Reward Selection with Noisy Observations

Kamyar Azizzadenesheli KAMYARA@NVIDIA.COM
NVIDIA

Trung Dang DDDTRUNG@CS.UTEXAS.EDU
University of Texas at Austin

Aranyak Mehta ARANYAK@GOOGLE.COM
Google Research

Alexandros Psomas APSOMAS@PURDUE.EDU
Purdue University and Google Research

Qian Zhang ZHAN3761@PURDUE.EDU
Purdue University

Editors: Matus Telgarsky and Jonathan Ullman

Abstract

We study a fundamental problem in optimization under uncertainty. There are n boxes; each box i contains a hidden reward x_i. Rewards are drawn i.i.d. from an unknown distribution $\mathcal{D}$. For each box i, we see y_i, an unbiased estimate of its reward, which is drawn from a Normal distribution with known standard deviation σ_i (and an unknown mean x_i). Our task is to select a single box, with the goal of maximizing our reward. This problem captures a wide range of applications, e.g. ad auctions, where the hidden reward is the click-through rate of an ad. Previous work in this model (Bax et al., 2012) proves that the naive policy, which selects the box with the largest estimate y_i, is suboptimal, and suggests a linear policy, which selects the box i with the largest $y_i - c \cdot \sigma_i$, for some $c > 0$. However, no formal guarantees are given about the performance of either policy (e.g., whether their expected reward is within some factor of the optimal policy's reward).

In this work, we prove that both the naive policy and the linear policy are arbitrarily bad compared to the optimal policy, even when $\mathcal{D}$ is well-behaved, e.g. has monotone hazard rate (MHR), and even under a "small tail" condition, which requires that not too many boxes have arbitrarily large noise. On the flip side, we propose a simple threshold policy that gives a constant approximation to the reward of a prophet (who knows the realized values $x_1, \dots, x_n$) under the same "small tail" condition. We prove that when this condition is not satisfied, even an optimal clairvoyant policy (that knows $\mathcal{D}$) cannot get a constant approximation to the prophet, even for MHR distributions, implying that our threshold policy is optimal against the prophet benchmark, up to constants. En route to proving our results, we show a strong concentration result for the maximum of n i.i.d. samples from an MHR random variable that might be of independent interest.

Keywords: optimization under uncertainty, auctions, order statistics, threshold policies.

1. Introduction

Suppose that you are given n boxes, with box i containing a hidden reward x_i. Rewards are drawn independently and identically (i.i.d.) from an unknown distribution $\mathcal{D}$. For each box i, you see an unbiased estimate y_i of its reward: nature draws noise $\epsilon_i \sim \mathcal{N}(0, \sigma_i)$ with known σ_i, and you observe $y_i = x_i + \epsilon_i$. Your goal is to select the box with the highest reward x_i. This fundamental problem, originally introduced by Bax et al. (2012), captures a wide range of applications.

The original motivation of Bax et al. (2012) is ad auctions. As a concrete example, consider the following problem: there is a single ad slot for sale, and n advertisers. The platform knows the advertisers' bids $b_1, \ldots, b_n$, and has access to a machine learning model which outputs click-through-rate (CTR) predictions $y_1, \ldots, y_n$. Such algorithms typically have different amounts of data across different populations; therefore, there is different variance in the error of each y_i. If the platform chooses advertiser i, its profit will be $b_i \cdot x_i$, where x_i is the (true) click-through-rate of advertiser i; the platform wants to maximize this profit. For the sake of exposition, let's assume that all bids are equal: $b_1 = b_2 = \cdots = b_n$. If x_is are drawn from a *known* distribution $\mathcal{D}$, the platform should just calculate the posterior expectation $R_i(y_i) = \mathbb{E}[X_i \mid Y_i = y_i]$ for each advertiser i, and select the one with the largest $R_i(y_i)$. If x_is are drawn from an *unknown* distribution $\mathcal{D}$, this calculation is, of course, not possible. Then, shouldn't the platform just pick the advertiser i with the largest predicted click-through-rate y_i?

Bax et al. (2012) prove that Naive, the policy that picks the box/advertiser with the largest observation y_i, is suboptimal when the noise ϵ_i is drawn from $\mathcal{N}(0, \sigma_i)$. Specifically, they consider a family of *linear* policies. A linear policy with parameter c selects the box with the largest $y_i - c \cdot \sigma_i$; for $c = 0$ this corresponds to Naive. Bax et al. (2012) show that the derivative of the expected reward is strictly positive at $c = 0$; that is, the Naive policy is *not* optimal, even within the family of linear policies. However, and this brings us to our interest here, no other formal guarantees are given. Is the best linear policy, or even the Naive policy, a good (e.g., constant) approximation to the optimal policy? Are there better policies, outside the family of linear policies?

1.1. Our contribution

Without loss of generality, we assume that $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$ satisfies $\sigma_1 \leq \ldots \leq \sigma_n$. Naturally, if σ_i is large for almost all i, no policy, including a clairvoyant policy that knows $\mathcal{D}$, can hope to achieve any non-trivial performance guarantees (e.g., perform better than picking a random box). We start by making this intuition precise, i.e. quantifying "large" and "almost all." An important quantity for us will be $\mathbb{E}[\mathcal{D}_{n:n}]$,[1] the expected reward of a prophet that knows the rewards $x_1, \ldots, x_n$. Our definition of "largeness" for the noise is a logarithmic factor off of the reward of a prophet that can only pick from a specific number of boxes.

Informally, given $\mathcal{D}$, n and $c \in [0, 1]$, we say that $\boldsymbol{\sigma}$ has "large noise" if σ_{cn}, the cn-th smallest σ_i, is at least $\tilde{\Omega}(\mathbb{E}[\mathcal{D}_{cn:cn}])$. That is, some boxes are allowed to have very small noise, but enough boxes ($(1-c)n$) have noise at least $\tilde{\Omega}(\mathbb{E}[\mathcal{D}_{cn:cn}])$. Under this condition, we show that, even for the case of a distribution $\mathcal{D}$ with monotone hazard rate (MHR),[2] even an optimal clairvoyant policy (which knows $\mathcal{D}$) has reward comparable to the reward of picking a box uniformly at random. See Section 2 for the precise definitions, and Section 3 for the formal statements and proofs. We henceforth assume that the environment has "small noise:" informally, the cn-th smallest σ_i is at most $\tilde{O}(\mathbb{E}[\mathcal{D}_{cn:cn}])$. We first analyze the performance of known policies under this assumption.

We can show that the Naive policy, which selects the box with the highest reward, is not only suboptimal, but that it can be made suboptimal for *every* distribution $\mathcal{D}$ (Theorem 18). Specifically, given an arbitrary distribution $\mathcal{D}$, there exist choices for n and $\boldsymbol{\sigma}$ (satisfying the "small noise" assumption) such that the optimal (non-clairvoyant) policy has reward at least $\mathbb{E}[\mathcal{D}_{n:n}]/2$, while the Naive policy has a reward of at most $4\mathbb{E}[\mathcal{D}]$. Our construction has a small number, $\Theta(\log(n))$,

1. $\mathcal{D}_{k:n}$ is the k-th lowest of n i.i.d. samples from $\mathcal{D}$.
2. A distribution has monotone hazard rate (MHR) if $\frac{1-F(x)}{f(x)}$ is a non-increasing function.

boxes with large noise, with the remaining boxes having no noise. The intuition is that, with high probability, a random large noise box is chosen by Naive, while picking among the no noise boxes yields a reward of almost $\mathbb{E}[\mathcal{D}_{n:n}]$. Selecting $\mathcal{D}$ such that $\mathbb{E}[\mathcal{D}_{n:n}] \in \Theta(n\mathbb{E}[\mathcal{D}])$, we have that Naive provides only a trivial approximation to the optimal reward. Similarly, linear policies (the family of policies suggested by Bax et al. (2012)), can also be made suboptimal in a similarly strong way. Given an arbitrary MHR distribution $\mathcal{D}$, there exist choices for n and $\boldsymbol{\sigma}$ satisfying a "small noise for MHR" assumption, such that the optimal policy has reward at least a constant factor of $\mathbb{E}[\mathcal{D}_{n:n}]$, but no linear policy gets expected reward more than a constant factor of $\mathbb{E}[\mathcal{D}]$ (Theorem 25). By letting $\mathcal{D}$ be the exponential distribution, we get a lower bound of $\Omega(\log(n))$ for the approximation ratio of linear policies. Due to space constraints, these results are deferred to Section 5.1 and Section 5.2.

Theorem 25, the counter-example for linear policies, requires a delicate construction. En route to proving Theorem 25, we show a lemma about the concentration of the maximum of n i.i.d. samples from an MHR distribution, which might be of independent interest. Order statistics of MHR distributions satisfy the MHR condition (Barlow and Proschan, 1996). Additionally, MHR distributions exceed their mean with probability of at least $1/e$. Therefore, $\Pr[\mathcal{D}_{n:n} \geq \mathbb{E}[\mathcal{D}_{n:n}]] \geq 1/e$. We prove that $\mathcal{D}_{n:n}$ does not exceed twice its mean with high probability: $\Pr[\mathcal{D}_{n:n} \leq 2\mathbb{E}[\mathcal{D}_{n:n}]] \geq 1 - \frac{1}{n^{3/5}}$ (Lemma 8). The proof of this result is based on a new lemma (which again might be of independent interest) which states that the $(1 - 1/n)$-quantile value of an MHR distribution $\mathcal{D}$ is within a constant factor of $\mathbb{E}[\mathcal{D}_{n:n}]$.

Combined, Theorems 18 and 25 show that, even if we know that $\mathcal{D}$ belongs to the (arguably very well-behaved) family of MHR distributions, we need a new algorithm. At a high level, the downfall of both Naive and linear policies is that they treat very different types of boxes in a virtually identical manner: Naive does not take in the noise information at all, while linear policies utilize this information in a crude way, and discount boxes with different order of noise using the same weight. Intuitively, a good policy should identify large noise boxes and ignore them. However, a nontrivial obstacle is that the definition of "large noise" is relative to $\mathcal{D}$, which is unknown. In Section 4 we propose our new policy that circumvents this issue. The policy is quite elegant: pick $\alpha \sim U[0, 1]$, and run Naive on the α fraction of the boxes with the lowest noise. If a c fraction of the boxes has low noise, and specifically, if $\sigma_{cn} \leq \frac{\mathbb{E}[\mathcal{D}_{cn:cn}]}{5\sqrt{2\ln(n)}}$, then our policy gives a $\frac{c^2}{20}$ approximation to $\mathbb{E}[\mathcal{D}_{n:n}]$, the expected reward of a prophet (Theorem 15). Clearly, if c is a constant, we get a constant approximation. Interestingly, our policy provides the same guarantees even in a setting with a lot less information, where the σ_is are *unknown*, and only their order is available to the policy.

For the case of MHR distributions that satisfy the (incomparable) "small noise for MHR" assumption, we give an additional algorithm (i.e., for MHR distributions both algorithms are valid). The policy itself has a slight twist: pick $\alpha \sim U[0, 1]$, and run Naive on the n^α boxes with the lowest noise (i.e. boxes 1 through n^α). This time, we require that n^c boxes (a lot fewer than cn) have bounded noise. However, our bound on the noise is a lot smaller: $\sigma_{n^c} \leq \frac{\mathbb{E}[\mathcal{D}_{n^c:n^c}]}{18\sqrt{2\ln(n^c)}}$. Under these conditions (noting that these are the same conditions needed for the lower bound on linear policies in Theorem 25), this version of our policy guarantees a $c^2/576$ approximation to the prophet (Theorem 16). For a constant c, our approximation to the prophet is again a constant, and we only require n^c boxes with bounded noise.

Overall, ignoring logarithmic factors, our results provide a complete picture of when it is, and when it is not, possible to provide positive guarantees (algorithms that give a constant approximation to the prophet) in our setting. See Section 2.1, in particular Figures 1(*a*) and 1(*b*), for a visualization of the conditions required for the positive results to hold.

1.2. Related Work

Bax et al. (2012), whose contribution we already discussed, and Mahdian et al. (2023), are the two works most closely related to ours. Mahdian et al. (2023) study a very similar model to ours, where the reward x_i for each box i is not stochastic, but adversarial, and the noise distribution is not $\mathcal{N}(0, \sigma_i)$, but an arbitrary (known) zero-mean distribution A_i. Mahdian et al. (2023) are interested in finding policies with small worst-case *regret*, defined as the difference between the maximum reward and the expected performance of the policy, where the expectation is over only the random noise. A policy is then a constant approximation if its regret is within a constant of the optimal regret; in contrast, for us, a policy is a constant approximation if its expected *reward* is within a constant of the expected reward of the optimal policy/a prophet. Mahdian et al. (2023) show that in their model as well, the naive policy which picks the box with the highest observation y_i is arbitrarily bad (in terms of regret) even in the $n = 2$ case. Similar to our results here, Mahdian et al. (2023) show that there is a function θ from random variables to positive reals, such that picking the box with the largest $y_i - \theta(A_i)$ is a constant approximation (in terms of regret) to the optimal policy. Note that, in the case of our policy, this function is especially simple: $\theta(A_i) = 0$ if σ_i is small, otherwise $\theta(A_i)$ is infinite. We note, tangential to this discussion, a phenomenon related to the naive policy being suboptimal, both in the model studied here (the model of Bax et al. (2012)) as well as the model of Mahdian et al. (2023): the winner's curse (Thaler, 1988). In this phenomenon, multiple bidders with the same ex-post value for an item estimate this value independently and submit bids based on those estimates; the winner tends to have a bid that's an overestimate of the true value.

Related to the problem studied here is *delegated choice*, introduced by Khodabakhsh et al. (2024) with some recent work by Bowers et al. (2025). In delegated choice, the principal must pick one of n actions, each of which has a stochastic reward x_i without directly viewing the rewards. They can choose a subset of these actions and delegate them to an agent, who views biased rewards $x_i + b_i$ and picks an action that maximizes biased reward. Our problem is closely related to the case where the biases b_i are also random variables. Another related problem is robust optimization, where we seek solutions that are robust with respect to the realization of uncertainty; see Bertsimas et al. (2011) for a survey.

Apart from the aforementioned settings, our problem bears some (superficial) similarity to the area of algorithms with predictions; see Mitzenmacher and Vassilvitskii (2022); Balkanski et al. (2024) for recent surveys. As opposed to our problem, predictions in this literature are not random variables. Instead, the goal is to, given a prediction (side information) about the instance, design an algorithm that is near-optimal when the prediction is accurate (consistency) and whose performance degrades gracefully as a function of the prediction error, while retaining worst-case guarantees (robustness). Finally, there has been a lot of work on the related problem of finding the maximum (or the top k elements) given noisy information, see, e.g., Feige et al. (1994); Braverman et al. (2016, 2019); Cohen-Addad et al. (2020).

Some of our results assume that $\mathcal{D}$ is MHR. MHR distributions satisfy numerous useful properties, see Barlow and Proschan (1996) for a textbook. In algorithmic economics, such properties have been exploited to enable positive results for several problems, e.g., the sample complexity of revenue maximization (Dhangwatnotai et al., 2010; Cole and Roughgarden, 2014; Huang et al., 2015; Guo et al., 2019, 2021), the competition complexity of dynamic auctions (Liu and Psomas, 2018), and the design of optimal and approximately optimal auctions (Hartline and Roughgarden, 2009; Daskalakis and Weinberg, 2012; Cai and Daskalakis, 2011; Allouah and Besbes, 2020; Giannakopoulos et al., 2021).

2. Preliminaries

There are n boxes. The i-th box contains a reward x_i. These rewards are drawn i.i.d. from an *unknown* distribution $\mathcal{D}$ with a cumulative distribution function F and density function f. We assume that $\mathcal{D}$ is supported on $[0, \infty)$. Rewards are not observed by our algorithm. Instead, nature draws unbiased estimates, $y_1, \dots, y_n$, where y_i is drawn from a normal distribution with (an unknown) mean x_i and a *known* standard deviation σ_i. We refer to y_i as the i-th observation. We often write X_i and Y_i for the random variable for the i-th reward and i-th observation, respectively. Note that Y_i can be equivalently thought of as $Y_i = X_i + \epsilon_i$, where the noise ϵ_i is drawn from $\mathcal{N}(0, \sigma_i)$. Our goal is to select a single box i with the goal of maximizing the (expected) realized reward.

Policies and expected rewards. Formally, a policy A maps the public information, the pair $(\boldsymbol{\sigma}, \boldsymbol{y})$, $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_n)$ and $\boldsymbol{y} = (y_1, \dots, y_n)$, to a distribution over boxes. We write $R_A(\mathcal{D}, \boldsymbol{\sigma}, \boldsymbol{y})$ for the expected reward of a policy A under true reward distribution $\mathcal{D}$ and observations $\boldsymbol{y} = (y_1, \dots, y_n)$, where the standard deviation of the noise is according to $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_n)$, and where this expectation is with respect to the randomness of A and the randomness of the rewards. In order to evaluate a policy under a fixed reward distribution $\mathcal{D}$ we need to take an additional expectation over the random observations $\boldsymbol{y} = (y_1, \dots, y_n)$. We overload notation and write $R_A(\mathcal{D}, \boldsymbol{\sigma}) = \mathbb{E}_{\boldsymbol{y}}\left[R_A(\mathcal{D}, \boldsymbol{\sigma}, \boldsymbol{y})\right]$ for the expected reward of a policy A under true reward distribution $\mathcal{D}$, where the standard deviation of the noise is according to $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_n)$.

Previous policies and benchmarks. Bax et al. (2012) consider two simple policies. The Naive policy always selects the box i with the largest observation y_i. A linear policy Linear_γ, parameterized by a function $\gamma : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, chooses the box i which maximizes $y_i - \gamma(\boldsymbol{\sigma}, \boldsymbol{y}) \cdot \sigma_i$.

We use the following two policies as useful benchmarks: the *optimal clairvoyant policy*, and the *prophet*. The optimal clairvoyant policy for a distribution $\mathcal{D}$, $\mathsf{Opt}_{\mathcal{D}}$, selects the box i with maximum $\mathbb{E}\left[X_i \mid Y_i = y_i\right]$. Its expected reward in outcome $\boldsymbol{y}$ is precisely $\max_i \mathbb{E}\left[X_i \mid Y_i = y_i\right]$: $R_{\mathsf{Opt}_{\mathcal{D}}}(\mathcal{D}, \boldsymbol{\sigma}) = \mathbb{E}_{\boldsymbol{y}}\left[\max_{i\in[n]} \mathbb{E}\left[X_i \mid Y_i = y_i\right]\right]$. Finally, the (expected) reward of a prophet who knows $x_1, \dots, x_n$ is equal to $\mathbb{E}[\mathcal{D}_{n:n}]$, the expected maximum of n i.i.d. draws from $\mathcal{D}$.

