript.

REFERENCES

Pablo D Azar, Robert Kleinberg, and S Matthew Weinberg. Prophet inequalities with limited information. In *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*, pp. 1358–1377. SIAM, 2014.

Achraf Azize, Mathieu Molina, Hugo Richard, and Vianney Perchet. Prophet inequalities under local differential privacy. In *The 27th conference on Economics and Computation (EC 26)*, 2026.

Yiling Chen, Stephen Chong, Ian A. Kash, Tal Moran, and Salil P. Vadhan. Truthful mechanisms for agents that value privacy. *ACM Trans. Economics and Comput.*, 4(3):13:1–13:30, 2016.

Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam D. Smith. Calibrating noise to sensitivity in private data analysis. *J. Priv. Confidentiality*, 7(3):17–51, 2016.

Paul Dütting, Thomas Kesselheim, Brendan Lucier, Rebecca Reiffenhäuser, and Sahil Singla. Online combinatorial allocations and auctions with few samples. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 1231–1250, 2024. doi: 10.1109/FOCS61266.2024.00081.

Michal Feldman, Nick Gravin, and Brendan Lucier. Combinatorial auctions via posted prices. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '15, pp. 123–135, USA, 2015. Society for Industrial and Applied Mathematics.

Simone Di Gregorio, Federico Fusco, Stefano Leonardi, and Chris Schwiegelshohn. Private learning in bilateral trade. *CoRR*, abs/2606.02050, 2026. doi: 10.48550/ARXIV.2606.02050. URL https://doi.org/10.48550/arXiv.2606.02050.

Zhiyi Huang and Sampath Kannan. The exponential mechanism for social welfare: Private, truthful, and nearly optimal. In *FOCS*, pp. 140–149. IEEE Computer Society, 2012.

Ulrich Krengel and Louis Sucheston. On semiamarts, amarts, and processes with finite value. In *Probability on Banach spaces*, volume 4, pp. 197–266. Dekker New York, 1978.

Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *FOCS*, pp. 94–103. IEEE Computer Society, 2007.

Kobbi Nissim, Rann Smorodinsky, and Moshe Tennenholtz. Approximately optimal mechanism design via differential privacy. In *ITCS*, pp. 203–213. ACM, 2012.

Aaron Roth. Buying private data at auction: the sensitive surveyor's problem. *SIGecom Exch.*, 11(1):1–8, 2012. doi: 10.1145/2325713.2325714. URL https://doi.org/10.1145/2325713.2325714.

Aviad Rubinstein, Jack Z. Wang, and S. Matthew Weinberg. Optimal single-choice prophet inequalities from samples. In Thomas Vidick (ed.), *11th Innovations in Theoretical Computer Science Conference, ITCS 2020, January 12-14, 2020, Seattle, Washington, USA*, volume 151 of *LIPIcs*, pp. 60:1–60:10. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. doi: 10.4230/LIPICS.ITCS.2020.60. URL https://doi.org/10.4230/LIPIcs.ITCS.2020.60.

A THRESHOLD PRIVACY

Proof of Theorem 2.1. Changing one sample changes every C_j by at most one, and hence every score u_j by at most one. Algorithm 1 is therefore ε-differentially private.

Let

$$\rho := \left\lceil \frac{2}{\varepsilon} \log(16(k+1)) \right\rceil = O\left(\frac{\log k}{\varepsilon}\right).$$

Fix a constant $m \geq 8$, and set $r := m + \rho$ and $h := m + 2\rho + 1$. Since $C_{-1} = n$, $C_k = 0$, and $(C_j)_{j=-1}^{k}$ is nonincreasing, some interval $[C_{j+1}, C_j]$ contains r, so the maximum score is zero. Hence the exponential mechanism gives

$$\Pr(\mathrm{dist}(r, [C_{J+1}, C_J]) > \rho) \leq (k+1)e^{-\varepsilon\rho/2} \leq \frac{1}{16}.$$

On the complementary event,

$$C_J \geq r - \rho = m, \qquad C_{J+1} \leq r + \rho = h - 1.$$

If $s^{(h)} > B_0$, then $J \neq -1$, since otherwise $C_{J+1} = C_0 \geq h$. Therefore $T = b_J$, and

$$\frac{s^{(h)}}{2} \leq T < s^{(m)}. \tag{1}$$

Indeed, $C_J \geq m$ gives $T = b_J < s^{(m)}$, while $C_{J+1} \leq h - 1$ gives $s^{(h)} \leq b_{J+1} \leq 2b_J = 2T$.

We next use the standard two-sample symmetrization. For each i, draw two independent values from $\mathcal{D}_i$ and, independently across coordinates, choose uniformly which becomes the sample s_i and which becomes the online value x_i. Condition on the resulting $2n$ numerical values, breaking ties by a fixed rule independent of these assignments, and let v be their $4h$-th largest value.

With probability at least $7/8$, at least h of the largest $4h$ combined values become samples. Indeed, suppose c complete pairs occur among these $4h$ values. These contribute exactly c samples, while the remaining $4h-2c$ values belong to distinct pairs and independently become samples with probability $1/2$. Thus, if Y is the number of samples among the largest $4h$ values,

$$\mathbb{E}[Y] = 2h, \qquad \mathrm{Var}(Y) \leq h,$$

and Chebyshev's inequality gives

$$\Pr(Y < h) \leq \frac{1}{h} \leq \frac{1}{8}.$$

Hence $s^{(h)} \geq v$ with probability at least $7/8$.

Similarly, with probability at least $15/16$, at least one of the largest m combined values becomes online. If these values contain a complete pair this is automatic; otherwise they belong to distinct pairs, and the probability that all become samples is $2^{-m} \leq 1/16$. On this event,

$$\max_i x_i \geq s^{(m)},$$

since $s^{(m)}$ is at most the m-th largest combined value.