Formalizing "small" and "large" noise environments. Clearly, if σ_i is large for almost all $i \in [n]$, then no policy can hope to get a non-trivial guarantee. Therefore, we intuitively need a condition that captures the fact that we need small noise for enough boxes. In the following couple of definitions, we formalize precisely what we mean by "small" and "enough."

Definition 1 (Small noise) *For any distribution $\mathcal{D}$, any n and any $c \in (0, 1]$, let $\mathcal{S}_{(\mathcal{D},n,c)}$ be the set of vectors $\boldsymbol{\sigma} \in \mathbb{R}^n_+$ where at least cn values in $\boldsymbol{\sigma}$ are small, namely at most $\frac{\mathbb{E}[\mathcal{D}_{cn:cn}]}{5\sqrt{2\ln n}}$. Formally, $\mathcal{S}_{(\mathcal{D},n,c)} = \{\boldsymbol{\sigma} \in \mathbb{R}^n_+ \mid \sigma_1 \leq \dots \leq \sigma_n$ and $\sigma_{cn} \leq \frac{\mathbb{E}[\mathcal{D}_{cn:cn}]}{5\sqrt{2\ln n}}\}$.*

For the case of MHR distributions, the following incomparable condition also suffices to guarantee strong positive results. We also use this condition in our lower bound on linear policies.

Definition 2 (Small noise for MHR) *For any MHR distribution $\mathcal{D}$ and any n, let $\mathcal{S}^{\text{MHR}}_{(\mathcal{D},n,c)}$ be the set of vectors $\boldsymbol{\sigma} \in \mathbb{R}^n_+$ where at least n^c values in $\boldsymbol{\sigma}$ are small, namely at most $\frac{\mathbb{E}[\mathcal{D}_{n^c:n^c}]}{18\sqrt{2\ln(n^c)}}$. Formally, $\mathcal{S}^{\text{MHR}}_{(\mathcal{D},n,c)} = \{\boldsymbol{\sigma} \in \mathbb{R}^n_+ \mid \sigma_1 \leq \cdots \leq \sigma_n$ and $\sigma_{n^c} \leq \frac{\mathbb{E}[\mathcal{D}_{n^c:n^c}]}{18\sqrt{2\ln(n^c)}}\}$.*

Ideally, whenever $\boldsymbol{\sigma} \notin \mathcal{S}_{(\mathcal{D},n,c)}$ or $\boldsymbol{\sigma} \notin \mathcal{S}^{\text{MHR}}_{(\mathcal{D},n,c)}$, we would like strong negative results for the optimal policy. Our negative results are with respect to the optimal *clairvoyant* policy and hold even for MHR distributions (so, very strong negative results), but our condition is a bit weaker than $\boldsymbol{\sigma}$ not being in the complement of $\mathcal{S}_{(\mathcal{D},n,c)}$: we lose an extra $\sqrt{\ln(n)/\ln(cn)}$ factor. Under the following "large noise" condition, we cannot do better than picking a box uniformly at random (Theorem 11).

Definition 3 (Large noise) *For any distribution $\mathcal{D}$, any n and any $c \in (0, 1]$, let $\mathcal{L}_{(\mathcal{D},n,c)}$ be the set of vectors $\boldsymbol{\sigma} \in \mathbb{R}^n_+$ where at most cn values in $\boldsymbol{\sigma}$ are small, namely at most $\frac{\mathbb{E}[\mathcal{D}_{cn:cn}]\cdot\sqrt{\ln n}}{\ln(cn)}$. Formally, $\mathcal{L}_{(\mathcal{D},n,c)} = \{\boldsymbol{\sigma} \in \mathbb{R}^n_+ \mid \sigma_1 \leq \cdots \leq \sigma_n$ and $\sigma_{cn} > \frac{\mathbb{E}[\mathcal{D}_{cn:cn}]\cdot\sqrt{\ln n}}{\ln(cn)}\}$.*

We emphasize that all three definitions above are parameterized by a free parameter c, which intuitively represents the fraction of "small noise" boxes. This means that a noise vector $\boldsymbol{\sigma}$ may satisfy a definition for multiple choices of c. For example, a zero noise vector satisfies Definitions 1 and 2 for all choices of $c \in (0, 1]$. Observe a direct relationship between this fraction c and the magnitude of what constitutes "small noise" in all three definitions: if we only have a lower fraction c of "small noise" boxes, then the σ's of these boxes also need to be substantially lower. This makes sense intuitively: with a lower fraction of "small noise" boxes, we need these boxes to behave more deterministically to extract better reward from them.

2.1. Visualizing our main positive results

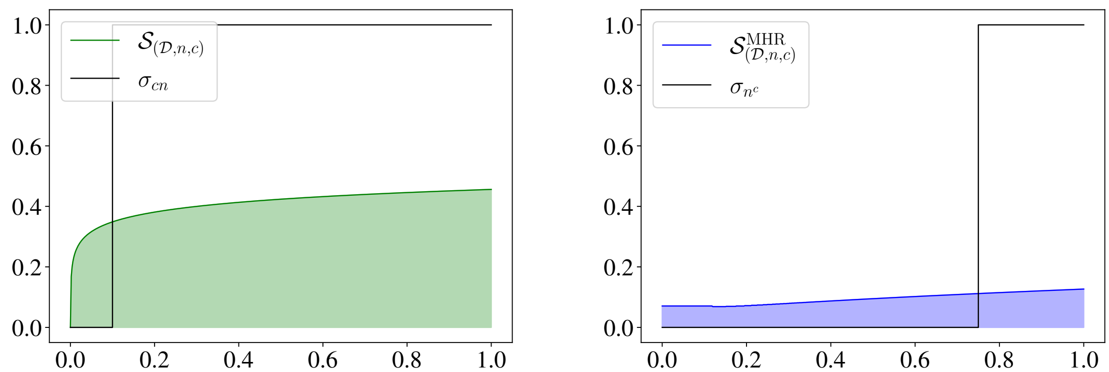
Given the definitions so far, we informally restate and visualize our main positive results.

1. Theorem 15: For all distributions $\mathcal{D}$, if $\boldsymbol{\sigma} \in \mathcal{S}_{(\mathcal{D},n,c)}$ for some $c \in [1/n, 1]$, then the IgnoreLarge policy obtains an expected reward of at least $\frac{c^2}{20}$ times that of the prophet.
2. Theorem 16: For all MHR distributions $\mathcal{D}$, if $\boldsymbol{\sigma} \in \mathcal{S}^{\text{MHR}}_{(\mathcal{D},n,c)}$ for some $c \in (0, 1]$, then the IgnoreLargeExp policy obtains an expected reward of at least $\frac{c^2}{576}$ times that of the prophet.

Let us first visualize the first result. For any fixed distribution $\mathcal{D}$, number of boxes n, and standard deviations $\sigma_1 \leq \cdots \leq \sigma_n$ sorted in non-decreasing order, we plot the two central quantities in Definition 1 for $\mathcal{S}_{(\mathcal{D},n,c)}$: $\frac{\mathbb{E}[\mathcal{D}_{cn:cn}]}{5\sqrt{2\ln n}}$ and σ_{cn}. For any point c where σ_{cn} is below $\frac{\mathbb{E}[\mathcal{D}_{cn:cn}]}{5\sqrt{2\ln n}}$, if $\boldsymbol{\sigma}$ lies in $\mathcal{S}_{(\mathcal{D},n,c)}$ we obtain a positive result from Theorem 15 for that particular c.

In Figure 1(*a*), we chose $\mathcal{D}$ to be the standard exponential distribution, $n = 10\,000$, and $\sigma_i = 0$ for all $i \leq 1000$ and 1 otherwise. We have that σ_{cn} stays below $\mathcal{S}_{(\mathcal{D},n,c)}$ for c up to 0.1, implying that, by Theorem 15, IgnoreLarge can achieve a $\frac{0.1^2}{20}$ approximation against the prophet.

Similarly, in Figure 1(*b*) we plot the two central quantities in Definition 2 for $\mathcal{S}^{\text{MHR}}_{(\mathcal{D},n,c)}$: $\frac{\mathbb{E}[\mathcal{D}_{n^c:n^c}]}{18\sqrt{2\ln(n^c)}}$ and σ_{n^c}, for the same choice of $\mathcal{D}$, n, and σ_i's. Here, σ_{n^c} is below $\mathcal{S}^{\text{MHR}}_{(\mathcal{D},n,c)}$ for c up to 0.75, implying, by Theorem 16, that IgnoreLargeExp achieves a $\frac{0.75^2}{576}$ (better than $0.1^2/20$) approximation against the prophet.



(*a*) Plot comparing σ_{cn} and $\mathcal{S}_{(\mathcal{D},n,c)}$ for $n = 10\,000$, and $\sigma_i = 0$ for all $i \leq 1000$ and 1 otherwise.

(*b*) Plot comparing σ_{n^c} and $\mathcal{S}^{\text{MHR}}_{(\mathcal{D},n,c)}$ for $n = 10\,000$, and $\sigma_i = 0$ for all $i \leq 1000$ and 1 otherwise.

Figure 1: Visualizations of our main positive results.

2.2. Technical lemmas

Here, we present some definitions and a few technical lemmas that will be useful throughout the paper. We often use the following lemma (Lemma 4) about the CDF of the standard normal distribution, and a lemma (Lemma 5; proof is deferred to Appendix B) about the relation between the expected maximum of a and b i.i.d. samples from an arbitrary distribution $\mathcal{D}$. We write $\mathcal{D}_{k:n}$ for the k-th lowest order statistic out of n i.i.d. samples, that is, $\mathcal{D}_{1:n} \leq \mathcal{D}_{2:n} \leq \cdots \leq \mathcal{D}_{n:n}$. Throughout the paper, $\Phi(x)$ is the CDF of the standard normal distribution, and $\phi(x)$ is the PDF of the standard normal distribution.

Lemma 4 (Gordon (1941)) *For all $t > 0$, we have $1 - \frac{1}{\sqrt{2\pi}} \frac{1}{t} e^{-t^2/2} \leq \Phi(t) \leq 1 - \frac{1}{\sqrt{2\pi}} \frac{t}{t^2+1} e^{-t^2/2}$. Furthermore, this implies directly that for all $t > 0$, $1 - \frac{\phi(t)}{t} \leq \Phi(t) \leq 1 - \frac{t\phi(t)}{t^2+1}$.*

Lemma 5 *For any distribution $\mathcal{D}$ and integers $1 \leq a < b$, $\frac{\mathbb{E}[\mathcal{D}_{a:a}]}{a} \geq \frac{\mathbb{E}[\mathcal{D}_{b:b}]}{b}$.*

The following definitions will be crucial in describing our lower bounds.

Definition 6 *Let $\alpha_m^{(\mathcal{D})} = \inf\{x \mid F(x) \geq 1 - \frac{1}{m}\}$ be the $(1 - \frac{1}{m})$-th quantile of a distribution $\mathcal{D}$.*

Definition 7 *For a distribution $\mathcal{D}$, let $\beta_m^{(\mathcal{D})} = \inf\{x : \mathbb{E}[\mathcal{D} \mid \mathcal{D} \geq x] \Pr[\mathcal{D} \geq x] \leq \frac{\mathbb{E}[\mathcal{D}]}{m}\}$ be the smallest x such that the contribution to $\mathbb{E}[\mathcal{D}]$ from values at least x is at most $\frac{\mathbb{E}[\mathcal{D}]}{m}$.*

Technical lemmas for MHR distributions. We prove the following technical lemma for the concentration of the maximum of n i.i.d. samples of an MHR distribution, which might be of independent interest.

Lemma 8 *For any MHR distribution $\mathcal{D}$ and any $n \geq 4$, we have $\Pr[\mathcal{D}_{n:n} < 2\,\mathbb{E}[\mathcal{D}_{n:n}]] \geq 1 - \frac{1}{n^{3/5}}$.*

It is known that the maximum of i.i.d. draws from an MHR distribution is also MHR (Barlow and Proschan, 1996). This implies that the probability that the maximum exceeds its mean, $\Pr[\mathcal{D}_{n:n} \geq \mathbb{E}[\mathcal{D}_{n:n}]]$, is at least $1/e$. In Lemma 8 we show that, in fact, this maximum concentrates around its mean: it does not exceed twice its mean with high probability. We note that a related, but incomparable, statement is given by Cai and Daskalakis (2011), who show that at least a $(1-\epsilon)$-fraction of $\mathbb{E}[\max_i X_i]$ is contributed by values no larger than $\mathbb{E}[\max_i X_i] \cdot \log(\frac{1}{\epsilon})$, where the X_is are (possibly not identical) MHR distributions.

Lemma 8 is an immediate consequence of the following two lemmas. Missing proofs are deferred to Appendix B.

Lemma 9 (Cai and Daskalakis (2011); Lemma 34) *If the distribution of a random variable X satisfies MHR, then for all $m \geq 1$ and $d \geq 1$, we have $d\alpha_m^{(X)} \geq \alpha_{m^d}^{(X)}$.*

Lemma 10 *For any MHR distribution $\mathcal{D}$ and any $n \geq 4$, we have $\frac{\mathbb{E}[\mathcal{D}_{n:n}]}{3} \leq \alpha_n^{(\mathcal{D})} \leq \frac{5\,\mathbb{E}[\mathcal{D}_{n:n}]}{4}$.*

Proof of Lemma 8 We have $\alpha_{n^{8/5}}^{(\mathcal{D})} \leq^{\text{(Lemma 9)}} \frac{8}{5}\alpha_n^{(\mathcal{D})} \leq^{\text{(Lemma 10)}} 2\mathbb{E}[\mathcal{D}_{n:n}]$. Therefore,

$$\Pr[\mathcal{D}_{n:n} \leq 2\mathbb{E}[\mathcal{D}_{n:n}]] \geq \Pr[\mathcal{D}_{n:n} \leq \alpha_{n^{8/5}}^{(\mathcal{D})}] = \left(1 - \frac{1}{n^{8/5}}\right)^n \geq^{\text{(Bernoulli's inequality)}} 1 - \frac{1}{n^{3/5}}.$$

■

3. Negative results for large noise environments

Before discussing small noise environments, we show strong lower bounds for the optimal clairvoyant policy (an optimal policy that knows $\mathcal{D}$) in large noise environments, even under the assumption that the distribution $\mathcal{D}$ is MHR. All missing proofs can be found in Appendix C.

Theorem 11 shows that, for an MHR distribution $\mathcal{D}$, when $\boldsymbol{\sigma} \in \mathcal{L}_{(\mathcal{D},n,c)}$, then the optimal clairvoyant policy is comparable to the policy that picks a random box. First, as we discussed in Section 2, note that $\mathcal{S}_{(\mathcal{D},n,c)}$ is *almost*, but not exactly, the complement of $\mathcal{L}_{(\mathcal{D},n,c)}$; the complement of $\mathcal{L}_{(\mathcal{D},n,c)}$ includes $\boldsymbol{\sigma}$ where at least cn values in $\boldsymbol{\sigma}$ are at most $\frac{\mathbb{E}[\mathcal{D}_{cn:cn}]\cdot\sqrt{\ln n}}{\ln(cn)}$, while $\mathcal{S}_{(\mathcal{D},n,c)}$ is characterized by $\boldsymbol{\sigma}$'s containing at least cn values upper bounded by $\frac{\mathbb{E}[\mathcal{D}_{cn:cn}]}{5\sqrt{2\ln n}}$, implying that $\mathcal{S}_{(\mathcal{D},n,c)}$ is a strict subset of the complement of $\mathcal{L}_{(\mathcal{D},n,c)}$ as $c \leq 1$ (since $\frac{\sqrt{\ln n}}{\ln cn} \geq \frac{1}{\sqrt{\ln n}} > \frac{1}{5\sqrt{2\ln n}}$). This leaves a gap (arguably insignificant, but a gap nonetheless) in our understanding. On the flip side, our negative result holds against the (well-behaved) class of MHR distributions, even against the strong benchmark of the optimal clairvoyant policy.

Theorem 11 *There exists an MHR distribution $\mathcal{D}$ where $\mathbb{E}[\mathcal{D}_{k:k}] \in \omega(\mathbb{E}[D])$ for $k \in \omega(1)$, such that for all $n \geq n_0$, for some constant n_0, for all $c \in [1/n, 1]$, and for all $\boldsymbol{\sigma} \in \mathcal{L}_{(\mathcal{D},n,c)}$, we have*

$$R_{\mathsf{Opt}_{\mathcal{D}}}(\mathcal{D}, \boldsymbol{\sigma}) \in O\left(\sqrt{\ln(cn)} \cdot \mathbb{E}[\mathcal{D}]\right).$$

One way to interpret this theorem is: given any constant target ratio α and any large enough n, one can pick c small enough (e.g. such that $cn \in O(1)$) and $\boldsymbol{\sigma}$ that satisfies the "large" noise condition, such that the optimal clairvoyant policy is not α times better than the policy that picks a box uniformly at random. The $\mathbb{E}[\mathcal{D}_{k:k}] \in \omega(\mathbb{E}[D])$ is crucial in this theorem, since, for the theorem to have bite, it must be that $\sqrt{\ln(cn)} \cdot \mathbb{E}[\mathcal{D}]$ is a lot smaller than $\mathbb{E}[\mathcal{D}_{n:n}]$ the reward of a prophet. We include the fact that $\mathbb{E}[\mathcal{D}_{k:k}] \in \omega(\mathbb{E}[D])$, to highlight that the distribution is not trivial.

The distribution that witnesses Theorem 11 is the half-normal distribution $\mathcal{D} = |\mathcal{N}(0,1)|$. We start by proving that this distribution is MHR, and bounding its expected maximum value.

Lemma 12 *Consider the distribution $\mathcal{D} = |\mathcal{N}(0,1)|$. $\mathcal{D}$ is MHR, and it holds that (i) $\mathbb{E}[\mathcal{D}] = \sqrt{\frac{2}{\pi}}$, and (ii) $\frac{4}{5}\sqrt{\ln n} \le \mathbb{E}[\mathcal{D}_{n:n}] \le 3\sqrt{2}\sqrt{\ln n}$, for $n \ge 8$.*

Since order statistics are preserved under affine transformations, we get the following corollary.