The private-rank event has probability at least $15/16$ conditional on every sample vector. Therefore, by a union bound, all three events above hold simultaneously with probability at least

$$1 - \frac{1}{8} - \frac{1}{16} - \frac{1}{16} = \frac{3}{4}.$$

If $v > B_0$, then on this event $s^{(h)} \geq v > B_0$, so equation (1) gives

$$\frac{v}{2} \leq \frac{s^{(h)}}{2} \leq T < s^{(m)} \leq \max_i x_i.$$

Thus some online value crosses the threshold, and the algorithm obtains reward at least $v/2$. Conditioning on the fixed $2n$ numerical values therefore gives

$$\mathbb{E}[\mathsf{ALG}] \geq \frac{3}{8}(v - B_0)_+.$$

Finally, the $4h$-th largest online value is pointwise at most the $4h$-th largest combined value v. Averaging over the combined values,

$$\mathbb{E}[\mathsf{ALG}] \geq \frac{3}{8}\mathbb{E}[(x^{(4h)} - B_0)_+] \geq \frac{3}{8}(\mathsf{OPT}_{4h} - B_0).$$

Set $q := 4h$. Since $h = m + 2\rho + 1$, we have

$$q = O\left(\frac{\log k}{\varepsilon}\right).$$

If $\mathsf{OPT}_q \le B_0$, the desired guarantee follows from $\mathsf{ALG} \ge 0$. Otherwise, since $3/8 > \gamma = (1 - e^{-1})/2$,

$$\mathbb{E}[\mathsf{ALG}] \ge \gamma(\mathsf{OPT}_q - B_0),$$

as claimed. □

A.1 Lower Bound

We now show that the $\log k/\varepsilon$ dependence is necessary. The idea is to construct many deterministic instances at well-separated value scales. On each instance, good utility forces the threshold to lie with constant probability in a different interval. Privacy then forces some probability on each of these intervals even on the same all-zero sample vector. Since the intervals are disjoint, this is possible only if q is sufficiently large.

Theorem A.1 (Lower bound)**.** *Fix a constant $\gamma > 1/7$. Let $1 \le q < n$. Suppose an ε-differentially private threshold mechanism satisfies*

$$\mathbb{E}[\mathsf{ALG}] \ge \gamma(\mathsf{OPT}_q - B_0)$$

for every product distribution supported on $[0, B]$. Then, for sufficiently large k,

$$q = \Omega\left(\frac{\log k}{\varepsilon}\right).$$

Proof. Let

$$K := \left\lfloor \log_8 \frac{B}{B_0} \right\rfloor = \Theta(k).$$

For each $j = 1, \dots, K$, define $v_j := 8^j B_0$, $b_j := v_j/8$, and consider the deterministic instance

$$(b_j, \underbrace{v_j, \dots, v_j}_{q \text{ copies}}, 0, \dots, 0).$$

Here $\mathsf{OPT}_q = v_j$. Let $I_j := (b_j, v_j]$. If the threshold lies in I_j, the algorithm skips b_j and obtains v_j; outside I_j, its reward is at most b_j. The utility guarantee therefore forces

$$\Pr(T(s^{(j)}) \in I_j) \ge \kappa$$

for some universal constant $\kappa > 0$.

Now compare $s^{(j)}$ with the all-zero sample vector $s^\star$. They differ in exactly $q + 1$ coordinates, so pure differential privacy gives

$$\Pr(T(s^\star) \in I_j) \ge \kappa e^{-\varepsilon(q+1)}.$$

The intervals $I_1, \dots, I_K$ are pairwise disjoint, so their probabilities on the same dataset $s^\star$ must sum to at most one. Hence

$$1 \ge K\kappa e^{-\varepsilon(q+1)},$$

and therefore $\varepsilon(q + 1) \ge \log(\kappa K)$. Since $K = \Theta(k)$, this yields

$$q = \Omega\left(\frac{\log k}{\varepsilon}\right).$$

Combining Theorems 2.1 and A.1, we obtain matching upper and lower bounds:

$$\boxed{q = \Theta\left(\frac{\log k}{\varepsilon}\right).}$$

Thus pure differential privacy requires the rank to grow logarithmically with the number of relevant value scales, and the upper-bound mechanism achieves exactly this dependence up to universal constant factors.

A.2 MULTIPLE SAMPLES

We now consider the setting in which the mechanism observes L independent sample vectors

$$\mathbf{s}_1, \dots, \mathbf{s}_L \sim \mathcal{D},$$

where $\mathbf{s}_\ell = (s_{\ell,1}, \dots, s_{\ell,n})$, before posting a single threshold. An independent online vector $\mathbf{x} \sim \mathcal{D}$ then arrives, and the algorithm accepts the first value that meets or exceeds the posted threshold. As before, all distributions are supported on $[0, B]$, and

$$\mathsf{OPT} = \mathbb{E}_{\mathbf{x}\sim\mathcal{D}}\left[\max_i x_i\right].$$

Unlike the single-sample setting, with sufficiently many samples we can compete directly with OPT, rather than with a lower order statistic.

Theorem A.2. *Let $\varepsilon > 0$. There exists an $(\varepsilon, 0)$-differentially private threshold mechanism such that*

$$\mathbb{E}_{\mathbf{s}\sim\mathcal{D}^L, \mathbf{x}\sim\mathcal{D}}[\mathsf{ALG}(T(\mathbf{s}), \mathbf{x})] \geq \frac{1}{3}\mathsf{OPT} - B\exp\left(-\exp\left(\frac{\min\{1,\varepsilon\}L}{10^5}\right)\right).$$

The main difficulty is to privately identify the right threshold. For each sample vector, $m_\ell = \max_i s_{\ell,i}$ is an independent sample from the same distribution as the online maximum $M = \max_i x_i$. The proof below shows that it is enough to identify where a typical value of M lies. Two issues arise. First, this typical value could be anywhere between very small values and B. We therefore take logarithms, so values of very different sizes can be treated on the same scale. Second, the distribution of M may have point masses, which makes its CDF jump. We therefore add an independent $U_\ell \sim \mathrm{Unif}[0,1]$, which smooths the distribution and makes its CDF continuous and 1-Lipschitz.

The mechanism is as follows.

Algorithm 3 PRIVATE MEDIAN-SCALE THRESHOLD

Require: Sample vectors $\mathbf{s}_1, \dots, \mathbf{s}_L$, support bound B, privacy parameter ε.

1: Set

$$\eta = B\exp\left(-\exp\left(\frac{\min\{1,\varepsilon\}L}{10^5}\right)\right), \qquad R = \log(1 + B/\eta) + 1.$$

2: **for** $\ell = 1, \dots, L$ **do**
3: $\quad m_\ell \leftarrow \max_i s_{\ell,i}$.
4: $\quad$ Draw $U_\ell \sim \mathrm{Unif}[0,1]$.
5: $\quad z_\ell \leftarrow \log(1 + m_\ell/\eta) + U_\ell$.
6: **end for**
7: Let

$$\widehat{G}(v) = \frac{1}{L}\sum_{\ell=1}^{L} \mathbf{1}\{z_\ell \leq v\}.$$

8: Draw $V \in [0, R]$ with density f satisfying

$$f(v) \propto \exp\left(-\frac{\varepsilon L}{2}\left|\widehat{G}(v) - \frac{1}{2}\right|\right).$$

9: Draw $U \sim \mathrm{Unif}[0,1]$.
10: **return** $T = \eta(e^{V-U} - 1)$.

Proof. We first verify privacy. Conditional on $U_1, \dots, U_L$, changing one entry of the sample matrix changes at most one transformed value z_ℓ. Hence, for every v, the count

$$L\widehat{G}(v) = \sum_{\ell=1}^{L} \mathbf{1}\{z_\ell \le v\}$$

changes by at most one. The score

$$-\left| L\widehat{G}(v) - \frac{L}{2} \right|$$

therefore has sensitivity one, so sampling V with density proportional to

$$\exp\left(-\frac{\varepsilon}{2} \left| L\widehat{G}(v) - \frac{L}{2} \right| \right)$$

is an application of the exponential mechanism. The perturbations are independent of the data, and the final threshold is randomized post-processing of V; thus the mechanism is $(\varepsilon, 0)$-differentially private.

We next prove the welfare guarantee. Let

$$M = \max_i x_i, \qquad Z = \log(1 + M/\eta) + U,$$

where $U \sim \mathrm{Unif}[0,1]$ is independent, and let G denote the CDF of Z. Adding the independent uniform variable makes G 1-Lipschitz.

Fix v, let $T_v = \eta(e^{v-U} - 1)$, and define

$$q(v) := \Pr(M < T_v) = G(v), \qquad p_v(\mathbf{x}) := \Pr_U(M(\mathbf{x}) \ge T_v).$$

The equality $q(v) = G(v)$ follows directly from the definition of Z; strict versus weak inequalities are immaterial because U is continuous. We claim that

$$\mathbb{E}[\mathsf{ALG}(T_v, \mathbf{x})] \ge \min\left\{ q(v), 1 - \frac{1}{\mathbb{E}[e^{p_v(\mathbf{x})}]} \right\} \mathsf{OPT} - \eta. \tag{6}$$

To see this, let $F_M(t^-) = \Pr(M < t)$. For any deterministic threshold t, independence of the online values gives

$$\mathbb{E}[\mathsf{ALG}(t, \mathbf{x})] \ge (1 - F_M(t^-))t + F_M(t^-)\mathbb{E}[(M - t)^+].$$

Averaging over T_v, and using that T_v decreases with U, gives

$$\mathbb{E}[(1 - F_M(T_v^-))T_v] \ge \eta e^{v-1}\big(\mathbb{E}[e^{p_v(\mathbf{x})}] - 1\big) - \eta.$$

Moreover, $T_v \le \eta(e^v - 1)$, and hence

$$\mathbb{E}[F_M(T_v^-)(M - T_v)^+] \ge q(v)\mathbb{E}[(M + \eta - \eta e^v)^+].$$

Writing

$$A_v := \eta e^{v-1}\mathbb{E}[e^{p_v(\mathbf{x})}], \qquad C_v := \mathbb{E}[(M + \eta - \eta e^v)^+],$$

we therefore obtain

$$\mathbb{E}[\mathsf{ALG}(T_v, \mathbf{x})] \ge \min\left\{ q(v), 1 - \frac{1}{\mathbb{E}[e^{p_v(\mathbf{x})}]} \right\} (A_v + C_v) - \eta.$$

It remains to relate $A_v + C_v$ to OPT. Using the tail-integral representation of M,

$$\begin{aligned}\mathsf{OPT} &= \int_0^\infty \Pr(M > t)\, dt \\ &\le \eta e^{v-1} + \int_{\eta e^{v-1}}^{\eta e^v} \Pr(M + \eta > t)\, dt + \mathbb{E}[(M + \eta - \eta e^v)^+] \\ &= \eta e^{v-1}\mathbb{E}[e^{p_v(\mathbf{x})}] + \mathbb{E}[(M + \eta - \eta e^v)^+] \\ &= A_v + C_v.\end{aligned}$$

Substituting this above proves equation 6.