Corollary 13 *For all $\sigma > 0$, $\frac{4}{5} \cdot \sigma\sqrt{\ln n} \le \mathbb{E}\left[|\mathcal{N}(0,\sigma^2)|_{n:n}\right] \le 3\sqrt{2} \cdot \sigma\sqrt{\ln n}$ for $n \ge 8$.*

Towards bounding the optimal policy, we get the following bound on $\mathbb{E}[X_i \mid Y_i = y_i]$.

Lemma 14 *Let $U_\sigma(y) = \sqrt{\frac{2}{\pi}} + \max\left\{0, \frac{y}{\sigma^2+1}\right\}$, then $\mathbb{E}[X_i \mid Y_i = y_i] \le U_{\sigma_i}(y_i)$ for all σ_i and y_i.*

We are now ready to prove Theorem 11.

Proof of Theorem 11 Let $\mathcal{D} = |\mathcal{N}(0,1^2)|$, and consider $\boldsymbol{\sigma} \in \mathcal{L}_{(\mathcal{D},n,c)}$ where, without loss of generality, we have $\sigma_1 \le \sigma_2 \le \cdots \le \sigma_n$. This means that $\sigma_{cn} > \frac{\mathbb{E}[\mathcal{D}_{cn:cn}] \cdot \sqrt{\ln(n)}}{\ln(cn)}$. Note that the expected reward of the optimal policy is at most the expected reward of the optimal policy that picks 2 boxes u and v where $u \in [1, cn-1]$ and $v \in [cn, n]$, and *then enjoys the rewards of both boxes*.

The expected reward from choosing box u is at most $\mathbb{E}[\max_{i\in[1,cn-1]} x_i] \le \mathbb{E}[\mathcal{D}_{cn:cn}]$. The expected reward from choosing box v is at most the expected reward of $\mathsf{Opt}_{\mathcal{D}}$ conditioned on it choosing boxes from cn to n, which in turn is at most $\max_{i\in[cn,n]} \mathbb{E}[X_i \mid Y_i = y_i]$. Therefore, the expected reward from box v is upper bounded by:

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{y}}\left[\max_{i\in[cn,n]} \mathbb{E}[X_i \mid Y_i = y_i]\right] &\le^{\text{(Lemma 14)}} \mathbb{E}_{\boldsymbol{y}}\left[\max_{i\in[cn,n]} U_{\sigma_i}(y_i)\right] \\
&= \mathbb{E}\left[\max_{i\in[cn,n]} U_{\sigma_i}\left(X_i + \mathcal{N}(0,\sigma_i^2)\right)\right] \\
&\le^{(U_{\sigma_i}(y)\text{ is monotone})} \mathbb{E}\left[\max_{i\in[cn,n]} U_{\sigma_i}\left(X_i + |\mathcal{N}(0,\sigma_i^2)|\right)\right] \\
&= \mathbb{E}\left[\max_{i\in[cn,n]} \sqrt{\frac{2}{\pi}} + \frac{(X_i + |\mathcal{N}(0,\sigma_i^2)|)}{\sigma_i^2+1}\right] \\
&\le \mathbb{E}\left[\sqrt{\frac{2}{\pi}} + \max_{i\in[cn,n]} \frac{X_i}{\sigma_i^2} + \max_{i\in[cn,n]} \frac{|\mathcal{N}(0,\sigma_i^2)|}{\sigma_i^2}\right] \\
&\le \sqrt{\frac{2}{\pi}} + \frac{\mathbb{E}\left[|\mathcal{N}(0,1)|_{n:n}\right]}{\sigma_{cn}^2} + \mathbb{E}\left[\max_{i\in[cn,n]} \left|\mathcal{N}\left(0, \frac{1}{\sigma_i^2}\right)\right|\right]
\end{aligned}$$

$$
\begin{aligned}
&\leq^{\text{(Corollary 13)}} \sqrt{\frac{2}{\pi}} + \frac{3\sqrt{2}\cdot\sqrt{\ln n}}{\sigma_{cn}^2} + \frac{1}{\sigma_{cn}}\cdot 3\sqrt{2}\cdot\sqrt{\ln n} \\
&\leq \mathbb{E}[\mathcal{D}] + \frac{6\sqrt{2}\cdot\sqrt{\ln n}}{\sigma_{cn}} \\
&\leq^{\left(\sigma_{cn} > \frac{\mathbb{E}[\mathcal{D}_{cn:cn}]\cdot\sqrt{\ln(n)}}{\ln(cn)}\right)} \mathbb{E}[\mathcal{D}] + \frac{6\sqrt{2}\cdot\ln(cn)}{\mathbb{E}[\mathcal{D}_{cn:cn}]} \\
&\leq^{\text{(Lemma 12)}} \mathbb{E}[\mathcal{D}] + \frac{6\sqrt{2}\cdot\frac{25}{16}(\mathbb{E}[\mathcal{D}_{cn:cn}])^2}{\mathbb{E}[\mathcal{D}_{cn:cn}]} \\
&\leq^{(cn\geq 1)} 15\mathbb{E}[\mathcal{D}_{cn:cn}].
\end{aligned}
$$

Combining, we get $R_{\mathsf{Opt}_{\mathcal{D}}}(\mathcal{D},\boldsymbol{\sigma}) \leq 16\mathbb{E}[\mathcal{D}_{cn:cn}]$. As

$$\mathbb{E}[\mathcal{D}_{cn:cn}] \leq 3\sqrt{2}\sqrt{\ln(cn)} = 3\sqrt{\pi}\sqrt{\ln(cn)}\mathbb{E}[\mathcal{D}]$$

by Lemma 12, we have

$$R_{\mathsf{Opt}_{\mathcal{D}}}(\mathcal{D},\boldsymbol{\sigma}) \leq 16\cdot 3\sqrt{\pi}\sqrt{\ln(cn)}\mathbb{E}[\mathcal{D}] \leq 86\sqrt{\ln(cn)}\mathbb{E}[\mathcal{D}].$$

■

4. A threshold algorithm for selecting the best box in small noise environments

In this section, we propose a new policy, IgnoreLarge, and give sufficient conditions under which IgnoreLarge's expected reward is at least a constant factor of the expected reward of a prophet who knows $x_1, \dots, x_n$. We will describe two versions of this policy. The first version works for all distributions (including MHR distributions). The second one is a slight modification that works only for MHR distributions, under a different condition on the instance. Without loss of generality, we will assume that boxes are ordered in increasing σ_i, that is, $\sigma_1 \leq \sigma_2 \leq \cdots \leq \sigma_n$.

- IgnoreLarge: Pick $\alpha \in [0,1]$ uniformly at random. Return $\arg\max_{1\leq i\leq \alpha n} y_i$.
- IgnoreLargeExp: Pick $\alpha \in [0,1]$ uniformly at random. Return $\arg\max_{1\leq i\leq n^\alpha} y_i$.

Note that α needs to be randomized since $\mathcal{D}$ is not known. Therefore, even if one knows that ignoring large noise boxes works, without a reference frame for the variance and expectation of $\mathcal{D}_{n:n}$ it is not clear whether a specific σ_i at hand is big or small, and hence which boxes should be ignored. Picking α at random is a necessary ingredient to circumvent this issue.

In Theorem 15 we present our guarantee for arbitrary distributions. Intuitively, if there is a universal constant c, e.g. $c = 0.01$, such that a c fraction of the boxes have bounded noise (specifically, $\sigma_{cn} \leq \frac{\mathbb{E}[\mathcal{D}_{cn:cn}]}{5\sqrt{2\ln n}}$), then our policy gives a constant approximation to the reward of a prophet.

Theorem 15 *For all $c \in [1/n, 1]$, for all distributions $\mathcal{D}$, all $n \geq 4$, and all $\boldsymbol{\sigma} \in \mathcal{S}_{(\mathcal{D},n,c)}$, we have $R_{\mathsf{IgnoreLarge}}(\mathcal{D},\boldsymbol{\sigma}) \geq \frac{c^2}{20}\cdot\mathbb{E}[\mathcal{D}_{n:n}]$.*

Proof of Theorem 15 Consider $\boldsymbol{\sigma} \in \mathcal{S}_{(\mathcal{D},n,c)}$ where, without loss of generality, we have $\sigma_1 \le \sigma_2 \le \cdots \le \sigma_n$. As $\boldsymbol{\sigma} \in \mathcal{S}_{(\mathcal{D},n,c)}$, we have $\sigma_{cn} \le \frac{\mathbb{E}[\mathcal{D}_{cn:cn}]}{5\sqrt{2\ln n}}$. Consider the event that $|\epsilon_i| \le \sigma_i\sqrt{2\ln n}$ for all $1 \le i \le cn$. For any such box i, we have

$$\begin{aligned}
\Pr\left[|\epsilon_i| \le \sigma_i\sqrt{2\ln n}\right] &= \Pr\left[|\mathcal{N}(0,\sigma_i^2)| \le \sigma_i\sqrt{2\ln n}\right] \\
&= 2\Phi\left(\sqrt{2\ln n}\right) - 1 \\
&\ge^{\text{(Lemma 4)}} 2\left(1 - \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\ln n}}\exp\left(-\frac{1}{2}\cdot 2\ln n\right)\right) - 1 \\
&= 1 - \frac{1}{n\sqrt{\pi\ln n}},
\end{aligned}$$

and therefore

$$\Pr\left[|\epsilon_i| \le \sigma_i\sqrt{2\ln n}, \forall i \in [1,cn]\right] \ge \left(1 - \frac{1}{n\sqrt{\pi\ln n}}\right)^{cn} \ge^{\text{(Bernoulli's inequality)}} 1 - \frac{c}{\sqrt{\pi\ln n}} \ge \frac{1}{2},$$

where the last inequality holds for all $n \ge 4 \ge e^{\frac{4c^2}{\pi}}$. Observe that, since $\sigma_i \le \frac{\mathbb{E}[\mathcal{D}_{cn:cn}]}{5\sqrt{2\ln n}}$ for all $i \in [1,cn]$, we can conclude that $\Pr[\max_{i\in[1,cn]}|\epsilon_i| \le \frac{1}{5}\cdot\mathbb{E}[\mathcal{D}_{cn:cn}]] \ge \frac{1}{2}$. Conditioned on this event we have $x_i - \frac{1}{5}\cdot\mathbb{E}[\mathcal{D}_{cn:cn}] \le y_i \le x_i + \frac{1}{5}\cdot\mathbb{E}[\mathcal{D}_{cn:cn}]$ for all $i \in [1,cn]$; therefore, for all $k \le cn$, we have $\max_{i\in[1,k]} y_i \ge \max_{i\in[1,k]} x_i - \frac{2}{5}\cdot\mathbb{E}[\mathcal{D}_{cn:cn}]$.

We analyze the performance of IgnoreLarge under this event. Recall that IgnoreLarge draws $\alpha \in [0,1]$ uniformly at random, and then outputs $\arg\max_{i\in[1,\alpha n]} y_i$. There are two cases for α:

Case $\alpha > c$: we will lower bound the expected reward of IgnoreLarge by 0.

Case $\alpha \le c$: IgnoreLarge is going to pick the box with the largest y_i among the first αn boxes. By our observation, IgnoreLarge's reward in this case is at least $\max_{i\in[1,\alpha n]} x_i - \frac{2}{5}\cdot\mathbb{E}[\mathcal{D}_{cn:cn}]$, and therefore the expected reward of IgnoreLarge in this case is at least

$$E[\mathcal{D}_{\alpha n:\alpha n}] - \frac{2}{5}\cdot\mathbb{E}[\mathcal{D}_{cn:cn}] \ge^{\text{(Lemma 5)}} \frac{\alpha}{c}\mathbb{E}[\mathcal{D}_{cn:cn}] - \frac{2}{5}\cdot\mathbb{E}[\mathcal{D}_{cn:cn}].$$

Therefore, conditioned on the event that $\max_{i\in[1,cn]}|\epsilon_i| \le \frac{1}{5}\cdot\mathbb{E}[\mathcal{D}_{cn:cn}]$, IgnoreLarge's expected reward is lower bounded by

$$\int_{\alpha=0}^{c}\frac{\alpha}{c}\mathbb{E}[\mathcal{D}_{cn:cn}] - \frac{2}{5}\cdot\mathbb{E}[\mathcal{D}_{cn:cn}]\,d\alpha = \frac{c}{10}\cdot\mathbb{E}[\mathcal{D}_{cn:cn}].$$

When this event does not occur, we lower bound IgnoreLarge's expected reward by 0. Combining everything together, IgnoreLarge's expected reward is

$$R_{\mathsf{IgnoreLarge}}(\mathcal{D},\boldsymbol{\sigma}) \ge \frac{1}{2}\cdot\frac{c}{10}\cdot\mathbb{E}[\mathcal{D}_{cn:cn}] \ge^{\text{(Lemma 5)}} \frac{c^2}{20}\cdot\mathbb{E}[\mathcal{D}_{n:n}].$$

■

We only used the fact that the noise distribution is Gaussian to lower bound $\Pr\left[|\epsilon_i| \le \sigma_i\sqrt{2\ln n}\right]$. Similar bounds on this quantity hold if the noise distribution is sub-Gaussian,[3] and therefore our

3. A distribution X is d sub-Gaussian if $\Pr[|X| \ge t] \le 2\exp(-t^2/d^2)$.

guarantees carry through to this variation of the problem as well, with an appropriate definition of the small noise regime.

For MHR distributions, the established condition (i.e. $\boldsymbol{\sigma} \in \mathcal{S}^{\text{MHR}}_{(\mathcal{D},n,c)}$) required by Theorem 16 suffices. We can prove strong positive results using a different sufficient condition: if for some universal constant c, such that n^c boxes have bounded noise (and specifically, σ_{n^c} at most $\frac{\mathbb{E}[\mathcal{D}_{n^c:n^c}]}{18\sqrt{2c\ln n}}$), then our policy gives a constant approximation to the reward of a prophet. The number of boxes is a lot smaller (n^c versus cn), but the bound on the noise is also smaller: $\frac{\mathbb{E}[\mathcal{D}_{n^c:n^c}]}{18\sqrt{2c\ln n}}$ versus $\frac{\mathbb{E}[\mathcal{D}_{cn:cn}]}{5\sqrt{2\ln n}}$. Therefore, the two theorems are incomparable.

Theorem 16 *For all $c \in (0,1]$, all MHR distributions $\mathcal{D}$, all $n \geq e^{\frac{4}{c\pi}}$, and all $\boldsymbol{\sigma} \in \mathcal{S}^{\text{MHR}}_{(\mathcal{D},n,c)}$, we have $R_{\mathsf{IgnoreLargeExp}}(\mathcal{D}, \boldsymbol{\sigma}) \geq \frac{c^2}{576} \cdot \mathbb{E}[\mathcal{D}_{n:n}]$.*

The proof of Theorem 16 follows a similar structure to the proof of Theorem 15 and is deferred to Appendix D. We note, however, that it uses the following useful lemma for MHR distribution, which bounds $\mathbb{E}[\mathcal{D}_{n^c:n^c}]$ as a function of $\mathbb{E}[\mathcal{D}_{n:n}]$.

Lemma 17 *For any MHR distribution $\mathcal{D}$, $n \geq 4$ and $a \geq 1$, $\mathbb{E}[\mathcal{D}_{n^a:n^a}] \leq 4a\, \mathbb{E}[\mathcal{D}_{n:n}]$.*

Proof of Lemma 17 Since $n \geq 4$ we have that $n^a \geq 4$ for all $a \geq 1$. Therefore,

$$\mathbb{E}[\mathcal{D}_{n^a:n^a}] \leq^{(\text{Lem. } 10)} 3\alpha^{(\mathcal{D})}_{n^a} \leq^{(\text{Lem. } 9)} 3a\, \alpha^{(\mathcal{D})}_{n} \leq^{(\text{Lem. } 10)} \frac{15a}{4}\, \mathbb{E}[\mathcal{D}_{n:n}] < 4a\, \mathbb{E}[\mathcal{D}_{n:n}].$$

■

5. Negative results for small noise environments

In this section, we show negative results for algorithms proposed by Bax et al. (2012), namely Naive (Section 5.1) and Linear_γ (Section 5.2). All missing proofs can be found in Appendix E.

5.1. Warm-up: Negative results for Naive

Theorem 18 *For every distribution $\mathcal{D}$, all $n \geq 46$, and all $c \leq \frac{n-6\ln(n)}{n}$, there exists $\boldsymbol{\sigma}^* = (\sigma^*_1, \ldots, \sigma^*_n)$ such that $\boldsymbol{\sigma}^* \in \mathcal{S}_{(\mathcal{D},n,c)}$, and*

$$R_{\mathsf{Naive}}(\mathcal{D}, \boldsymbol{\sigma}^*) \leq \frac{8\mathbb{E}[\mathcal{D}]}{\mathbb{E}[\mathcal{D}_{n:n}]} \cdot R_{\mathsf{Opt}_{\mathcal{D}}}(\mathcal{D}, \boldsymbol{\sigma}^*).$$

As an immediate consequence of Theorem 18, by picking a distribution $\mathcal{D}$ such that $\mathbb{E}[\mathcal{D}_{n:n}] \in \Theta(n\mathbb{E}[\mathcal{D}])$, we get that Naive only gives a (trivial) n approximation to the optimal policy.

Corollary 19 *For all $n \geq 46$ and $c \leq \frac{n-6\ln(n)}{n}$, there exists $\mathcal{D}$ and $\boldsymbol{\sigma}^* \in \mathcal{S}_{(\mathcal{D},n,c)}$ such that $R_{\mathsf{Opt}_{\mathcal{D}}}(\mathcal{D}, \boldsymbol{\sigma}^*) \in \Omega(n)\, R_{\mathsf{Naive}}(\mathcal{D}, \boldsymbol{\sigma}^*)$.*

Proof of Corollary 19 Consider the distribution $\mathcal{D}$ that takes the value 0 with probability $1 - 1/n$, and the value n with probability $1/n$. Then, $\mathbb{E}[\mathcal{D}] = 1$, and $\mathbb{E}[\mathcal{D}_{n:n}] = n \cdot \left(1 - \left(1 - \frac{1}{n}\right)^n\right) \geq n \cdot \left(1 - \frac{1}{e}\right)$. Applying Theorem 18 implies the corollary. ■

Our construction of $\boldsymbol{\sigma}^*$ works as follows, where $c_b = 6 \ln n$ and $\sigma_b = 6\beta_{n^2}^{(\mathcal{D}_{n:n})}\sqrt{\ln n}$:

$$\sigma_i^* = \begin{cases} 0 & i \in [1, n - c_b] \\ \sigma_b & i \in [n - c_b + 1, n] \end{cases}$$

We refer to the boxes with $\sigma_i^* = 0$ as "exact", while the boxes with $\sigma_i^* = \sigma_b$ as having "large noise." It is straightforward to confirm that $\boldsymbol{\sigma}^* \in \mathcal{S}_{(\mathcal{D},n,c)}$, for $c \leq \frac{n - 6\ln(n)}{n}$ (according to Definition 1).