Call v *good* if

$$\left| G(v) - \frac{1}{2} \right| \le \frac{1}{16}.$$

For such a v, $q(v) \ge 7/16$. Moreover, $\mathbb{E}[p_v(\mathbf{x})] = 1 - q(v)$, so Jensen's inequality gives

$$\mathbb{E}[e^{p_v(\mathbf{x})}] \ge e^{1-q(v)} \ge e^{7/16} \ge \frac{20}{13}.$$

Thus both terms in the minimum in equation 6 are at least $7/20$, and therefore

$$\mathbb{E}[\mathsf{ALG}(T_v, \mathbf{x})] \ge \frac{7}{20}\mathsf{OPT} - \eta \qquad \text{for every good } v. \tag{7}$$

It remains to show that the mechanism selects a good V with high probability. Let

$$\mathcal{B} = \left\{ v \in [0, R] : \left| G(v) - \frac{1}{2} \right| > \frac{1}{16} \right\}, \qquad \mathcal{E} = \left\{ \sup_{v \in [0,R]} |\widehat{G}(v) - G(v)| \le \frac{1}{64} \right\}.$$

The variables $z_1, \ldots, z_L$ are i.i.d. with CDF G, so the Dvoretzky–Kiefer–Wolfowitz inequality gives

$$\Pr(\mathcal{E}^c) \le 2e^{-L/2048}.$$

Since G is continuous, let τ satisfy $G(\tau) = 1/2$. Moreover, $G(0) = 0$, $G(R) = 1$, and G is 1-Lipschitz, so

$$\tau \in [1/2, R - 1/2].$$

Hence

$$\mathcal{I} = \left[\tau - \frac{1}{64}, \tau + \frac{1}{64} \right] \subseteq [0, R].$$

Conditioned on $\mathcal{E}$, every $v \in \mathcal{I}$ satisfies

$$\left| \widehat{G}(v) - \frac{1}{2} \right| \le \frac{1}{32},$$

whereas every $v \in \mathcal{B}$ satisfies

$$\left| \widehat{G}(v) - \frac{1}{2} \right| \ge \frac{3}{64}.$$

Comparing the exponential-mechanism mass assigned to $\mathcal{B}$ and $\mathcal{I}$ therefore gives

$$\Pr(V \in \mathcal{B} \mid \mathcal{E}) \le 32R\, e^{-\varepsilon L/128}.$$

Consequently,

$$\Pr(V \in \mathcal{B}) \le 32R, e^{-\varepsilon L/128} + 2e^{-L/2048}. \tag{8}$$

Finally, set

$$y := \min\{1, \varepsilon\}L.$$

By the definition of η,

$$R = \log(1 + B/\eta) + 1 \le 3e^{y/10^5}.$$

Since $\varepsilon L \ge y$ and $L \ge y$, and since $1/128 - 10^{-5} \ge 1/200$, equation 8 implies

$$\Pr(V \in \mathcal{B}) \le 96e^{-y/200} + 2e^{-y/2048}.$$

For $y \ge 8000$, the right-hand side is at most $1/21$. Hence $\Pr(V \notin \mathcal{B}) \ge 20/21$. Using equation 7 on this event and non-negativity otherwise,

$$\begin{aligned}\mathbb{E}[\mathsf{ALG}] &\ge \Pr(V \notin \mathcal{B})\frac{7}{20}\mathsf{OPT} - \eta \\ &\ge \frac{20}{21} \cdot \frac{7}{20}\mathsf{OPT} - \eta \\ &= \frac{1}{3}\mathsf{OPT} - \eta.\end{aligned}$$

If $y < 8000$, then

$$\eta = Be^{-e^{y/10^5}} > Be^{-e^{0.08}} > \frac{B}{3} \ge \frac{\mathsf{OPT}}{3},$$

so the desired lower bound is non-positive and follows immediately from $\mathsf{ALG} \ge 0$. Substituting the definition of η completes the proof. □

A.2.1 Lower Bound

The upper bound above is essentially tight. The idea is that privacy prevents the threshold from adapting too sharply to the scale of the instance.

Consider K well-separated scales $v_j = B \cdot 8^{-j}$. At scale v_j, consider the deterministic instance

$$(v_j/8, v_j, 0, \ldots, 0).$$

Here $\mathsf{OPT} = v_j$. To obtain substantially more than $v_j/8$, the threshold must lie in $(v_j/8, v_j]$: if it is at most $v_j/8$, the algorithm accepts the first value $v_j/8$, while if it is larger than v_j, it accepts nothing.

These intervals are disjoint across different scales. Hence, on the all-zero sample dataset, there must be some j for which

$$\Pr(T(\mathbf{0}) \in (v_j/8, v_j]) \le \frac{1}{K}.$$

On the instance at scale v_j, the L sample vectors differ from the all-zero dataset in only $2L$ entries. Pure differential privacy therefore gives

$$\Pr\Big(T(\mathbf{s}^{(j)}) \in (v_j/8, v_j]\Big) \le \frac{e^{2\varepsilon L}}{K}.$$

Taking $K = \lceil 8e^{2\varepsilon L} \rceil$, this probability is at most $1/8$. Thus

$$\mathbb{E}[\mathsf{ALG}] \le \frac{1}{8} v_j + \frac{7}{8}\frac{v_j}{8} < \frac{1}{4} v_j.$$

Since $\mathsf{OPT} = v_j$, we lose at least $v_j/12$ compared with $\mathsf{OPT}/3$. Moreover, $j \le K$, so $v_j \ge B8^{-K}$. This gives the following lower bound.

Proposition A.3. *For every $(\varepsilon, 0)$-differentially private threshold mechanism using L sample vectors, there exists an instance supported on $[0, B]$ such that*

$$\mathbb{E}[\mathsf{ALG}] < \frac{1}{3}\mathsf{OPT} - B\exp\left(-54e^{2\varepsilon L}\right).$$

Thus the doubly exponential additive error in Theorem A.2 is unavoidable up to constants in the exponent. In particular, for $0 < \varepsilon \le 1$ and sufficiently small γ/B, if $L^*(\varepsilon, B, \gamma)$ denotes the minimum number of sample vectors needed to guarantee

$$\mathbb{E}[\mathsf{ALG}] \ge \frac{1}{3}\mathsf{OPT} - \gamma$$

for every instance, then

$$L^*(\varepsilon, B, \gamma) = \Theta\left(1 + \frac{1}{\varepsilon}\log\log\frac{B}{\gamma}\right).$$

Thus the optimal sample complexity is doubly logarithmic in B/γ.

B Index Privacy

B.1 Upper Bound

Theorem 3.1 (Index privacy from pooled samples)**.** *For every $n \ge 1$ and $L \ge 512$, there is an online mechanism that is $(512/L, 0)$-index private and, for every tuple $\mathcal{D}$ of independent distributions on $[0, \infty)$ with continuous cumulative distribution functions, is $(1/10)$-competitive against OPT:*

$$\mathbb{E}_{\mathbf{s}\sim\mathcal{D}^L, \mathbf{x}\sim\mathcal{D}}[\mathsf{ALG}(\mathbf{s}, \mathbf{x})] \ge \frac{1}{10}\mathsf{OPT}.$$

In fact, its privacy inequality holds for every fixed online sequence $\mathbf{x} \in \mathbb{R}^n_{\ge 0}$ and all events in $[n] \cup \{\perp\}$, including the outcome of selecting no element.

Proof of Theorem 3.1. We analyze the mechanism $\mathcal{M}$ given by Algorithm 2. We have $\varepsilon = 512/L$, the rank cutoff $k = \lfloor L/4 \rfloor$ and write

$$c(\mathbf{s}, t) := \#\{(\ell, j) \in [L] \times [n] : s_{\ell,j} > t\} \tag{9}$$

for the number of sample entries above t. The algorithm uses a shared noise $\eta_0 \sim \mathrm{Lap}(2/\varepsilon)$ and query noises $\eta_i \sim \mathrm{Lap}(4/\varepsilon)$, all mutually independent and independent of the data. It accepts the first index i satisfying

$$c(\mathbf{s}, x_i) \le k + \eta_i + \eta_0, \tag{10}$$

and returns $\perp$ if no such index exists. For the analysis, we draw all query noises in advance, including those for indices the algorithm never reaches. Thus the inequality is defined for every i; if the algorithm is still active when x_i arrives, it accepts i exactly when this inequality holds.

We prove privacy by showing that a small translation of the noises preserves the selected index when one sample entry changes.

Claim 1. *For every fixed online sequence* $\mathbf{x} \in \mathbb{R}^n_{\geq 0}$*, neighboring datasets* $\mathbf{s}, \mathbf{s}' \in \mathbb{R}^{L \times n}$*, and* $A \subseteq [n] \cup \{\perp\}$*,*

$$\mathbb{P}(\mathcal{M}(\mathbf{s}, \mathbf{x}) \in A) \leq e^{\varepsilon} \mathbb{P}(\mathcal{M}(\mathbf{s}', \mathbf{x}) \in A). \tag{11}$$

Proof of Claim 1. The datasets and online sequence are fixed, so probabilities in this claim are only over the mechanism's noises. By equation 9, changing one sample entry changes $c(\mathbf{s}, x_j)$ by at most one, for every j. Fix an index i. On the region of noise vectors for which the mechanism on $\mathbf{s}$ selects i, apply the translation

$$\eta_0 \longmapsto \eta_0 - 1, \qquad \eta_i \longmapsto \eta_i + 2, \tag{12}$$

leaving all other noises unchanged. We check that the translation in equation 12 preserves the selected index. Every earlier index $j < i$ is still rejected: its original rejection under equation 10 gives

$$c(\mathbf{s}', x_j) \geq c(\mathbf{s}, x_j) - 1 > k + \eta_j + (\eta_0 - 1).$$

Likewise, the original acceptance of i under equation 10 gives

$$c(\mathbf{s}', x_i) \leq c(\mathbf{s}, x_i) + 1 \leq k + (\eta_i + 2) + (\eta_0 - 1).$$

The final inequality is precisely the acceptance condition after equation 12. Thus the translation maps the output-i region for $\mathbf{s}$ injectively into the output-i region for $\mathbf{s}'$.

The shared noise has Laplace scale $2/\varepsilon$, so decreasing it by one changes its density by a factor of at most $e^{\varepsilon/2}$. The query noise has scale $4/\varepsilon$, so shifting it by two has the same cost. Integrating over the output-i region, and using that the translation has Jacobian one, gives

$$\mathbb{P}(\mathcal{M}(\mathbf{s}, \mathbf{x}) = i) \leq e^{\varepsilon} \mathbb{P}(\mathcal{M}(\mathbf{s}', \mathbf{x}) = i).$$

For output $\perp$, decreasing only η_0 by one keeps every index rejected and costs at most $e^{\varepsilon/2}$. Summing the singleton inequalities over A proves equation 11. □

We now prove the reward guarantee. Unless otherwise specified, probabilities and expectations are over the samples, online values, and mechanism's noises. Define the expected number of online values above some threshold t as

$$\lambda(t) := \sum_{i=1}^{n} \mathbb{P}(x_i > t), \qquad \lambda(a) = \frac{1}{2}, \qquad \lambda(b) = \frac{1}{8}. \tag{13}$$

Finite levels $0 < a < b$ exist because λ is continuous, $\lambda(0) = n$, and $\lambda(t) \to 0$ as $t \to \infty$. These levels are used only in the analysis. Moreover, let

$$H := \sum_{i=1}^{n} \mathbb{E}[(x_i - a)\mathbf{1}_{\{x_i \geq b\}}]. \tag{14}$$

That is, H is the total expected excess above a contributed by arrivals with value at least b.

The prophet's reward and our target. For every realization, $M(\mathbf{x}) \leq a + \sum_i (x_i - a)_+$. To bound the expected excess, separate the values into three ranges. Values below a contribute zero. Values in $[a, b)$ contribute at most $b - a$ each, and their expected number is $\lambda(a) - \lambda(b) = 3/8$ by continuity and equation 13. Values at least b contribute H in expected excess by (14). Taking expectations therefore gives

$$\mathsf{OPT} \leq a + \sum_i \mathbb{E}[(x_i - a)_+] \leq a + \frac{3}{8}(b - a) + H. \tag{15}$$

It remains to express $b - a$ in terms of H. Every value at least b has excess at least $b - a$, so equations (13) and (14) imply

$$H \geq (b - a) \sum_i \mathbb{P}(x_i \geq b) = (b - a)\lambda(b) = \frac{b - a}{8}.$$