Theorem 18 will be an immediate consequence of two facts. First, intuitively, a large noise box will have large ϵ_i with high probability, and therefore be selected by Naive, but its expected reward won't be much better than $4\mathbb{E}[\mathcal{D}]$ (Lemma 21). On the other hand, even the policy that selects the best exact box gets reward at least $\frac{1}{2}\mathbb{E}[\mathcal{D}_{n:n}]$ (Lemma 20).

Lemma 20 *For every distribution $\mathcal{D}$, for all $n \geq 46$, $R_{\mathsf{Opt}_{\mathcal{D}}}(\mathcal{D}, \boldsymbol{\sigma}^*) \geq \frac{1}{2}\mathbb{E}[\mathcal{D}_{n:n}]$.*

Proof of Lemma 20 The optimal policy is at least as good as the policy that selects the box with the largest y_i among the exact boxes. Since $x_i = y_i$ for these boxes, the reward of this policy is at least

$$\mathbb{E}[\mathcal{D}_{n-c_b:n-c_b}] \geq^{\text{(Lemma 5)}} \frac{n - c_b}{n} \cdot \mathbb{E}[\mathcal{D}_{n:n}] = \frac{n - 6\ln n}{n} \cdot \mathbb{E}[\mathcal{D}_{n:n}] \geq^{(n \geq 46)} \frac{1}{2}\mathbb{E}[\mathcal{D}_{n:n}].$$

■

Lemma 21 *For every distribution $\mathcal{D}$, for all $n \geq 22$, we have that $R_{\mathsf{Naive}}(\mathcal{D}, \boldsymbol{\sigma}^*) \leq 4\mathbb{E}[\mathcal{D}]$.*

On a high level, our proof works as follows. Consider the event $\mathcal{E}^*$ that $X_i \leq \beta_{n^2}^{(\mathcal{D}_{n:n})}$ for all boxes i. We prove that conditioned on $\mathcal{E}^*$, Naive gets an expected reward of at most $3\mathbb{E}[\mathcal{D}]$. On the other hand, when $\mathcal{E}^*$ does not occur, even if Naive performs as well as taking $\mathcal{D}_{n:n} = \max_i X_i$, the contribution to the final expected reward is also upper bounded by $\mathbb{E}[\mathcal{D}]$. The second fact can be shown directly from the definition of $\beta_{n^2}^{(\mathcal{D}_{n:n})}$. For the first fact, we first show that with high probability ϵ_i is not too small for some large box i (Lemma 22); conditioned on $\mathcal{E}^*$ and this event, this implies that Naive picks a large noise box. It is also true that with high probability ϵ_i is not too big, for any large noise box i (Lemma 23). Additionally conditioning on ϵ_i being not too big for every large noise box, we have that both the noise and the reward are not too big (and there is a box with large noise). We can then upper bound the reward of Naive by the reward of a "clairvoyant" policy which knows $\mathcal{D}$, but is required to pick a large noise box; for this step, we need a technical lemma (Lemma 24) that will also be useful in our lower bound for linear policies. In all other events, we upper bound Naive by $\max_i X_i$.

Lemma 22 *With probability at least $1 - \frac{1}{n^3}$, $\epsilon_i > \beta_{n^2}^{(\mathcal{D}_{n:n})}$ for at least one large noise box i.*

Lemma 23 *For any large noise box i, we have $\Pr\left[\epsilon_i \leq 12\beta_{n^2}^{(\mathcal{D}_{n:n})} \ln n\right] \geq 1 - \frac{1}{n^2}$.*

Lemma 24 *For any non-negative and bounded random variable Z supported on $[0, V]$ and any $\sigma > 2V$, we have that $\mathbb{E}[Z \mid Z + \mathcal{N}(0, \sigma^2) = y] \leq 2\mathbb{E}[Z]$ for all $y \leq \frac{\sigma^2}{2V}$.*

The proofs of Lemmas 22 to 24 can be found in Appendix E.1. Next, we prove Lemma 21.
Proof of Lemma 21 We define the following events.

- $\mathcal{E}_1$ be the event that $\epsilon_j \leq 12\beta_{n^2}^{(\mathcal{D}_{n:n})} \ln n$ for all large noise boxes j.
- $\mathcal{E}_1'$ be the event that $Y_j \leq 18\beta_{n^2}^{(\mathcal{D}_{n:n})} \ln n$ for all large noise boxes j.
- $\mathcal{E}_2$ be the event that $\epsilon_j > \beta_{n^2}^{(\mathcal{D}_{n:n})}$ for at least one large noise box j.
- $\mathcal{E}_2'$ be the event that $Y_j > \beta_{n^2}^{(\mathcal{D}_{n:n})}$ for at least one large noise box j.

Recall that $\mathcal{E}^*$ is the event that $X_i \leq \beta_{n^2}^{(\mathcal{D}_{n:n})}$ for all $i \in [n]$.

We first explore the relationship between these events. First, notice that if $X_i \leq \beta_{n^2}^{(\mathcal{D}_{n:n})}$ and $\epsilon_i \leq 12\beta_{n^2}^{(\mathcal{D}_{n:n})} \ln n$, we have that

$$Y_i = X_i + \epsilon_i \leq \beta_{n^2}^{(\mathcal{D}_{n:n})} + 12\beta_{n^2}^{(\mathcal{D}_{n:n})} \ln n \leq 18\beta_{n^2}^{(\mathcal{D}_{n:n})} \ln n.$$

Therefore, $\mathcal{E}_1 \cap \mathcal{E}^* \subseteq \mathcal{E}_1' \cap \mathcal{E}^*$. Since $X_i \geq 0$ for all i, $\mathcal{E}_2'$ occurs every time $\mathcal{E}_2$ occurs, i.e. $\mathcal{E}_2 \subseteq \mathcal{E}_2'$, and thus $\mathcal{E}_2 \cap \mathcal{E}^* \subseteq \mathcal{E}_2' \cap \mathcal{E}^*$. Therefore, $\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}^* \subseteq \mathcal{E}_1' \cap \mathcal{E}_2' \cap \mathcal{E}^*$, or $\overline{\mathcal{E}_1 \cap \mathcal{E}_2} \cap \mathcal{E}^* \supseteq \overline{\mathcal{E}_1' \cap \mathcal{E}_2'} \cap \mathcal{E}^*$.

First, we will bound $\mathbb{E}[\max X_i \mid \overline{\mathcal{E}_1' \cap \mathcal{E}_2'} \cap \mathcal{E}^*] \cdot \Pr[\overline{\mathcal{E}_1' \cap \mathcal{E}_2'} \mid \mathcal{E}^*]$, which is an upper bound on the contribution of outcomes in $\overline{\mathcal{E}_1' \cap \mathcal{E}_2'} \cap \mathcal{E}^*$ to the overall expected reward of Naive. Since the contribution of an event A to the expectation of a random variable ($\mathbb{E}[X|A]\Pr[A]$) is smaller than the contribution of an event B to the expectation if $A \subseteq B$, we have

$$\mathbb{E}[\max_i X_i \mid \overline{\mathcal{E}_1' \cap \mathcal{E}_2'} \cap \mathcal{E}^*] \cdot \Pr[\overline{\mathcal{E}_1' \cap \mathcal{E}_2'} \mid \mathcal{E}^*] \leq \mathbb{E}[\max_i X_i \mid \overline{\mathcal{E}_1 \cap \mathcal{E}_2} \cap \mathcal{E}^*] \cdot \Pr[\overline{\mathcal{E}_1 \cap \mathcal{E}_2} \mid \mathcal{E}^*].$$

By Lemma 23, $\Pr[\mathcal{E}_1] \geq \left(1 - \frac{1}{n^2}\right)^{c_b} \geq 1 - \frac{6\ln n}{n^2}$. By Lemma 22, $\Pr[\mathcal{E}_2] \geq 1 - \frac{1}{n^3}$. Therefore, $\Pr[\mathcal{E}_1 \cap \mathcal{E}_2] \geq \Pr[\mathcal{E}_1] + \Pr[\mathcal{E}_2] - 1 \geq 1 - \frac{6\ln n}{n^2} + 1 - \frac{1}{n^3} - 1 \geq 1 - \frac{7\ln n}{n^2}$. Observe that, $\mathcal{E}_1$ and $\mathcal{E}_2$ are independent of the X_is, while $\mathcal{E}^*$ only dependent on X_is. Therefore, $\mathcal{E}_1 \cap \mathcal{E}_2$ and $\mathcal{E}^*$ are independent, and hence $\Pr[\mathcal{E}_1 \cap \mathcal{E}_2 \mid \mathcal{E}^*] = \Pr[\mathcal{E}_1 \cap \mathcal{E}_2] \geq 1 - \frac{7\ln n}{n^2}$, or $\Pr[\overline{\mathcal{E}_1 \cap \mathcal{E}_2} \mid \mathcal{E}^*] \leq \frac{7\ln n}{n^2}$. Additionally, $\mathbb{E}[\max_i X_i \mid \overline{\mathcal{E}_1 \cap \mathcal{E}_2} \cap \mathcal{E}^*] = \mathbb{E}[\max_i X_i \mid \mathcal{E}^*]$, as $\mathcal{E}_1$ and $\mathcal{E}_2$ are events regarding ϵ_is and therefore is independent of X_i. Furthermore, $\mathbb{E}[\max_i X_i \mid \mathcal{E}^*] = \mathbb{E}[\max_i X_i \mid X_i \leq \beta_{n^2}^{(\mathcal{D}_{n:n})}] \leq \mathbb{E}[\max_i X_i] = \mathbb{E}[\mathcal{D}_{n:n}]$. Putting everything together, we have

$$\begin{aligned}
\mathbb{E}[\max_i X_i \mid \overline{\mathcal{E}_1' \cap \mathcal{E}_2'} \cap \mathcal{E}^*] \cdot \Pr[\overline{\mathcal{E}_1' \cap \mathcal{E}_2'} \mid \mathcal{E}^*] &\leq \mathbb{E}[\max_i X_i \mid \overline{\mathcal{E}_1 \cap \mathcal{E}_2} \cap \mathcal{E}^*] \cdot \Pr[\overline{\mathcal{E}_1 \cap \mathcal{E}_2} \mid \mathcal{E}^*] \\
&\leq \mathbb{E}[\mathcal{D}_{n:n}] \cdot \frac{7\ln n}{n^2} \\
&\leq^{(\text{Lemma 5})} \frac{7\ln n}{n^2} \cdot n \cdot \mathbb{E}[\mathcal{D}] \\
&\leq^{(n \geq 22)} \mathbb{E}[\mathcal{D}].
\end{aligned}$$

Second, we will upper bound the contribution of outcomes in $\mathcal{E}_1' \cap \mathcal{E}_2' \cap \mathcal{E}^*$ to the expected reward of Naive. Note that in such outcomes, Naive must choose a large noise box, by the definition of $\mathcal{E}_2'$

($Y_j > \beta_{n^2}^{(\mathcal{D}_{n:n})}$ for some large noise box j) and $\mathcal{E}^*$ ($X_i \leq \beta_{n^2}^{(\mathcal{D}_{n:n})}$ for all i, and therefore the exact boxes). Therefore, in such an outcome, the reward of Naive is at most the reward of an optimal policy which also knows $\mathcal{D}$, but is conditioned to pick a large noise box. When selecting box i such a policy makes expected reward $\mathbb{E}[X_i \mid Y_i = y_i, \mathcal{E}^*, \mathcal{E}'_1, \mathcal{E}'_2] = \mathbb{E}[X_i \mid Y_i = y_i, \mathcal{E}^*]$, where the equality holds since X_i is independent of Y_j, for $j \neq i$, and $\mathcal{E}'_1 \cap \mathcal{E}'_2$ have less information about Y_i than $\{Y_i = y_i\}$. Let $R_i(y_i) = \mathbb{E}[X_i \mid Y_i = y_i, \mathcal{E}^*]$. The reward of an optimal policy which knows $\mathcal{D}$ and is conditioned to pick a large noise box is then $\mathbb{E}_{\boldsymbol{y}}\left[\max_{i\in[n-c_b+1,n]} R_i(y_i) \mid \mathcal{E}'_1 \cap \mathcal{E}'_2 \cap \mathcal{E}^*\right]$. We prove that $R_i(y_i) \leq 2\mathbb{E}[\mathcal{D}]$ for all y_i consistent with $\mathcal{E}'_1 \cap \mathcal{E}'_2 \cap \mathcal{E}^*$, which in turn implies an upper bound of $2\mathbb{E}[\mathcal{D}]$ for the expected reward of Naive conditioned on in $\mathcal{E}'_1 \cap \mathcal{E}'_2 \cap \mathcal{E}^*$.

Consider any large noise box i. Let $\overline{X}_i = X_i \mid X_i \leq \beta_{n^2}^{(\mathcal{D}_{n:n})}$.[4] Then, conditioned on $\mathcal{E}'_1 \cap \mathcal{E}'_2 \cap \mathcal{E}^*$, for any realization of $\boldsymbol{y}$, we note that $R_i(y_i) = \mathbb{E}[X_i \mid Y_i = y_i, \mathcal{E}^*] = \mathbb{E}[\overline{X}_i \mid \overline{X}_i + \mathcal{N}(0,\sigma_i^2) = y_i]$. Furthermore, as y_i is a realization conditioned on $\mathcal{E}'_1 \cap \mathcal{E}'_2 \cap \mathcal{E}^*$, we have $y_i \leq 18\beta_{n^2}^{(\mathcal{D}_{n:n})}\ln n$. Using Lemma 24 for $V = \beta_{n^2}^{(\mathcal{D}_{n:n})}$ and $\sigma = \sigma_b = 6\beta_{n^2}^{(\mathcal{D}_{n:n})}\sqrt{\ln n}$, we have $\mathbb{E}[\overline{X}_i \mid \overline{X}_i + \mathcal{N}(0,\sigma_i^2) = y_i] \leq 2\mathbb{E}[\overline{X}_i] \leq 2\mathbb{E}[X_i] = 2\mathbb{E}[\mathcal{D}]$.

Overall, conditioned on $\mathcal{E}^*$, if $\mathcal{E}'_1 \cap \mathcal{E}'_2$ occurs, Naive's expected reward is at most $2\mathbb{E}[\mathcal{D}]$; otherwise, the contribution to the expected reward is at most $\mathbb{E}[\mathcal{D}]$. Thus, the reward of Naive conditioned on $\mathcal{E}^*$ is at most

$$\begin{aligned}\Pr[\mathcal{E}'_1 \cap \mathcal{E}'_2 \mid \mathcal{E}^*] \cdot 2\mathbb{E}[\mathcal{D}] + \mathbb{E}[\max X_i \mid \overline{\mathcal{E}'_1 \cap \mathcal{E}'_2} \cap \mathcal{E}^*] \cdot \Pr[\overline{\mathcal{E}'_1 \cap \mathcal{E}'_2} \mid \mathcal{E}^*]\\ \leq 2\mathbb{E}[\mathcal{D}] + \mathbb{E}[\mathcal{D}] = 3\mathbb{E}[\mathcal{D}].\end{aligned}$$

Finally, conditioned on $\mathcal{E}^*$ not happening, the best Naive can do is $\mathcal{D}_{n:n} = \max_i X_i$, whose expected reward is $\mathbb{E}[\mathcal{D}_{n:n} \mid \overline{\mathcal{E}^*}]$. Therefore:

$$\begin{aligned}R_{\mathsf{Naive}}(\mathcal{D}, \boldsymbol{\sigma}^*) &\leq 3\mathbb{E}[\mathcal{D}] \cdot \Pr[\mathcal{E}^*] + \mathbb{E}[\mathcal{D}_{n:n} \mid \overline{\mathcal{E}^*}] \cdot \Pr[\overline{\mathcal{E}^*}]\\ &\leq 3\mathbb{E}[\mathcal{D}] + \mathbb{E}[\mathcal{D}_{n:n} \mid \mathcal{D}_{n:n} \geq \beta_{n^2}^{(\mathcal{D}_{n:n})}] \cdot \Pr[\mathcal{D}_{n:n} \geq \beta_{n^2}^{(\mathcal{D}_{n:n})}]\\ &\leq^{\text{(Definition 7)}} 3\mathbb{E}[\mathcal{D}] + \frac{\mathbb{E}[\mathcal{D}_{n:n}]}{n^2}\\ &\leq^{\text{(Lemma 5)}} 3\mathbb{E}[\mathcal{D}] + \frac{n \cdot \mathbb{E}[\mathcal{D}]}{n^2} \leq 4\mathbb{E}[\mathcal{D}].\end{aligned}$$

■

Given all our lemmas thus far, the proof of Theorem 18 is straightforward.

Proof of Theorem 18 The theorem is implied by Lemmas 20 and 21. ■

5.2. Negative results for linear policies

In this section, we give our negative results for linear policies. Recall that a linear policy parameterized by $\gamma : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ selects the box which maximizes $y_i - \gamma(\boldsymbol{\sigma}, \boldsymbol{y}) \cdot \sigma_i$.

Theorem 25 *For every MHR distribution $\mathcal{D}$, for all $n \geq n_0$, for some constant n_0, there exists $\boldsymbol{\sigma}^*$, such that $\boldsymbol{\sigma}^* \in \mathcal{S}^{\mathrm{MHR}}_{(\mathcal{D},n,1/5626)}$, and for every function $\gamma : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, we have*

$$R_{\mathsf{Linear}_\gamma}(\mathcal{D}, \boldsymbol{\sigma}^*) \in O\left(\frac{\mathbb{E}[\mathcal{D}]}{\mathbb{E}[\mathcal{D}_{n:n}]}\right) R_{\mathsf{Opt}_\mathcal{D}}(\mathcal{D}, \boldsymbol{\sigma}^*).$$

4. Equivalently, we can think of sampling from $\overline{X}_i$ by sampling from X_i, until $X_i \leq \beta_{n^2}^{(\mathcal{D}_{n:n})}$.