Substituting $b - a \leq 8H$ into equation 15 yields

$$\mathsf{OPT} \leq a + 3H + H = a + 4H. \tag{16}$$

On the other hand, denote by R be the value selected by Algorithm 2, with $R = 0$ when no value is selected. A selection worth at least a contributes a base amount a, and a selection worth at least b also contributes its excess above a. Thus

$$\mathbb{E}[R] \geq a\mathbb{P}(R \geq a) + \mathbb{E}[(R-a)\mathbf{1}_{\{R\geq b\}}]. \tag{17}$$

To match one tenth of equation 16, we aim for $\mathbb{P}(R \geq a) \geq 1/10$ and an expected excess of at least $2H/5$. For the latter, it is sufficient to show that $\mathbb{P}(\mathcal{M}(\mathbf{s}, \mathbf{x}) = i \mid x_i = x) \geq 2/5$ for every i and $x \geq b$. Indeed, conditioning on each arriving value and summing over the possible selected indices then gives

$$\mathbb{E}[(R-a)\mathbf{1}_{\{R\geq b\}}] = \sum_i \int_{[b,\infty)} (x-a)\mathbb{P}(\mathcal{M}(\mathbf{s}, \mathbf{x}) = i \mid x_i = x)\, \mathcal{D}_i(dx) \geq \frac{2}{5}H,$$

where the inequality uses $x - a \geq 0$ on $[b, \infty)$ and the definition of H.

First, we record a simple identity that we use for summing probability bounds that depend on a value only through its aggregate tail mass.

Claim 2 (Integrating aggregate tail mass)**.** *For every nonnegative Borel function f,*

$$\sum_{i=1}^{n} \int f(\lambda(x))\, \mathcal{D}_i(dx) = \int_0^n f(u)\, du. \tag{18}$$

Proof of Claim 2. The mixture $n^{-1}\sum_i \mathcal{D}_i$ has continuous tail function λ/n by equation 13. For a draw x from this mixture, the probability-integral transform makes $\lambda(x)$ uniform on $[0, n]$. The sum of the marginal expectations is n times the mixture expectation, so

$$\sum_i \int f(\lambda(x))\, \mathcal{D}_i(dx) = n \int_0^n f(u)\, \frac{du}{n} = \int_0^n f(u)\, du.$$

This proves (18); the marginals need not have densities. □

We next establish two bounds: selecting a value below a is rare, and a fixed value at least b satisfies the acceptance inequality with high probability.

Claim 3 (Low selections and high-value acceptance)**.** *We have*

$$\mathbb{P}(R < a,\ \mathcal{M}(\mathbf{s}, \mathbf{x}) \neq \perp) \leq \frac{1}{1000}. \tag{19}$$

For every index i and fixed value $x \geq b$,

$$\mathbb{P}(c(\mathbf{s}, x) > k + \eta_i + \eta_0) \leq \frac{1}{12}. \tag{20}$$

The probability in equation 20 is over the samples and noises, without conditioning on reaching i.

Proof of Claim 3. Low selections. Fix $z < a$. We bound the probability that the acceptance inequality holds at this value. The sample count has mean $L\lambda(z)$, and $k \leq L/4 \leq L\lambda(z)/2$ by (13). Consequently, equation 10 requires at least one of

$$c(\mathbf{s}, z) \leq \frac{3L\lambda(z)}{4}, \qquad \eta_0 \geq \frac{L\lambda(z)}{8}, \qquad \eta_i \geq \frac{L\lambda(z)}{8}.$$

A Chernoff bound with relative deviation $1/4$ bounds the first event's probability by $e^{-L\lambda(z)/32}$. The other two probabilities are $\frac{1}{2}e^{-32\lambda(z)}$ and $\frac{1}{2}e^{-16\lambda(z)}$, since the Laplace scales are $L/256$ and $L/128$, respectively. Summing these bounds and using $L \geq 512$ to bound each exponential by $e^{-16\lambda(z)}$ gives

$$\begin{aligned}&\mathbb{P}(c(\mathbf{s}, z) \leq k + \eta_i + \eta_0) \\ &\quad \leq e^{-L\lambda(z)/32} + \frac{1}{2}e^{-32\lambda(z)} + \frac{1}{2}e^{-16\lambda(z)} \leq 2e^{-16\lambda(z)}.\end{aligned} \tag{21}$$

If the algorithm selects a value below a, some $x_i < a$ satisfies equation 10. We apply a union bound over indices, conditioning on $x_i = z$ and integrating equation 21 over $z \in [0, a)$.

To integrate over values of $z < a$, recall that $z < a$ implies $\lambda(z) \geq 1/2$ by (13). We therefore apply equation 18 with $f(u) = e^{-16u}\mathbf{1}_{\{u \geq 1/2\}}$. Finally, extending the resulting integral to infinity and evaluating it gives

$$\begin{aligned}\mathbb{P}(R < a,\ \mathcal{M}(\mathbf{s}, \mathbf{x}) \neq \perp) &\leq 2 \sum_i \int_{[0,a)} e^{-16\lambda(z)}\, \mathcal{D}_i(dz) \\ &\leq 2 \int_{1/2}^{n} e^{-16u}\, du \\ &\leq \frac{e^{-8}}{8} < \frac{1}{1000}.\end{aligned}$$

High-value acceptance. Fix $z \geq b$. We bound the probability that its count exceeds the noisy cutoff. By equation 13, the count has mean at most $L/8$ and, being a sum of independent indicators, variance at most $L/8$. Exceeding $3L/16$ therefore requires an upward deviation of at least $L/16$, to which we apply Chebyshev's inequality.