An immediate corollary is that linear policies give, in the worst case, a logarithmic approximation, even for MHR distributions, by considering $\mathcal{D}$ to be the exponential distribution with parameter $\lambda = 1$, for which $\mathbb{E}[\mathcal{D}_{n:n}] = \sum_{i=1}^{n} \frac{1}{i} \geq \ln n$. Note also that $\mathbb{E}[\mathcal{D}_{n:n}] \leq \ln n + 1$ for all MHR random variables (Lemma 36 in Appendix E.2), so the exponential distribution minimizes the ratio in Theorem 25 (up to constants).

Corollary 26 *There exists $\mathcal{D}$, such that for all $n \geq n_0$, for some constant n_0, there exists $\boldsymbol{\sigma}^* \in \mathcal{S}^{\text{MHR}}_{(\mathcal{D},n,1/5626)}$ such that $R_{\mathsf{Opt}_{\mathcal{D}}}(\mathcal{D}, \boldsymbol{\sigma}^*) \in \Omega(\ln(n)) \cdot R_{\mathsf{Linear}_\gamma}(\mathcal{D}, \boldsymbol{\sigma}^*)$.*

Our construction of $\boldsymbol{\sigma}^*$ works as follows. It contains one box such that $\sigma^* = 0$, a small number of boxes with some small noise σ_s, and the remaining boxes have large noise σ_b:

$$\sigma_i^* = \begin{cases} 0 & i = 1 \\ \sigma_s & i \in [2, c_s + 1] \\ \sigma_b & i \in [c_s + 2, n] \end{cases}$$

where $c_s = n^{1/5626}$, $\sigma_s = \frac{37}{9\sqrt{2}} \frac{\mathbb{E}[\mathcal{D}_{c_s:c_s}]}{\sqrt{\ln n}}$, and $\sigma_b = 6\alpha^{(\mathcal{D}_{n-c_s:n-c_s})}_{n^{1/10000}} \sqrt{\ln n}$. We refer to the first box as the "exact box," the boxes with $\sigma_i^* = \sigma_s$ as "small noise" boxes, and the rest as "large noise" boxes. One can easily confirm that $\boldsymbol{\sigma}^* \in \mathcal{S}^{\text{MHR}}_{(\mathcal{D},n,1/5626)}$.

Next, we lower bound the expected reward of the optimal policy.

Lemma 27 *For every MHR distribution $\mathcal{D}$, all $n \geq n_0$ for some constant n_0, $R_{\mathsf{Opt}}(\mathcal{D}, \boldsymbol{\sigma}^*) \in \Omega\left(\mathbb{E}[\mathcal{D}_{n:n}]\right)$.*

Proof The optimal policy is at least as good as the policy that picks the box with the largest y_i among the small noise boxes. Consider the event that $|\epsilon_i| \leq \frac{\sqrt{2}}{75}\sigma_s\sqrt{\ln n}$ for all small noise boxes i:

$$\begin{aligned}
&\Pr\left[\max_{i\in[2,c_s+1]} |\epsilon_i| \leq \frac{\sqrt{2}}{75}\sigma_s\sqrt{\ln n}\right] \\
&\quad= \Pr\left[|\mathcal{N}(0,\sigma_s^2)| \leq \frac{\sqrt{2}}{75}\sigma_s\sqrt{\ln n}\right]^{c_s} \\
&\quad= \left(2\Phi\left(\frac{\sqrt{2\ln n}}{75}\right) - 1\right)^{c_s} \\
&\quad\geq^{(\text{Lemma 4})} \left(2\left(1 - \frac{1}{\sqrt{2\pi}}\frac{75}{\sqrt{2\ln n}}\exp\left(-\frac{1}{2}\cdot\frac{2}{5625}\ln n\right)\right) - 1\right)^{c_s} \\
&\quad= \left(1 - \frac{75}{\sqrt{\pi\ln n}} n^{-1/5625}\right)^{n^{1/5626}} \\
&\quad\geq^{(\text{Bernoulli's inequality})} 1 - \frac{75}{\sqrt{\pi \ln n}} n^{1/5626 - 1/5625} \\
&\quad\geq 1 - \frac{1}{\ln n}.
\end{aligned}$$

When this event occurs, the reward from picking a small noise box i is at least $y_i - \frac{\sqrt{2}}{75}\sigma_s\sqrt{\ln n}$, and therefore the overall reward of picking from small noise boxes is at least $\max_{i=2,\dots,c_s+1} x_i - \frac{2\sqrt{2}}{75}\sigma_s\sqrt{\ln n}$. Noting that the noise and reward are independent random variables, we have:

$$\begin{aligned}
R_{\mathsf{Opt}}(\mathcal{D}, \boldsymbol{\sigma}^*) &\geq \left(1 - \frac{1}{\ln n}\right) \cdot \left(\mathbb{E}\left[\max_{i \in [2, c_s+1]} X_i\right] - \frac{2\sqrt{2}}{75}\sqrt{\ln n} \cdot \sigma_s\right) \\
&= \left(1 - \frac{1}{\ln n}\right)\left(\mathbb{E}[\mathcal{D}_{c_s:c_s}] - \frac{2\sqrt{2}}{75}\sqrt{\ln n} \cdot \frac{5}{\sqrt{2}}\frac{\mathbb{E}[\mathcal{D}_{c_s:c_s}]}{\sqrt{\ln n}}\right) \\
&\geq \left(1 - \frac{1}{\ln n}\right) \cdot \frac{1}{75}\mathbb{E}[\mathcal{D}_{c_s:c_s}]. \\
&\geq^{\text{(Lemma 17)}} \left(1 - \frac{1}{\ln n}\right)\frac{1}{75} \cdot \frac{1}{4 \cdot 5626}\mathbb{E}[\mathcal{D}_{n:n}] \\
&\geq^{(n \geq e^{606})} \frac{1}{2\,000\,000}\mathbb{E}[\mathcal{D}_{n:n}].
\end{aligned}$$

■

Our next (and final) task is to upper bound the expected reward of Linear.

Lemma 28 *For every MHR distribution $\mathcal{D}$, for all $n \geq n_0$, for some constant n_0, and for all γ, it holds that $R_{\mathsf{Linear}_\gamma}(\mathcal{D}, \boldsymbol{\sigma}^*) \leq 8\mathbb{E}[\mathcal{D}]$.*

The proof structure of Lemma 28 is similar to that of Lemma 21. We first prove that conditioned on an event $\mathcal{E}^*$, Linear_γ's expected reward is upper bounded, while the contribution to the reward of other events is negligible, even if Linear_γ performs as well as taking $\max_i X_i$. Here, $\mathcal{E}^*$ is the event that $X_i \leq \alpha_{n^{1/10000}}^{(\mathcal{D}_{c_s:c_s})}$ *for all small noise boxes i, and* $X_j \leq \alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})}$ *for all remaining boxes j.*

Lemma 29 *For every MHR distribution $\mathcal{D}$, for all $n \geq n_0$, for some constant n_0, and for all γ, the expected reward of a policy* Linear_γ *conditioned on the event $\mathcal{E}^*$ is at most $7\mathbb{E}[\mathcal{D}]$.*

To prove Lemma 29, we first consider a slightly different family of policies. Let $\mathsf{LinearFixed}_c$ be the policy that chooses the box with the largest $y_i - c\sigma_i$, where c is a constant independent of $\boldsymbol{y}$ and $\boldsymbol{\sigma}$. We show that with high probability, all $\mathsf{LinearFixed}$ policies make poor choices. We can use this fact to get bounds on the performance of Linear_γ (conditioned on certain events), since, fixing $\boldsymbol{y}$ and $\boldsymbol{\sigma}$, Linear_γ is only as good as the best $\mathsf{LinearFixed}$ policy. We consider two cases on c: $c > \theta^*$ and $c \leq \theta^*$, where $\theta^* = \sqrt{\frac{\ln n}{2}}$.

We defer the proof of Lemma 29, and how it can be used to prove Lemma 28, to Appendix E.2. Theorem 25 is a direct corollary of Lemmas 27 and 28.

Acknowledgments

Trung Dang is supported in part by NSF award CCF-2217069. Alexandros Psomas is supported in part by an NSF CAREER award CCF-2144208, and a research award from the Herbert Simon Family Foundation.

References

Amine Allouah and Omar Besbes. Prior-independent optimal auctions. *Management Science*, 66 (10):4417–4432, 2020.

Eric Balkanski, Vasilis Gkatzelis, and Xizhi Tan. Mechanism design with predictions: An annotated reading list. *ACM SIGecom Exchanges*, 21(1):54–57, 2024.

Richard E. Barlow and Frank Proschan. *Mathematical Theory of Reliability*. Society for Industrial and Applied Mathematics, 1996. doi: 10.1137/1.9781611971194. URL https://epubs.siam.org/doi/abs/10.1137/1.9781611971194.

Eric Bax, Anand Kuratti, Preston Mcafee, and Julian Romero. Comparing predicted prices in auctions for online advertising. *International Journal of Industrial Organization*, 30(1):80–88, 2012.

Dimitris Bertsimas, David B Brown, and Constantine Caramanis. Theory and applications of robust optimization. *SIAM review*, 53(3):464–501, 2011.

Robin Bowers, Marius Garbea, Emmanouil Pountourakis, and Samuel Taggart. Polynomial-time approximation schemes via utility alignment: Unit-demand pricing and more. In *2025 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2025.

Mark Braverman, Jieming Mao, and S Matthew Weinberg. Parallel algorithms for select and partition with noisy comparisons. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 851–862, 2016.

Mark Braverman, Jieming Mao, and Yuval Peres. Sorted top-k in rounds. In *Conference on Learning Theory*, pages 342–382. PMLR, 2019.

Yang Cai and Constantinos Daskalakis. Extreme-value theorems for optimal multidimensional pricing. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pages 522–531, 2011. doi: 10.1109/FOCS.2011.76.

Vincent Cohen-Addad, Frederik Mallmann-Trenn, and Claire Mathieu. Instance-optimality in the noisy value-and comparison-model* accept, accept, strong accept: Which papers get in? In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2124–2143. SIAM, 2020.

Richard Cole and Tim Roughgarden. The sample complexity of revenue maximization. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 243–252, 2014.

Constantinos Daskalakis and Seth Matthew Weinberg. Symmetries and optimal multi-dimensional mechanism design. In *Proceedings of the 13th ACM conference on Electronic commerce*, pages 370–387, 2012.

Peerapong Dhangwatnotai, Tim Roughgarden, and Qiqi Yan. Revenue maximization with a single sample. In *Proceedings of the 11th ACM conference on Electronic commerce*, pages 129–138, 2010.

Uriel Feige, Prabhakar Raghavan, David Peleg, and Eli Upfal. Computing with noisy information. *SIAM Journal on Computing*, 23(5):1001–1018, 1994.

Yiannis Giannakopoulos, Diogo Poças, and Keyu Zhu. Optimal pricing for mhr and λ-regular distributions. *ACM Transactions on Economics and Computation (TEAC)*, 9(1):1–28, 2021.

Robert David Gordon. Values of mills' ratio of area to bounding ordinate and of the normal probability integral for large values of the argument. *Annals of Mathematical Statistics*, 12:364–366, 1941.

Chenghao Guo, Zhiyi Huang, and Xinzhi Zhang. Settling the sample complexity of single-parameter revenue maximization. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 662–673, 2019.

Wenshuo Guo, Michael Jordan, and Emmanouil Zampetakis. Robust learning of optimal auctions. *Advances in Neural Information Processing Systems*, 34:21273–21284, 2021.

Jason D Hartline and Tim Roughgarden. Simple versus optimal mechanisms. In *Proceedings of the 10th ACM conference on Electronic commerce*, pages 225–234, 2009.

Zhiyi Huang, Yishay Mansour, and Tim Roughgarden. Making the most of your samples. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 45–60, 2015.

Norman Lloyd. Johnson, Samuel. Kotz, and N. Balakrishnan. *Continuous univariate distributions / Norman L. Johnson, Samuel Kotz, N. Balakrishnan.* Wiley series in probability and mathematical statistics. Wiley & Sons, New York, 2nd ed. edition, 1994. ISBN 0471584959.

Ali Khodabakhsh, Emmanouil Pountourakis, and Samuel Taggart. Simple delegated choice. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 569–590. SIAM, 2024.

Fred C Leone, Lloyd S Nelson, and RB Nottingham. The folded normal distribution. *Technometrics*, 3(4):543–550, 1961.

Siqi Liu and Christos-Alexandros Psomas. On the competition complexity of dynamic mechanism design. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2008–2025. SIAM, 2018.

Mohammad Mahdian, Jieming Mao, and Kangning Wang. Regret minimization with noisy observations. In *Proceedings of the 24th ACM Conference on Economics and Computation*, EC '23, page 963, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9798400701047. doi: 10.1145/3580507.3597789. URL https://doi.org/10.1145/3580507.3597789.

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

Richard H Thaler. Anomalies: The winner's curse. *Journal of economic perspectives*, 2(1):191–202, 1988.

Appendix A. A technical lemma

The following technical lemma will be useful throughout this appendix.

Lemma 30 *For a random variable $Y = X + \epsilon$, where $\epsilon \sim \mathcal{N}(0, \sigma^2)$, it holds that $\mathbb{E}[X \mid Y = y]$ is monotone non-decreasing in y.*

Proof of Lemma 30 Let $A(y) = \int_0^\infty x \cdot f(x) \cdot f_{\mathcal{N}}(y-x)\, dx$ and $B(y) = \int_0^\infty f(x) \cdot f_{\mathcal{N}}(y-x)\, dx$, then $\mathbb{E}[X \mid Y = y] = \frac{A(y)}{B(y)}$. We first compute the derivative of $f_{\mathcal{N}}(y-x)$:

$$\begin{aligned}
\frac{df_{\mathcal{N}}(y-x)}{dy} &= \frac{d}{dy}\left(\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2} \cdot \left(\frac{y-x}{\sigma}\right)^2\right)\right) \\
&= \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2} \cdot \left(\frac{y-x}{\sigma}\right)^2\right) \cdot \frac{x-y}{\sigma^2} \\
&= f_{\mathcal{N}}(y-x) \cdot \frac{x-y}{\sigma^2}.
\end{aligned}$$

Let $C(y) = \int_0^\infty x^2 \cdot f(x) \cdot f_{\mathcal{N}}(y-x)\, dx$. The derivative for $A(y)$ is

$$\begin{aligned}
\frac{dA(y)}{dy} &= \frac{d}{dy}\left(\int_0^\infty x \cdot f(x) \cdot f_{\mathcal{N}}(y-x)\, dx\right) \\
&= \int_0^\infty x \cdot f(x) \cdot f_{\mathcal{N}}(y-x) \cdot \frac{x-y}{\sigma^2}\, dx \\
&= \frac{1}{\sigma^2}\left(C(y) - y \cdot A(y)\right).
\end{aligned}$$

The derivative for $B(y)$ is

$$\begin{aligned}
\frac{dB(y)}{dy} &= \frac{d}{dy}\left(\int_0^\infty f(x) \cdot f_{\mathcal{N}}(y-x)\, dx\right) \\
&= \int_0^\infty f(x) \cdot f_{\mathcal{N}}(y-x) \cdot \frac{x-y}{\sigma^2}\, dx \\
&= \frac{1}{\sigma^2}\left(A(y) - y \cdot B(y)\right).
\end{aligned}$$

Finally, the derivative for $\mathbb{E}[X \mid Y = y]$ is

$$\begin{aligned}
\frac{d}{dy}\mathbb{E}[X \mid Y = y] &= \frac{d}{dy}\frac{A(y)}{B(y)} \\
&= \frac{\frac{dA(y)}{dy} \cdot B(y) - \frac{dB(y)}{dy} \cdot A(y)}{B(y)^2} \\
&= \frac{\left(\frac{1}{\sigma^2}\left(C(y) - yA(y)\right)\right) \cdot B(y) - \left(\frac{1}{\sigma^2}\left(A(y) - yB(y)\right)\right) \cdot A(y)}{B(y)^2} \\
&= \frac{B(y)C(y) - yA(y)B(y) - A(y)^2 + yA(y)B(y)}{(\sigma B(y))^2}
\end{aligned}$$

$$= \frac{B(y)C(y) - A(y)^2}{(\sigma B(y))^2}.$$

Since $(x \cdot f(x) \cdot f_{\mathcal{N}}(y-x))^2 = (f(x) \cdot f_{\mathcal{N}}(y-x)) \cdot (x^2 \cdot f(x) \cdot f_{\mathcal{N}}(y-x))$, the Cauchy-Schwarz inequality implies that $B(y)C(y) \geq A(y)^2$. Therefore $\frac{d}{dy}\mathbb{E}[X \mid Y = y] = \frac{B(y)C(y)-A(y)^2}{(\sigma B(y))^2} \geq 0$. ■

Appendix B. Deferred proofs from Section 2

Proof of Lemma 5 It is sufficient to prove that $\frac{\mathbb{E}[\mathcal{D}_{k:k}]}{k} \geq \frac{\mathbb{E}[\mathcal{D}_{k+1:k+1}]}{k+1}$ for all integers $k \geq 1$, or $\frac{\int_0^\infty 1-F(x)^k}{k} dx \geq \frac{\int_0^\infty 1-F(x)^{k+1}}{k+1} dx$. For all $t \in [0,1]$, we have

$$\sum_{i=0}^{k-1} t^i \geq kt^k \Leftrightarrow (1-t)\sum_{i=0}^{k-1} t^i \geq k(1-t)t^k \Leftrightarrow 1 - t^k \geq k(t^k - t^{k+1})$$
$$\Leftrightarrow k + 1 - (k+1)t^k \geq k - kt^{k+1} \Leftrightarrow \frac{1-t^k}{k} \geq \frac{1-t^{k+1}}{k+1}.$$

Letting $t = F(x)$ and integrating both sides, we get our desired result. ■

Proof of Lemma 10 We will heavily use the fact that order statistics of MHR distributions are also MHR (Theorem 5.5 on page 39 of Barlow and Proschan (1996)):

Lemma 31 (Barlow and Proschan (1996)) *For any MHR random variable X and any integers $1 \leq k \leq n$, $X_{k:n}$ is MHR.*

Define $\zeta_p^{(\mathcal{D})} = \inf\{x \mid F(x) \geq p\}$ as the p-th quantile of $\mathcal{D}$. We also use the following result from Barlow and Proschan (1996) (Theorem 4.6 on page 30):

Lemma 32 (Barlow and Proschan (1996)) *Let X be MHR with mean μ. If $p \leq 1 - \frac{1}{e}$, then*

$$\ln\left(\frac{1}{1-p}\right)\mu \leq \zeta_p^{(X)} \leq \frac{\ln\left(\frac{1}{1-p}\right)}{p}\mu.$$

For the lower bound, we first observe that $\Pr[\mathcal{D}_{n:n} \leq \alpha_n^{(\mathcal{D})}] = \Pr[\mathcal{D} \leq \alpha_n^{(\mathcal{D})}]^n = (1 - \frac{1}{n})^n$, where with $n \geq 4$ we get $\frac{81}{256} \leq (1 - \frac{1}{n})^n \leq \frac{1}{e}$. Therefore, $\zeta_{81/256}^{(\mathcal{D}_{n:n})} \leq \alpha_n^{(\mathcal{D})} \leq \zeta_{1/e}^{(\mathcal{D}_{n:n})}$. From Lemma 31, we know that $\mathcal{D}_{n:n}$ is also MHR. Since $\frac{81}{256} \leq 1/e \leq 1 - 1/e$, we can invoke Lemma 32 on $\zeta_{81/256}^{(\mathcal{D}_{n:n})}$ and $\zeta_{1/e}^{(\mathcal{D}_{n:n})}$. We have $\alpha_n^{(\mathcal{D})} \geq \zeta_{81/256}^{(\mathcal{D}_{n:n})} \geq -\ln(1 - 81/256) \cdot \mathbb{E}[\mathcal{D}_{n:n}] \geq \frac{1}{3} \cdot \mathbb{E}[\mathcal{D}_{n:n}]$.