For the noises, $L/64$ and $L/32$ are four times the respective Laplace scales. Each lower-tail event below thus has probability $\frac{1}{2}e^{-4}$, so their union has probability at most e^{-4}. Using also $L \geq 512$, we obtain

$$\begin{aligned}\mathbb{P}\left(c(\mathbf{s}, z) > \frac{3L}{16}\right) &\leq \frac{L/8}{(L/16)^2} \leq \frac{1}{16}, \\ \mathbb{P}\left(\eta_0 < -\frac{L}{64} \text{ or } \eta_i < -\frac{L}{32}\right) &\leq e^{-4}.\end{aligned} \tag{22}$$

Outside these events, the count is at most $3L/16$ and the sum of the noises is at least $-3L/64$. Since $k \geq L/4 - 1$ and $L \geq 512$, these bounds ensure acceptance:

$$c(\mathbf{s}, z) \leq \frac{3L}{16} \leq k - \frac{3L}{64} \leq k + \eta_0 + \eta_i.$$

Thus the acceptance inequality can fail only on the events bounded in equation 22. A union bound gives a failure probability at most $1/16 + e^{-4} < 1/12$, proving (20). □

For the base reward, it is enough that a high value satisfies the acceptance inequality: the algorithm either selects it or has already selected something. By (19), an earlier selection is also worth at least a with high probability.

Claim 4 (Base reward)**.** *We have*

$$\mathbb{P}(R \geq a) \geq \frac{1}{10}. \tag{23}$$

Proof of Claim 4. We first bound the probability that a value at least b appears. Independence expresses the probability that all values are below b as a product. By continuity and equation 13, the sum of their probabilities of exceeding b is $\lambda(b) = 1/8$. Applying $1 - v \leq e^{-v}$ to the product, and then $1 - e^{-v} \geq v - v^2/2$, gives

$$\mathbb{P}(M(\mathbf{x}) \geq b) = 1 - \prod_i (1 - \mathbb{P}(x_i \geq b)) \geq 1 - e^{-1/8} \geq \frac{15}{128}. \tag{24}$$

We next explain why such an arrival provides the base reward. Fix an online sequence with $M(\mathbf{x}) \geq b$. By (20), the acceptance inequality at an index attaining the maximum holds with probability at least $11/12$. Whenever it holds, the algorithm either selects that index or has already selected an earlier one. Either outcome gives $R \geq a$ unless the algorithm selects a value below a.

We therefore average over the online sequence and subtract the probability of selecting a value below a, which is at most $1/1000$ by equation 19. This yields

$$\mathbb{P}(R \geq a) \geq \frac{11}{12}\mathbb{P}(M(\mathbf{x}) \geq b) - \frac{1}{1000}. \tag{25}$$

Finally, substituting equation 24 into (25) gives

$$\mathbb{P}(R \geq a) \geq \frac{11}{12} \cdot \frac{15}{128} - \frac{1}{1000} = \frac{55}{512} - \frac{1}{1000} > \frac{1}{10}.$$

□

For the excess reward, we need to select high values themselves. We show that the mechanism reaches each arrival often enough that subtracting (20) leaves a selection probability of at least $2/5$ for every value at least b.

Claim 5 (Excess reward)**.** *We have*

$$\mathbb{E}[(R-a)\mathbf{1}_{\{R \geq b\}}] \geq \frac{2}{5}H. \tag{26}$$

Proof of Claim 5. We first show that each fixed arriving value at least b has probability at least $2/5$ of being selected. This requires both reaching its arrival and satisfying the acceptance inequality.

Let A_i be the event that the mechanism reaches arrival i. Stopping earlier requires either a selection below a or a preceding value at least a. The first event has probability at most $1/1000$ by equation 19. For the second, continuity and (13) give $\sum_{j<i} \mathbb{P}(x_j \geq a) \leq \lambda(a) = 1/2$. A union bound therefore yields

$$\begin{aligned} \mathbb{P}(A_i) &\geq 1 - \sum_{j<i} \mathbb{P}(x_j \geq a) - \mathbb{P}(R < a,\ \mathcal{M}(\mathbf{s},\mathbf{x}) \neq \perp) \\ &\geq \frac{1}{2} - \frac{1}{1000}. \end{aligned} \tag{27}$$

Now fix the current value $x_i = z \geq b$. The event A_i depends only on the samples, noises, and preceding online values, so $\mathbb{P}(A_i \mid x_i = z) = \mathbb{P}(A_i)$. The algorithm selects i whenever it reaches this arrival and the acceptance inequality holds. Subtracting the probability that the inequality fails therefore gives a lower bound on selection. Using equations (27) and (20), we obtain

$$\begin{aligned} \mathbb{P}(\mathcal{M}(\mathbf{s},\mathbf{x}) = i \mid x_i = z) &\geq \mathbb{P}(A_i) - \mathbb{P}(c(\mathbf{s},z) > k + \eta_i + \eta_0) \\ &\geq \frac{1}{2} - \frac{1}{1000} - \frac{1}{12} > \frac{2}{5}. \end{aligned} \tag{28}$$

We next convert this selection probability into expected excess. Since at most one index is selected, we can condition on each x_i and sum its expected contribution. For $z \geq b$, the excess $z - a$ is nonnegative, so multiplying equation 28 by this excess and integrating preserves the inequality. The resulting sum is H by equation 14. Thus

$$\begin{aligned} \mathbb{E}[(R-a)\mathbf{1}_{\{R \geq b\}}] &= \sum_i \int_{[b,\infty)} (z-a)\mathbb{P}(\mathcal{M}(\mathbf{s},\mathbf{x}) = i \mid x_i = z)\, \mathcal{D}_i(dx) \\ &\geq \frac{2}{5} \sum_i \int_{[b,\infty)} (z-a)\, \mathcal{D}_i(dx) = \frac{2}{5}H. \end{aligned}$$

All integrands are nonnegative, so Tonelli's theorem also justifies the calculation when $H = \infty$. □

Finally, insert (23) and (26) into the reward decomposition equation 17. Applying the prophet bound (16) then gives

$$\mathbb{E}_{\mathbf{s},\mathbf{x}}[\mathsf{ALG}(\mathbf{s},\mathbf{x})] = \mathbb{E}[R] \geq \frac{a}{10} + \frac{2}{5}H = \frac{1}{10}(a + 4H) \geq \frac{1}{10}\mathsf{OPT}.$$