For the upper bound, we have $\alpha_n^{(\mathcal{D})} \leq \zeta_{1/e}^{(\mathcal{D}_{n:n})} \leq -\frac{\ln(1-1/e)}{1/e} \cdot \mathbb{E}[\mathcal{D}_{n:n}] \leq \frac{5}{4} \cdot \mathbb{E}[\mathcal{D}_{n:n}]$. ■

Appendix C. Deferred proofs from Section 3

Proof of Lemma 12 $\mathbb{E}[\mathcal{D}] = \sqrt{2/\pi}$ is a standard property to the half-normal distribution (and can also be confirmed by computing the mean of a folded-normal with parameter $\mu = 0$ (Leone et al., 1961)).

For the MHR property, it suffices to show that $\frac{f_{\mathcal{D}}(x)}{1-F_{\mathcal{D}}(x)}$ is an increasing function. Note that its derivative is $\frac{f'_{\mathcal{D}}(x)(1-F_{\mathcal{D}}(x))+f^2_{\mathcal{D}}(x)}{(1-F_{\mathcal{D}}(x))^2}$, so we need the numerator to be non-negative.

As $f_{\mathcal{D}}(x) = \sqrt{\frac{2}{\pi}} \exp\left(\frac{-x^2}{2}\right) = 2\phi(x)$ and $F_{\mathcal{D}}(x) = erf\left(\frac{x}{\sqrt{2}}\right) = 2\Phi(x) - 1$, the numerator is

$$f'_{\mathcal{D}}(x)(1 - F_{\mathcal{D}}(x)) + f^2_{\mathcal{D}}(x) = -2x\phi(x)(2 - 2\Phi(x)) + 4\phi^2(x) = 4\phi(x)\left(\phi(x) - x(1 - \Phi(x))\right),$$

where the last quantity is non-negative as $\phi(x) \geq 0$ and by Lemma 4, proving our claim.

Finally, since $\mathcal{D}$ is MHR, we use results from Section 2.2 to bound $\mathbb{E}[\mathcal{D}_{n:n}]$. Observe that

$$\begin{aligned}
F_{\mathcal{D}}(\sqrt{\ln n}) &= 2\Phi(\sqrt{\ln n} - 1) \\
&\leq^{(\text{Lemma } 4)} 2\left(1 - \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\ln n}}{1 + \ln n} \exp\left(-\frac{1}{2} \cdot \ln n\right)\right) - 1 \\
&= 1 - \sqrt{\frac{2}{\pi}} \cdot \frac{\sqrt{\ln n}}{n^{1/2}(1 + \ln n)} \\
&\leq 1 - \frac{1}{n},
\end{aligned}$$

where the last inequality holds for all $n \geq 8$. Therefore, $\alpha_n^{(\mathcal{D})} \geq \sqrt{\ln n}$, which implies

$$\mathbb{E}[\mathcal{D}_{n:n}] \geq^{(\text{Lemma } 10)} \frac{4}{5}\sqrt{\ln n}.$$

Similarly,

$$\begin{aligned}
F_{\mathcal{D}}(\sqrt{2\ln n}) &= 2\Phi(\sqrt{2\ln n} - 1) \\
&\geq^{(\text{Lemma } 4)} 2\left(1 - \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2\ln n}} \exp\left(-\frac{1}{2} \cdot 2\ln n\right)\right) - 1 \\
&= 1 - \sqrt{\frac{2}{\pi}} \cdot \frac{1}{n\sqrt{2\ln n}} \\
&\geq 1 - \frac{1}{n}.
\end{aligned}$$

Therefore, $\alpha_n^{(\mathcal{D})} \leq \sqrt{2\ln n}$, which means $\mathbb{E}[\mathcal{D}_{n:n}] \leq^{(\text{Lemma } 10)} 3\sqrt{2}\sqrt{\ln n}$. ■

Proof of Lemma 14 Let's first compute $\mathbb{E}[X_i \mid Y_i = y_i]$ exactly. We have $\mathbb{E}[X_i \mid Y_i = y_i] = \frac{\int_0^\infty x \cdot f_{\mathcal{D}}(x) \cdot f_{\mathcal{N}(0,\sigma_i^2)}(y_i - x)\, dx}{\int_0^\infty f_{\mathcal{D}}(x) \cdot f_{\mathcal{N}(0,\sigma_i^2)}(y_i - x)\, dx}$. We transform the numerator.

$$\int_0^\infty f_{\mathcal{D}}(x) \cdot f_{\mathcal{N}(0,\sigma_i^2)}(y_i - x)\, dx = \int_0^\infty \frac{\sqrt{2}}{\sqrt{\pi}} \exp\left(\frac{-x^2}{2}\right) \cdot \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left(\frac{-(y_i - x)^2}{2\sigma_i^2}\right) dx$$

$$= \frac{1}{\sigma_i \pi} \int_0^\infty \exp\left(-\frac{1}{2}\left(x^2 + \left(\frac{y_i - x}{\sigma_i}\right)^2\right)\right) dx$$

Let's focus on $x^2 + \left(\frac{y_i-x}{\sigma_i}\right)^2$:

$$\begin{aligned}
x^2 + \left(\frac{y_i - x}{\sigma_i}\right)^2 &= \frac{(x\sigma_i)^2 + y_i^2 - 2y_i x + x^2}{\sigma_i^2} \\
&= \frac{\left(x\sqrt{\sigma_i^2+1}\right)^2 - 2y_i x + y_i^2}{\sigma_i^2} \\
&=^{(\text{let } \lambda = \sqrt{\sigma_i^2+1})} \frac{(\lambda x)^2 - 2\frac{y_i}{\lambda} \cdot \lambda x + \left(\frac{y_i}{\lambda}\right)^2 + y_i^2\left(1 - \frac{1}{\lambda^2}\right)}{\sigma_i^2} \\
&=^{\left(\text{let } \rho = \frac{y_i^2\left(1-\frac{1}{\lambda^2}\right)}{\sigma_i^2}\right)} \left(\frac{\lambda x - \frac{y}{\lambda}}{\sigma_i}\right)^2 + \rho
\end{aligned}$$

Observe that λ and ρ only depends on σ_i and y_i. Therefore, coming back to the previous integral:

$$\begin{aligned}
\int_0^\infty f_{\mathcal{D}}(x) \cdot f_{\mathcal{N}(0,\sigma_i^2)}(y_i - x)\, dx &= \frac{1}{\sigma_i \pi} \int_0^\infty \exp\left(-\frac{1}{2}\left(\left(\frac{\lambda x - \frac{y}{\lambda}}{\sigma_i}\right)^2 + \rho\right)\right) dx \\
&= \frac{e^{-\rho/2}\lambda\sqrt{2}}{\sqrt{\pi}} \int_0^\infty \frac{1}{\sqrt{2\pi}\cdot\lambda\sigma_i} \exp\left(-\frac{1}{2}\left(\frac{x - \frac{y_i}{\lambda^2}}{\lambda\sigma_i}\right)^2\right) dx \\
&= \frac{e^{-\rho/2}\lambda\sqrt{2}}{\sqrt{\pi}} \int_0^\infty f_{\mathcal{N}\left(\frac{y_i}{\lambda^2}, \left(\frac{\sigma_i}{\lambda}\right)^2\right)}(x)\, dx
\end{aligned}$$

Calculated similarly, we have

$$\int_0^\infty x \cdot f_{\mathcal{D}}(x) \cdot f_{\mathcal{N}(0,\sigma_i^2)}(y_i - x)\, dx = \frac{e^{-\rho/2}\lambda\sqrt{2}}{\sqrt{\pi}} \int_0^\infty x \cdot f_{\mathcal{N}\left(\frac{y_i}{\lambda^2}, \left(\frac{\sigma_i}{\lambda}\right)^2\right)}(x)\, dx$$

Therefore

$$\begin{aligned}
\mathbb{E}[X_i \mid Y_i = y_i] &= \frac{\int_0^\infty x \cdot f_{\mathcal{D}}(x) \cdot f_{\mathcal{N}(0,\sigma_i^2)}(y_i - x)\, dx}{\int_0^\infty f_{\mathcal{D}}(x) \cdot f_{\mathcal{N}(0,\sigma_i^2)}(y_i - x)\, dx} \\
&= \frac{\frac{e^{-\rho/2}\lambda\sqrt{2}}{\sqrt{\pi}} \int_0^\infty x \cdot f_{\mathcal{N}\left(\frac{y_i}{\lambda^2}, \left(\frac{\sigma_i}{\lambda}\right)^2\right)}(x)\, dx}{\frac{e^{-\rho/2}\lambda\sqrt{2}}{\sqrt{\pi}} \int_0^\infty f_{\mathcal{N}\left(\frac{y_i}{\lambda^2}, \left(\frac{\sigma_i}{\lambda}\right)^2\right)}(x)\, dx} \\
&= \frac{\int_0^\infty x \cdot f_{\mathcal{N}\left(\frac{y_i}{\lambda^2}, \left(\frac{\sigma_i}{\lambda}\right)^2\right)}(x)\, dx}{\int_0^\infty f_{\mathcal{N}\left(\frac{y_i}{\lambda^2}, \left(\frac{\sigma_i}{\lambda}\right)^2\right)}(x)\, dx} \\
&= \mathbb{E}\left[t \;\middle|\; t \sim \mathcal{N}\left(\frac{y_i}{\lambda^2}, \left(\frac{\sigma_i}{\lambda}\right)^2\right) \cap t \geq 0\right].
\end{aligned}$$

This last quantity is the mean of the normal distribution $\mathcal{N}\left(\frac{y_i}{\sigma_i^2+1}, \left(\frac{\sigma_i}{\sqrt{\sigma_i^2+1}}\right)^2\right)$ truncated to $[0,\infty)$ (as $\lambda = \sqrt{\sigma_i^2+1}$). From Johnson et al. (1994) (Chapter 13, Section 10.1), we can conclude that

$$\mathbb{E}[X_i \mid Y_i = y_i] = \frac{y_i}{\sigma_i^2+1} + \frac{\phi\left(\frac{-y_i}{\sigma_i\sqrt{\sigma_i^2+1}}\right)}{1-\Phi\left(\frac{-y_i}{\sigma_i\sqrt{\sigma_i^2+1}}\right)} \cdot \frac{\sigma_i}{\sqrt{\sigma_i^2+1}}.$$

Finally, we move on to the bound proposed in the statement. Let's first consider the case where $y_i \geq 0$. In this case, $U_{\sigma_i}(y_i) = \sqrt{\frac{2}{\pi}} + \frac{y_i}{\sigma_i^2+1}$. Observe that $\phi(x) \leq \frac{1}{\sqrt{2\pi}}$ for all x, $1-\Phi(x) \geq \frac{1}{2}$ for all $x \leq 0$, and $\frac{\sigma_i}{\sqrt{\sigma_i^2+1}} \leq 1$ for all $\sigma_i \geq 0$. Therefore,

$$\mathbb{E}[X_i \mid Y_i = y_i] = \frac{y_i}{\sigma_i^2+1} + \frac{\phi\left(\frac{-y_i}{\sigma_i\sqrt{\sigma_i^2+1}}\right)}{1-\Phi\left(\frac{-y_i}{\sigma_i\sqrt{\sigma_i^2+1}}\right)} \cdot \frac{\sigma_i}{\sqrt{\sigma_i^2+1}} \leq \frac{y_i}{\sigma_i^2+1} + \sqrt{\frac{2}{\pi}} = U_{\sigma_i}(y_i).$$

If $y_i < 0$, we use the property that $\mathbb{E}[X_i \mid Y_i = y_i] \leq \mathbb{E}[X_i \mid Y_i = 0]$ (this is due to the monotonicity of $\mathbb{E}[X_i \mid Y_i = y_i]$; see Lemma 30): $\mathbb{E}[X_i \mid Y_i = y_i] \leq \mathbb{E}[X_i \mid Y_i = 0] \leq U_{\sigma_i}(0) = U_{\sigma_i}(y_i)$. ■

Appendix D. Deferred proofs from Section 4.

Proof of Theorem 16 Consider $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_n) \in \mathcal{S}^{\text{MHR}}_{(\mathcal{D},n)}$ where, without loss of generality, we have $\sigma_1 \leq \sigma_2 \leq \cdots \leq \sigma_n$. As $\boldsymbol{\sigma} \in \mathcal{S}^{\text{MHR}}_{(\mathcal{D},n)}$, there exists a constant $c = c_{(\mathcal{D},n)} \in (0,1]$ such that $\sigma_{n^c} \leq \frac{\mathbb{E}[\mathcal{D}_{n^c:n^c}]}{18\sqrt{2c\ln n}}$.

Consider the event that $|\epsilon_i| \leq \sigma_i\sqrt{2c\ln n}$ for all $1 \leq i \leq n^c$. Following the same analysis as the proof of Theorem 15, for any box $i \in [1, n^c]$, we have

$$\begin{aligned}
\Pr\left[|\epsilon_i| \leq \sigma_i\sqrt{2c\ln n}\right] &= \Pr\left[|\epsilon_i| \leq \sigma_i\sqrt{2\ln n^c}\right] \\
&= \Pr\left[|\mathcal{N}(0,\sigma_i^2)| \leq \sigma_i\sqrt{2\ln n^c}\right] \\
&= 2\Phi\left(\sqrt{2\ln n^c}\right) - 1 \\
&\geq^{\text{(Lemma 4)}} 2\left(1 - \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\ln n^c}}\exp\left(-\frac{1}{2}\cdot 2\ln n^c\right)\right) - 1 \\
&= 1 - \frac{1}{n^c\sqrt{c\pi\ln n}},
\end{aligned}$$

and therefore

$$\Pr\left[|\epsilon_i| \leq \sigma_i\sqrt{2c\ln n}, \forall i \in [1,n^c]\right] \geq \left(1 - \frac{1}{n^c\sqrt{c\pi\ln n}}\right)^{n^c}$$

$$\geq^{\text{(Bernoulli's inequality)}} 1 - \frac{n^c}{n^c\sqrt{c\pi \ln n}} \geq \frac{1}{2},$$

where the last inequality holds for all $n \geq e^{\frac{4}{c\pi}}$.

Since $\sigma_i \leq \frac{\mathbb{E}[\mathcal{D}_{n^c:n^c}]}{18\sqrt{2c\ln n}}$ for all $i \in [1, n^c]$, we can conclude that $\Pr[\max_{i\in[1,n^c]} |\epsilon_i| \leq \frac{1}{18} \cdot \mathbb{E}[\mathcal{D}_{n^c:n^c}]] \geq \frac{1}{2}$. Conditioned on this event, for all $i \in [1, n^c]$, we have $x_i - \frac{1}{18} \cdot \mathbb{E}[\mathcal{D}_{n^c:n^c}] \leq y_i \leq x_i + \frac{1}{18} \cdot \mathbb{E}[\mathcal{D}_{n^c:n^c}]$; therefore, for all $k \leq n^c$, we have $\max_{i\in[1,k]} y_i \geq \max_{i\in[1,k]} x_i - \frac{1}{9} \cdot \mathbb{E}[\mathcal{D}_{n^c:n^c}]$.

We analyze the performance of IgnoreLargeExp conditioned on this event. Recall that the algorithm draws $\alpha \sim U[0, 1]$, and then outputs $\arg\max_{i\in[1,n^\alpha]} y_i$. We consider two cases for α:

Case $\alpha > c$: we will lower bound the expected reward of IgnoreLargeExp by 0.

Case $\alpha \leq c$: IgnoreLargeExp is going to pick the box with the largest y_i among the first n^α boxes. By our observation, IgnoreLargeExp's reward in this case is at least $\max_{i\in[1,n^\alpha]} x_i - \frac{1}{9} \cdot \mathbb{E}[\mathcal{D}_{n^c:n^c}]$, and therefore the expected reward of IgnoreLargeExp in this case is at least $\mathbb{E}[\mathcal{D}_{n^\alpha:n^\alpha}] - \frac{1}{9} \cdot \mathbb{E}[\mathcal{D}_{n^c:n^c}]$. By Lemma 17, since $\frac{c}{\alpha} \leq 1$, we have $\mathbb{E}[\mathcal{D}_{n^c:n^c}] \leq \frac{4c}{\alpha} \cdot \mathbb{E}[\mathcal{D}_{n^\alpha:n^\alpha}]$. Continuing our derivation, the expected reward of IgnoreLargeExp is at least

$$\mathbb{E}[\mathcal{D}_{n^\alpha:n^\alpha}] - \cdot\mathbb{E}[\mathcal{D}_{n^c:n^c}] \geq \frac{\alpha}{4c} \cdot \mathbb{E}[\mathcal{D}_{n^c:n^c}] - \frac{1}{9} \cdot \mathbb{E}[\mathcal{D}_{n^c:n^c}].$$

Therefore, conditioned on the event that $\max_{i\in[1,n^c} |\epsilon_i| \leq \frac{1}{18} \cdot \mathbb{E}[\mathcal{D}_{n^c:n^c}]$, IgnoreLargeExp's expected reward is lower bounded by

$$\int_{\alpha=0}^{c} \frac{\alpha}{4c} \cdot \mathbb{E}[\mathcal{D}_{n^c:n^c}] - \frac{1}{9} \cdot \mathbb{E}[\mathcal{D}_{n^c:n^c}]\, d\alpha = \frac{1}{72} \cdot \mathbb{E}[\mathcal{D}_{n^c:n^c}].$$

In outcomes outside this event, we can lower bound IgnoreLargeExp's expected reward by 0. Combining everything, IgnoreLargeExp's expected reward is

$$R_{\mathsf{IgnoreLargeExp}}(\mathcal{D}, \boldsymbol{\sigma}) \geq \frac{1}{2} \cdot \frac{1}{72} \cdot \mathbb{E}[\mathcal{D}_{n^c:n^c}] \geq^{\text{(Lemma 17)}} \frac{c^2}{576} \cdot \mathbb{E}[\mathcal{D}_{n:n}].$$

■

Appendix E. Deferred proofs from Section 5

E.1. Deferred proofs from Section 5.1

Proof of Lemma 22 Formally, this event is $\max_{i\in[n-c_b+1,n]} \epsilon_i > \beta_{n^2}^{(\mathcal{D}_{n:n})}$. We have

$$\begin{aligned}
\Pr\left[\max_{i\in[n-c_b+1,n]} \epsilon_i > \beta_{n^2}^{(\mathcal{D}_{n:n})}\right] &= 1 - \Pr\left[\max_{i\in[n-c_b+1,n]} \epsilon_i \leq \beta_{n^2}^{(\mathcal{D}_{n:n})}\right] \\
&= 1 - \Pr\left[\mathcal{N}(0, \sigma_b^2) \leq \beta_{n^2}^{(\mathcal{D}_{n:n})}\right]^{c_b} \\
&= 1 - \Pr\left[\mathcal{N}(0, \sigma_b^2) \leq \frac{\sigma_b}{6\sqrt{\ln n}}\right]^{c_b}
\end{aligned}$$

$$\geq 1 - \Pr\left[\mathcal{N}(0, \sigma_b^2) \leq \frac{\sigma_b}{6}\right]^{6 \ln n}$$

Using the fact that $\Pr\left[\mathcal{N}(\mu, \sigma^2) \leq x\right] = \Phi(\frac{x-\mu}{\sigma})$, where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$ is the CDF of the standard normal distribution, we have that $\Pr\left[\max_{i \in [n-c_b+1, n]} \epsilon_i > \beta_{n^2}^{(\mathcal{D}_{n:n})}\right] \geq 1 - \Phi\left(\frac{1}{6}\right)^{6 \ln n}$. Since $\Phi\left(\frac{1}{6}\right) < 0.6$ we have

$$\Pr\left[\max_{i \in [n-c_b+1, n]} \epsilon_i > \beta_{n^2}^{(\mathcal{D}_{n:n})}\right] \geq 1 - ((0.6)^2)^{3 \ln n} \geq 1 - \left(\frac{1}{e}\right)^{3 \ln n} \geq 1 - \frac{1}{n^3}.$$

■

Proof of Lemma 23 Note that as $\epsilon_i \sim \mathcal{N}(0, \sigma_b^2)$ and $\sigma_b = 6\beta_{n^2}^{(\mathcal{D}_{n:n})} \sqrt{\ln n}$ we have

$$\begin{aligned}
\Pr[\epsilon_i \leq 12\beta_{n^2}^{(\mathcal{D}_{n:n})} \ln n] &= \Pr[\epsilon_i \leq 2\sqrt{\ln n} \cdot \sigma_b)] \\
&= \Phi(2\sqrt{\ln n}) \\
&\geq^{\text{(Lemma 4)}} 1 - \frac{1}{\sqrt{2\pi}} \frac{1}{2\sqrt{\ln n}} \cdot \exp(-2 \ln n) \\
&= 1 - \frac{1}{2\sqrt{2\pi}} \frac{1}{n^2 \sqrt{\ln n}} \\
&\geq 1 - \frac{1}{n^2}.
\end{aligned}$$

■

Proof of Lemma 24 Slightly overloading notation, let $f(x)$ be the PDF of Z. Let $A(y) = \int_0^V x \cdot f(x) \cdot f_{\mathcal{N}}(y - x)\, dx$ and $B(y) = \int_0^V f(x) \cdot f_{\mathcal{N}}(y - x)\, dx$, then $\mathbb{E}[Z \mid Z + \mathcal{N}(0, \sigma^2) = y] = \frac{A(y)}{B(y)}$. From Lemma 30 we know that $\mathbb{E}[Z \mid Z + \mathcal{N}(0, \sigma^2) = y]$ is monotone non-decreasing in y.

Let $r = \frac{\sigma}{V}$. Consider $y^* = \frac{\sigma^2}{2V} = \sigma \cdot \frac{r}{2}$. As $\sigma > 2V$ or $r > 2$, we then have $y^* > \sigma > V$, which implies that $f_{\mathcal{N}}(y^* - V) \geq f_{\mathcal{N}}(y^* - x)$ for all $x \in [0, V]$. We then have the following bound on $A(y^*)$:

$$\begin{aligned}
A(y^*) &= \int_0^V x \cdot f(x) \cdot f_{\mathcal{N}}(y^* - x)\, dx \\
&\leq \int_0^V x \cdot f(x) \cdot f_{\mathcal{N}}(y^* - V)\, dx \\
&= \mathbb{E}[Z] \cdot f_{\mathcal{N}}(y^* - V) \\
&= \mathbb{E}[Z] \cdot \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y^* - V}{\sigma}\right)^2\right).
\end{aligned}$$

Recalling that $y^* = \sigma \cdot \frac{r}{2}$ and that $V = \frac{\sigma}{r}$, we have:

$$A(y^*) = \frac{1}{\sigma\sqrt{2\pi}} \mathbb{E}[Z] \cdot \exp\left(-\frac{1}{2}\left(\frac{r}{2} - \frac{1}{r}\right)^2\right)$$

$$
\begin{aligned}
&= \frac{1}{\sigma\sqrt{2\pi}}\mathbb{E}[Z]\cdot \exp\left(-\frac{r^2}{8}+\frac{1}{2}-\frac{1}{2r^2}\right)\\
&\le \frac{1}{\sigma\sqrt{2\pi}}\mathbb{E}[Z]\cdot \frac{\sqrt{e}}{\exp(r^2/8)}.
\end{aligned}
$$

Meanwhile, for $B(y^*)$, we have

$$
\begin{aligned}
B(y^*) &= \int_0^V f(x)\cdot f_{\mathcal{N}}(y^*-x)\,dx\\
&\ge^{(y^*\ge V)} \int_0^V f(x)\cdot f_{\mathcal{N}}(y^*)\,dx\\
&= f_{\mathcal{N}}(y^*)\cdot \int_0^V f(x)\,dx\\
&= f_{\mathcal{N}}(y^*)\\
&= \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{y^*}{\sigma}\right)^2\right)\\
&= \frac{1}{\sigma\sqrt{2\pi}}\cdot\frac{1}{\exp(r^2/8)}.
\end{aligned}
$$

Therefore $A(y^*) \le 2\mathbb{E}[Z]\cdot B(y^*)$, and thus $\mathbb{E}[Z \mid Z+\mathcal{N}(0,\sigma^2)=y^*] = \frac{A(y^*)}{B(y^*)}$ is at most $2\mathbb{E}[Z]$. Since $\mathbb{E}[Z \mid Z+\mathcal{N}(0,\sigma^2)=y]$ is monotone non-decreasing in y (Lemma 30), we can conclude that $\mathbb{E}[Z \mid Z+\mathcal{N}(0,\sigma^2)=y] \le 2\mathbb{E}[Z]$ for all $y \le y^* = \frac{\sigma^2}{2V}$. ■

E.2. Deferred proofs from Section 5.2

Our goal is to prove Lemma 29. To make the presentation cleaner, we define the following events.

Definition 33 *Let*

- $\mathcal{E}_1$ *be the event of* $\max_{i\in[2,c_s+1]}\epsilon_i \le \frac{\theta^*\sigma_s}{37}$.
- $\mathcal{E}_2$ *be the event of* $\max_{i\in[c_s+2,n]}\epsilon_i - \theta^*\sigma_b \ge \sigma_b$.
- $\mathcal{E}_2'$ *be the event of* $\max_{i\in[c_s+2,n]} Y_i - c\sigma_b \ge \sigma_b$ *for all* $c<\theta^*$.
- $\mathcal{E}_3$ *be the event of* $\max_{i\in[c_s+2,n]}\epsilon_i \le 12\alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})}\ln n$.
- $\mathcal{E}_3'$ *be the event of* $\max_{i\in[c_s+2,n]} Y_i \le 18\alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})}\ln n$.

Recall that $\mathcal{E}^*$ is the event that $X_i \le \alpha_{n^{1/10000}}^{(\mathcal{D}_{c_s:c_s})}$ for all small noise boxes i, and $X_j \le \alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})}$ for all remaining boxes j.

We first show some technical lemmas.

Lemma 34 *For all $n \ge n_0$, for some constant n_0,* $\Pr[\mathcal{E}_1] \ge 1-\frac{1}{\ln n}$.

Proof of Lemma 34 Observe that ϵ_i are values drawn from $\mathcal{N}(0, \sigma_s^2)$. We then have

$$\begin{aligned}
\Pr\left[\max_{i \in [2, c_s+1]} \epsilon_i \leq \frac{\theta^* \sigma_s}{37}\right] &= \Pr\left[\mathcal{N}(0, \sigma_s^2) \leq \frac{\theta^* \sigma_s}{37}\right]^{c_s} \\
&= \Phi\left(\frac{\theta^*}{37}\right)^{n^{1/5626}} \\
&\geq^{\text{(Lemma 4)}} \left(1 - \frac{1}{\sqrt{2\pi}} \frac{37\sqrt{2}}{\sqrt{\ln n}} \exp\left(-\frac{1}{2} \cdot \frac{1}{2738} \ln n\right)\right)^{n^{1/5626}} \\
&\geq^{\text{(Bernoulli's inequality)}} 1 - \frac{37}{\sqrt{\pi}\sqrt{\ln n}} n^{1/5626 - 1/5476} \\
&\geq 1 - \frac{1}{\ln n}.
\end{aligned}$$

■

Lemma 35 *For all $n \geq n_0$, for some constant n_0, $\Pr[\mathcal{E}_2] \geq 1 - \frac{1}{\ln n}$.*

Proof of Lemma 35 Observe that ϵ_i are values drawn from $\mathcal{N}(0, \sigma_b^2)$. We then have

$$\begin{aligned}
\Pr\left[\max_{i \in [c_s+2, n]} \epsilon_i - \theta^* \sigma_b \geq \sigma_b\right] &= 1 - \Pr\left[\max_{i \in [c_s+2, n]} \epsilon_i \leq \theta^* \sigma_b + \sigma_b\right] \\
&= 1 - \Pr\left[\mathcal{N}(0, \sigma_b^2) \leq \theta^* \sigma_b + \sigma_b\right]^{n - c_s - 1} \\
&= 1 - (\Phi(\theta^* + 1))^{n - c_s - 1} \\
&\geq 1 - \left(\Phi(\sqrt{2}\theta^*)\right)^{n/2} \\
&\geq^{\text{(Lemma 4)}} 1 - \left(1 - \frac{1}{\sqrt{2\pi}} \frac{\sqrt{2}\theta^*}{2(\theta^*)^2 + 1} \exp(-(\theta^*)^2)\right)^{\frac{n}{2}} \\
&\geq^{\text{(Bernoulli's inequality)}} 1 - \frac{1}{1 + \frac{n}{2} \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\ln n}}{\ln n + 1} \exp\left(-\frac{\ln n}{2}\right)} \\
&= 1 - \frac{1}{1 + \frac{\sqrt{n}}{2} \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\ln n}}{\ln n + 1}} \\
&\geq 1 - \frac{1}{\ln n}.
\end{aligned}$$

■

Lemma 36 *For any MHR distribution $\mathcal{D}$ supported on $[0, \infty)$ and for all $n \geq 1$, we have*

$$\mathbb{E}[\mathcal{D}_{n:n}] \leq (\ln n + 1) \cdot \mathbb{E}[\mathcal{D}].$$

Proof of Lemma 36 The lemma is an immediate consequence of the following result from Barlow and Proschan (1996) (Corollary 4.10 on page 33):

Lemma 37 (Barlow and Proschan (1996)) *If X_i, $i = 1, \ldots, n$, are MHR*[1] *random variables with mean μ_i and cdf $F_i(.)$, and $G_i(x) = 1 - \exp(-x/\mu_i)$, then:*

$$\int_0^\infty 1 - \prod_{i=1}^n F_i(x)dx \leq \int_0^\infty 1 - \prod_{i=1}^n G_i(x)dx.$$

Applying this result for the case of $F(x) = F_i(x)$ for all i, we have that

$$\mathbb{E}[\mathcal{D}_{n:n}] = \int_0^\infty 1 - F^n(x)dx \leq \int_0^\infty 1 - (1 - e^{-\frac{x}{\mathbb{E}[\mathcal{D}]}})^n dx = \mathbb{E}[\mathcal{D}] \sum_{i=1}^n \frac{1}{i}$$

Using the fact that $\sum_{i=1}^n \frac{1}{i} \leq \ln(n) + 1$, we get the lemma. ■

Lemma 38 *For any $n \geq 1$ and $m \geq 2$, we have*

$$\mathbb{E}[\mathcal{D}_{n:n} \mid \mathcal{D}_{n:n} > \alpha_m^{(\mathcal{D}_{n:n})}] \cdot \Pr[\mathcal{D}_{n:n} > \alpha_m^{(\mathcal{D}_{n:n})}] \leq \frac{15(\ln m + \ln n + 1)\mathbb{E}[\mathcal{D}]}{2m}.$$

Proof of Lemma 38 We use the following result from Cai and Daskalakis (2011):

Lemma 39 (Cai and Daskalakis (2011)) *For any MHR distribution $\mathcal{D}$ and any $m \geq 2$, we have*

$$\mathbb{E}[\mathcal{D} \mid \mathcal{D} \geq \alpha_m^{(\mathcal{D})}] \cdot \Pr[\mathcal{D} \geq \alpha_m^{(\mathcal{D})}] \leq \frac{6\alpha_m^{(\mathcal{D})}}{m}.$$

Since order statistics of MHR distributions are also MHR (Lemma 31), $\mathcal{D}_{n:n}$ and $(\mathcal{D}_{n:n})_{m:m} = \mathcal{D}_{nm:nm}$ are MHR. Then, by Lemma 10 we have that

$$\alpha_m^{(\mathcal{D}_{n:n})} \leq \frac{5}{4} \cdot \mathbb{E}[\mathcal{D}_{nm:nm}]. \tag{1}$$

Towards proving the lemma, we then get

$$\begin{aligned}
\mathbb{E}[\mathcal{D}_{n:n} \mid \mathcal{D}_{n:n} > \alpha_m^{(\mathcal{D}_{n:n})}] \cdot \Pr[\mathcal{D}_{n:n} > \alpha_m^{(\mathcal{D}_{n:n})}] &\leq^{(\text{Lemma } 39)} \frac{6\alpha_m^{(\mathcal{D}_{n:n})}}{m} \\
&\leq^{(\text{Equation } (1))} 6 \cdot \frac{5}{4} \cdot \frac{\mathbb{E}[\mathcal{D}_{nm:nm}]}{m} \\
&\leq^{(\text{Lemma } 36)} \frac{15(\ln(nm) + 1)}{2m} \cdot \mathbb{E}[\mathcal{D}] \\
&= \frac{15(\ln(n) + \ln(m) + 1)}{2m} \cdot \mathbb{E}[\mathcal{D}].
\end{aligned}$$

■

Lemma 40 *For all $n \geq n_0$, for some constant n_0, for any large noise box i,*

$$\Pr\left[Y_i \leq 18\alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})} \ln n\right] \geq 1 - \frac{1}{n^2}.$$

Proof of Lemma 40 The proof is similar to that of Lemma 23. Note that as $\epsilon_i \sim \mathcal{N}(0, \sigma_b^2)$ and $\sigma_b = 6\beta_{n^2}^{(\mathcal{D}_{n:n})}\sqrt{\ln n}$ we have

$$\begin{aligned}
\Pr[\epsilon_i \leq 12\beta_{n^2}^{(\mathcal{D}_{n:n})} \ln n] &= \Pr[\epsilon_i \leq 2\sqrt{\ln n} \cdot \sigma_b)] \\
&= \Phi(2\sqrt{\ln n}) \\
&\geq^{\text{(Lemma 4)}} 1 - \frac{1}{\sqrt{2\pi}} \frac{1}{2\sqrt{\ln n}} \cdot \exp(-2\ln n) \\
&= 1 - \frac{1}{2\sqrt{2\pi}} \frac{1}{n^2\sqrt{\ln n}} \\
&\geq 1 - \frac{1}{n^2}.
\end{aligned}$$

■

Lemma 41 *If $\mathcal{E}^* \cap \mathcal{E}_1$ occurs, for all $c \geq \theta^*$,* $\mathsf{LinearFixed}_c$ *does not choose a small noise box.*

Proof of Lemma 41 Consider any $c \geq \theta^*$. Observe that $Y_1 - c\sigma_1^* = X_1 \geq 0$. We show that conditioned on $\mathcal{E}_1 \cap \mathcal{E}^*$, we have $\max_{i \in [2, c_s+1]} Y_i \leq \theta^* \sigma_s$. We first note that from Lemma 8, we have $\Pr[\mathcal{D}_{c_s:c_s} < 2\mathbb{E}[\mathcal{D}_{c_s:c_s}]] \geq 1 - \frac{1}{c_s^{3/5}} = 1 - \frac{1}{n^{1/5626 \cdot 3/5}} > 1 - \frac{1}{n^{1/10000}}$. Therefore, by Definition 6, $2\mathbb{E}[\mathcal{D}_{c_s:c_s}] \geq \alpha_{n^{1/10000}}^{(\mathcal{D}_{c_s:c_s})}$. Then, conditioned on both $\mathcal{E}_1$ and $\mathcal{E}^*$, we have that for any small noise box i:

$$Y_i = X_i + \epsilon_i <^{\text{(Definition 33)}} \alpha_{n^{1/10000}}^{(\mathcal{D}_{c_s:c_s})} + \frac{\theta^* \sigma_s}{37} \leq 2\mathbb{E}[\mathcal{D}_{c_s:c_s}] + \frac{\theta^* \sigma_s}{37} = \frac{36\theta^* \sigma_s}{37} + \frac{\theta^* \sigma_s}{37} = \theta^* \sigma_s.$$

Therefore, conditioned on $\mathcal{E}_1$ and $\mathcal{E}^*$, we have $\max_{i \in [2, c_s+1]} Y_i - c\sigma_i^* \leq \theta^* \sigma_s - c\sigma_s < 0$, i.e. $Y_1 - c\sigma_1^* > Y_i - \max_{i \in [2, c_s+1]} Y_i - c\sigma_i^*$ and hence $\mathsf{LinearFixed}_c$ does not choose any small box i.
■

Lemma 42 *If $\mathcal{E}^* \cap \mathcal{E}_1 \cap \mathcal{E}_2'$ occurs, for all $c < \theta^*$,* $\mathsf{LinearFixed}_c$ *chooses some large noise box.*

Proof of Lemma 42 Consider any $c \geq \theta^*$. Observe that conditioned on $\mathcal{E}_2'$, $\max_{i \in [c_s+2, n]} Y_i - c\sigma_i^* \geq^{\text{(Definition 33)}} \sigma_b$.