Together with (11), this proves the theorem. □

B.2 Lower Bound

Lemma B.1. *Under the same assumptions as Theorem 3.2, for any $m \in [n]$, we have*

$$\sum_{r=1}^{n} e^{-\varepsilon L|r-m|} \leq \frac{1}{c}. \tag{4}$$

Proof of Lemma 3.3. Fix a scale ratio $R > \max\{\frac{1}{c}, 2\}$ and a width $h \in (0, 1)$. Put $b_j = R^j$ and define distributions

$$Q_0^h = \mathrm{Unif}[0, h], \qquad Q_j^h = \mathrm{Unif}[b_j, b_j + h] \quad (j \geq 1).$$

Consider the following product distribution supported on $\mathbb{R}^n$:

$$P_r^h := Q_1^h \otimes \cdots \otimes Q_r^h \otimes \underbrace{Q_0^h \otimes \ldots \otimes Q_0^h}_{n-r \text{ times}} \tag{29}$$

Intuitively, the first r coordinates are concentrated in small intervals whose starting points are exponentially increasing, while the remaining $n - r$ coordinates are positive random numbers smaller than h. Now, consider a sequence distributed as P_r^h. Let $p_r(R, h)$ be the probability, over a sample from $(P_r^h)^L$, an independent online input vector from P_r^h, and the internal randomness of $\mathcal{M}$, that the selected winner is r.

By definition of P_r^h and since the distributions Q_j^h are supported on disjoint intervals, the value in coordinate r is always the largest online value. If $r \geq 2$, selecting r earns at most $b_r + h$, whereas every other choice of winner earns at most $b_{r-1}+h$. Since $\mathsf{OPT}(P_r^h) = b_r + h/2$, c-competitiveness implies

$$\begin{aligned} c(b_r + h/2) &\leq \mathbb{E}[R_{\mathcal{M}}] \\ &\leq p_r(R, h)(b_r + h) + (1 - p_r(R, h))(b_{r-1} + h), \end{aligned} \tag{30}$$

and therefore

$$p_r(R, h) \geq \frac{c(b_r + h/2) - (b_{r-1} + h)}{b_r - b_{r-1}}. \tag{31}$$

For $r = 1$, every other choice of winner earns at most h, and the same argument gives

$$p_1(R, h) \geq \frac{c(b_r + h/2) - h}{b_1}. \tag{32}$$

It follows, for fixed R and finite n, that

$$\liminf_{h \downarrow 0} \min_{r \in [n]} p_r(R, h) \geq \alpha_R, \qquad \alpha_R := \frac{c \cdot R - 1}{R - 1}. \tag{33}$$

Indeed, the limit of equation 31 is α_R, while the limit of equation 32 is $c \geq \alpha_R$.

In the following, for ease of notation, we omit the superscript h and write P_r in place of P_r^h. The proof proceeds as follows: we (1) build a sequence of datasets $\mathbf{s}^r$ for each $r \in [n]$ such that $s^r \sim P_r$, (2) consider the output of $\mathcal{M}$ given $\mathbf{s}^r$ and an independent input sequence drawn from P_n and (3) use the differential privacy assumption. Our argument relies on the following key observation. By definition, the distributions P_n and P_r have identical first r marginals, for each $r \in [n]$. Moreover, by assumption, for a fixed sample matrix $\mathbf{s}$, the event that $\mathcal{M}$ outputs r depends only on the online prefix through time r and the mechanism's randomness. Therefore, for every fixed realization of $\mathbf{s}^r$,

$$\mathbb{P}_{\mathbf{x} \sim P_r, \omega}[\mathcal{M}(\mathbf{s}^r, \mathbf{x}; \omega) = r] = \mathbb{P}_{\mathbf{x} \sim P_n, \omega}[\mathcal{M}(\mathbf{s}^r, \mathbf{x}; \omega) = r]. \tag{34}$$

We build the datasets $(\mathbf{s}^r)_{r \in [n]}$ sequentially as follows. For $r = n$, $\mathbf{s}^r \sim (P_n)^L$. For each $1 \leq r < n$, we keep the first r columns of $\mathbf{s}_{r+1}$ and substitute the column indexed $r + 1$ with L samples from Q_0^h. Observe that, by construction, each $\mathbf{s}^r$ is a sample from P_r. Moreover, fixing an index $m \in [n]$, we know that $\mathbf{s}^r$ and $\mathbf{s}_m$ differ in exactly $L|r - m|$ entries. Now, let us consider input sequences $\mathbf{x}$ drawn from P_n and the event that index r is selected as the winner. By applying the privacy guarantee of $\mathcal{M}$ we get that:

$$\mathbb{P}_{\mathbf{x} \sim P_n, \omega}[\mathcal{M}(\mathbf{s}^m, \mathbf{x}; \omega) = r] \geq e^{-\varepsilon L|r-m|} \mathbb{P}_{\mathbf{x} \sim P_n, \omega}[\mathcal{M}(\mathbf{s}^r, \mathbf{x}; \omega) = r]. \tag{35}$$

We emphasize that the privacy guarantee holds for every possible sample matrix, regardless of whether it was sampled from the actual distribution P_n. Summing both sides of equation 35 over r and using equation 34, we obtain

$$1 \geq \sum_{r=1}^{n} e^{-\varepsilon L|r-m|} \mathbb{P}_{\mathbf{x}\sim P_r,\omega}[\mathcal{M}(\mathbf{s}^r, \mathbf{x}; \omega) = r]. \tag{36}$$

Now, taking expectations over the coupled datasets yields and by definition of $p_r(R, h)$:

$$1 \geq \sum_{r=1}^{n} e^{-\varepsilon L|r-m|} p_r(R, H) \geq \left(\min_{r\in[n]} p_r(R, h) \right) \sum_{r=1}^{n} e^{-\varepsilon L|r-m|} \tag{37}$$

where the second inequality follows from equation 33. Letting $h \downarrow 0$ and $R \to \infty$, and using $\alpha_R \to c$, by equation 33 we obtain exactly equation 4. □