From Lemma 8, we have $\Pr[\mathcal{D}_{c_s:c_s} < 2\mathbb{E}[\mathcal{D}_{c_s:c_s}]] \geq 1 - \frac{1}{c_s^{3/5}} = 1 - \frac{1}{n^{1/5626 \cdot 3/5}} > 1 - \frac{1}{n^{1/10000}}$. Therefore, $2\mathbb{E}[\mathcal{D}_{c_s:c_s}] \geq \alpha_{n^{1/10000}}^{(\mathcal{D}_{c_s:c_s})}$. Then, conditioned on $\mathcal{E}^* \cap \mathcal{E}_1$, we have that for all $i \in [2, c_s+1]$:

$$Y_i = X_i + \epsilon_i < \alpha_{n^{1/10000}}^{(\mathcal{D}_{c_s:c_s})} + \frac{\theta^* \sigma_s}{37} \leq 2\mathbb{E}[\mathcal{D}_{c_s:c_s}] + \frac{\theta^* \sigma_s}{37} = \frac{36\theta^* \sigma_s}{37} + \frac{\theta^* \sigma_s}{37} = \theta^* \sigma_s.$$

Furthermore, we can show that $\theta^* \sigma_s$ can be upper bounded. From Lemma 31, we know that $\mathcal{D}_{a:a}$ is MHR for any $a \geq 1$. Then, by Lemma 10 we have that

$$\alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})} \geq \frac{1}{3} \cdot \mathbb{E}[(\mathcal{D}_{n-c_s:n-c_s})_{n^{1/10000}:n^{1/10000}}] \tag{2}$$

Therefore,

$$\begin{aligned}
\sigma_b &= 6\alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})}\sqrt{\ln n} \\
&\geq^{\text{(Equation (2))}} 6 \cdot \frac{1}{3}\mathbb{E}[\mathcal{D}_{(n-c_s)\cdot n^{1/10000}:(n-c_s)\cdot n^{1/10000}}]\sqrt{\ln n} \\
&> \frac{5}{2}\mathbb{E}[\mathcal{D}_{(n-c_s)\cdot n^{1/10000}:(n-c_s)\cdot n^{1/10000}}] \\
&>^{(c_s=n^{1/5626})} \frac{5}{2}\mathbb{E}[\mathcal{D}_{c_s:c_s}] \\
&= \theta^*\sigma_s.
\end{aligned} \tag{3}$$

Putting everything together,

$$\begin{aligned}
\max_{i\in[1,c_s+1]} Y_i - c\sigma_i^* &= \max\{Y_1, \max_{i\in[2,c_s+1]} Y_i - c\sigma_s\} \\
&\leq \max\{X_1, \max_{i\in[2,c_s+1]} Y_i\} \\
&\leq^{\text{(Definition 33)}} \max\{\alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})}, \theta^*\sigma_s\} \\
&< \sigma_b,
\end{aligned}$$

where the last inequality follows from $\theta^*\sigma_s < \sigma_b$ by Equation (3) and that $\alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})} < \sigma_b = 6\alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})}\sqrt{\ln n}$. Therefore, $\max_{i\in[c_s+2,n]} Y_i - c\sigma_i^* > \max_{i\in[1,c_s+1]} Y_i - c\sigma_i^*$, and so $\mathsf{LinearFixed}_c$ chooses a large noise box.

■

In essence, Lemmas 41 and 42 say that if various combinations of the above events occur, $\mathsf{LinearFixed}_c$ policies make bad choices. We can now prove Lemma 29:

Proof of Lemma 29 We first explore the relationship between the events defined in Definition 33. First, note that $\mathcal{E}_2 \subseteq \mathcal{E}_2'$: if $\max_{i\in[c_s+2,n]} \epsilon_i - \theta^*\sigma_b \geq \sigma_b$, then for all $c < \theta^*$ we have

$$\max_{i\in[c_s+2,n]} Y_i - c\sigma_b = \max_{i\in[c_s+2,n]} (X_i + \epsilon_i) - c\sigma_b \geq \max_{i\in[c_s+2,n]} \epsilon_i - \theta^*\sigma_b \geq \sigma_b.$$

Second, note that $\mathcal{E}^* \cap \mathcal{E}_3 \subseteq \mathcal{E}_3'$, or $\mathcal{E}^* \cap \mathcal{E}_3 \subseteq \mathcal{E}^* \cap \mathcal{E}_3'$: if $\max_{i\in[c_s+2,n]} X_i \leq \alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})}$ and $\max_{i\in[c_s+2,n]} \epsilon_i \leq 12\alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})} \ln n$, then

$$\begin{aligned}
\max_{i\in[c_s+2,n]} Y_i &= \max_{i\in[c_s+2,n]} X_i + \epsilon_i \\
&\leq \max_{i\in[c_s+2,n]} X_i + \max_{i\in[c_s+2,n]} \epsilon_i \\
&\leq \alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})} + 12\alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})} \ln n \\
&\leq 18\alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})}.
\end{aligned}$$

Ultimately, we have $\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3 \cap \mathcal{E}^* \subseteq \mathcal{E}_1 \cap \mathcal{E}'_2 \cap \mathcal{E}'_3 \cap \mathcal{E}^*$, or $\overline{\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3} \cap \mathcal{E}^* \supseteq \overline{\mathcal{E}_1 \cap \mathcal{E}'_2 \cap \mathcal{E}'_3} \cap \mathcal{E}^*$.

We now bound $\mathbb{E}[\max_i X_i \mid \overline{\mathcal{E}_1 \cap \mathcal{E}'_2 \cap \mathcal{E}'_3} \cap \mathcal{E}^*] \cdot \Pr[\overline{\mathcal{E}_1 \cap \mathcal{E}'_2 \cap \mathcal{E}'_3} \mid \mathcal{E}^*]$, which is an upper bound on the contribution of outcomes in $\overline{\mathcal{E}_1 \cap \mathcal{E}'_2 \cap \mathcal{E}'_3} \cap \mathcal{E}^*$ to the overall expected reward of Linear_γ.

$$\begin{aligned}
&\mathbb{E}[\max_i X_i \mid \overline{\mathcal{E}_1 \cap \mathcal{E}'_2 \cap \mathcal{E}'_3} \cap \mathcal{E}^*] \cdot \Pr[\overline{\mathcal{E}_1 \cap \mathcal{E}'_2 \cap \mathcal{E}'_3} \mid \mathcal{E}^*] \\
&\qquad \le \mathbb{E}[\max_i X_i \mid \overline{\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3} \cap \mathcal{E}^*] \cdot \Pr[\overline{\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3} \mid \mathcal{E}^*]
\end{aligned}$$

By Lemma 34, $\Pr[\mathcal{E}_1] \ge 1 - \frac{1}{\ln n}$. By Lemma 35, $\Pr[\mathcal{E}_2] \ge 1 - \frac{1}{\ln n}$. Using Lemma 40, $\Pr[\mathcal{E}_3] \ge (1 - \frac{1}{n^2})^{n - c_s - 1} \ge 1 - \frac{1}{n}$. Therefore, by a union bound, $\Pr[\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3] \ge 1 - \frac{2}{\ln n} - \frac{1}{n} \ge 1 - \frac{3}{\ln n}$. Observe that, $\mathcal{E}_1$, $\mathcal{E}_2$, and $\mathcal{E}_3$ are independent of the X_is, while $\mathcal{E}^*$ is only dependent on X_is. Therefore, $\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3$ and $\mathcal{E}^*$ are independent, and hence $\Pr[\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3 \mid \mathcal{E}^*] = \Pr[\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3] \ge 1 - \frac{3}{\ln n}$, or $\Pr[\overline{\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3} \mid \mathcal{E}^*] \le \frac{3}{\ln n}$. Additionally, $\mathbb{E}[\max_i X_i \mid \overline{\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3} \cap \mathcal{E}^*] = \mathbb{E}[\max_i X_i \mid \mathcal{E}^*]$, as $\mathcal{E}_1$, $\mathcal{E}_2$, and $\mathcal{E}_3$ are events regarding ϵ_is and therefore independent of X_is. Finally, $\mathbb{E}[\max_i X_i \mid \mathcal{E}^*] \le \mathbb{E}[\max_i X_i] = \mathbb{E}[\mathcal{D}_{n:n}]$, as $\mathcal{E}^*$ is an event which upper bounds X_i. Putting everything together:

$$\begin{aligned}
&\mathbb{E}[\max_i X_i \mid \overline{\mathcal{E}_1 \cap \mathcal{E}'_2 \cap \mathcal{E}'_3} \cap \mathcal{E}^*] \cdot \Pr[\overline{\mathcal{E}_1 \cap \mathcal{E}'_2 \cap \mathcal{E}'_3} \mid \mathcal{E}^*] \\
&\qquad \le \mathbb{E}[\max_i X_i \mid \overline{\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3} \cap \mathcal{E}^*] \cdot \Pr[\overline{\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3} \mid \mathcal{E}^*] \\
&\qquad \le \frac{3}{\ln n} \cdot \mathbb{E}[\mathcal{D}_{n:n}] \\
&\qquad \le^{\text{(Lemma 36)}} \frac{3}{\ln n} \cdot (\ln n + 1)\mathbb{E}[\mathcal{D}] \\
&\qquad \le 4\mathbb{E}[\mathcal{D}].
\end{aligned}$$

Next, we will upper bound the contribution of outcomes in $\mathcal{E}_1 \cap \mathcal{E}'_2 \cap \mathcal{E}'_3 \cap \mathcal{E}^*$ to the expected reward of Linear_γ. Note that in such outcomes, for every $c_1 \ge \theta^*$ and $c_2 < \theta^*$, $\mathsf{LinearFixed}_{c_1}$ does not choose a small noise box (Lemma 41) and $\mathsf{LinearFixed}_{c_2}$ chooses some large noise box (Lemma 42). Hence, in such outcomes, Linear_γ does not choose a small noise box. Therefore, in such an outcome, the reward of Linear_γ is at most the reward of an optimal policy that knows $\mathcal{D}$, but is conditioned not to pick a small noise box. When selecting box i, such a policy has expected reward $\mathbb{E}[X_i \mid Y_i = y_i, \mathcal{E}^*, \mathcal{E}_1, \mathcal{E}'_2, \mathcal{E}'_3]$. We first observe that $\mathbb{E}[X_i \mid Y_i = y_i, \mathcal{E}^*, \mathcal{E}_1, \mathcal{E}'_2, \mathcal{E}'_3] = \mathbb{E}[X_i \mid Y_i = y_i, \mathcal{E}'_2, \mathcal{E}'_3]$ as $\mathcal{E}_1$ regards ϵ_j of all small noise boxes j, which are never picked in this policy. Secondly, $\mathbb{E}[X_i \mid Y_i = y_i, \mathcal{E}'_2, \mathcal{E}'_3] = \mathbb{E}[X_i \mid Y_i = y_i, \mathcal{E}^*]$ as X_i is independent of Y_j, for $j \ne i$, and $\mathcal{E}'_2 \cap \mathcal{E}'_3$ have less information about Y_i than $\{Y_i = y_i\}$.

Let $R_i(y_i) = \mathbb{E}[X_i \mid Y_i = y_i, \mathcal{E}^*]$. The reward of an optimal policy which knows $\mathcal{D}$ and is conditioned not to pick a small noise box is then

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{y}}\left[\max_{i \in \{1\} \cup [n - c_b + 1, n]} R_i(y_i) \mid \mathcal{E}_1 \cap \mathcal{E}'_2 \cap \mathcal{E}'_3 \cap \mathcal{E}^*\right] \\
&\quad \le^{(R_1(y_1) \ge 0)} \mathbb{E}_{\boldsymbol{y}}\left[R_1(y_1) + \max_{i \in [n - c_b + 1, n]} R_i(y_i) \mid \mathcal{E}_1 \cap \mathcal{E}'_2 \cap \mathcal{E}'_3 \cap \mathcal{E}^*\right] \\
&\quad =^{(\sigma_1 = 0)} \mathbb{E}[X_1 \mid \mathcal{E}_1 \cap \mathcal{E}'_2 \cap \mathcal{E}'_3 \cap \mathcal{E}^*] + \mathbb{E}_{\boldsymbol{y}}\left[\max_{i \in [n - c_b + 1, n]} R_i(y_i) \mid \mathcal{E}_1 \cap \mathcal{E}'_2 \cap \mathcal{E}'_3 \cap \mathcal{E}^*\right]
\end{aligned}$$

$$= \mathbb{E}[X_1 \mid \mathcal{E}^*] + \mathbb{E}_{\boldsymbol{y}}\left[\max_{i\in[n-c_b+1,n]} R_i(y_i) \mid \mathcal{E}_1 \cap \mathcal{E}_2' \cap \mathcal{E}_3' \cap \mathcal{E}^*\right],$$

where the last inequality holds since $\mathcal{E}_1$, $\mathcal{E}_2$, and $\mathcal{E}_3$ are events regarding small noise and large noise boxes, and hence is independent of X_1.

Consider any small noise box i. Let $\overline{X}_i = X_i \mid X_i \le \alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})}$. Then, conditioned on $\mathcal{E}_1 \cap \mathcal{E}_2' \cap \mathcal{E}_3' \cap \mathcal{E}^*$, for any realization of $\boldsymbol{y}$, we note that $R_i(y_i) = \mathbb{E}[X_i \mid Y_i = y_i, \mathcal{E}^*] = \mathbb{E}[\overline{X}_i \mid \overline{X}_i + \mathcal{N}(0,\sigma_i^2) = y_i]$. Furthermore, as y_i is a realization conditioned on $\mathcal{E}_1 \cap \mathcal{E}_2' \cap \mathcal{E}_3' \cap \mathcal{E}^*$, we have $y_i \le 18\alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})} \ln n$. Using Lemma 24 with $V = \beta_{n^2}^{(\mathcal{D}_{n:n})}$ and $\sigma = \sigma_b = 6\alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})}\sqrt{\ln n}$, we have $\mathbb{E}[\overline{X}_i \mid \overline{X}_i + \mathcal{N}(0,\sigma_i^2) = y_i] \le 2\mathbb{E}[\overline{X}_i] \le 2\mathbb{E}[X_i] = 2\mathbb{E}[\mathcal{D}]$. As this is true for any small noise box i on any realization of $\boldsymbol{y}$, we then have

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{y}}\left[\max_{i\in\{1\}\cup[n-c_b+1,n]} R_i(y_i) \mid \mathcal{E}_1 \cap \mathcal{E}_2' \cap \mathcal{E}_3' \cap \mathcal{E}^*\right] \\
&\le \mathbb{E}[X_1 \mid \mathcal{E}^*] + \mathbb{E}_{\boldsymbol{y}}\left[\max_{i\in[n-c_b+1,n]} R_i(y_i) \mid \mathcal{E}_1 \cap \mathcal{E}_2' \cap \mathcal{E}_3' \cap \mathcal{E}^*\right] \\
&\le \mathbb{E}[X_1 \mid X_1 \le \alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})}] + \mathbb{E}_{\boldsymbol{y}}[2\mathbb{E}[\mathcal{D}]] \\
&\le \mathbb{E}[X_1] + 2\mathbb{E}[\mathcal{D}] \\
&= 3\mathbb{E}[\mathcal{D}].
\end{aligned}$$

Overall, conditioned on $\mathcal{E}^*$, if $\mathcal{E}_1 \cap \mathcal{E}_2' \cap \mathcal{E}_3'$ occurs, Naive's expected reward is at most $3\mathbb{E}[\mathcal{D}]$, while otherwise, the contribution to the expected reward is at most $4\mathbb{E}[\mathcal{D}]$. Therefore, the reward of Naive conditioned on $\mathcal{E}^*$ is at most $7\mathbb{E}[\mathcal{D}]$. ■

With Lemma 29 at hand, we can prove Lemma 28.

Proof of Lemma 28 We decompose $\mathcal{E}^*$ as $\mathcal{E}_1^* \cap \mathcal{E}_2^*$, where $\mathcal{E}_1^*$ and $\mathcal{E}_2^*$ are two independent events defined as follows. $\mathcal{E}_1^*$ is the event that $X_i \le \alpha_{n^{1/10000}}^{(\mathcal{D}_{c_s:c_s})}$ for all small noise boxes $i \in [2, c_s+1]$. $\mathcal{E}_2^*$ is the event that $X_j \le \alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})}$ for all remaining boxes j.

Observe that $\Pr[\overline{\mathcal{E}_1^*}] = \Pr\left[\max_{i\in[2,c_s+1]} X_i > \alpha_{n^{1/10000}}^{(\mathcal{D}_{c_s:c_s})}\right] = \Pr[\mathcal{D}_{c_s:c_s} > \alpha_{n^{1/10000}}^{(\mathcal{D}_{c_s:c_s})}] = \frac{1}{n^{1/10000}}$. Similarly, $\Pr[\overline{\mathcal{E}_2^*}] = \frac{1}{n^{1/10000}}$. Therefore, $\Pr[\overline{\mathcal{E}^*}] = \Pr[\overline{\mathcal{E}_1^*} \cup \overline{\mathcal{E}_2^*}] \le \Pr[\overline{\mathcal{E}_1^*}] + \Pr[\overline{\mathcal{E}_2^*}] = \frac{2}{n^{1/10000}}$.

Next, we upper bound the contribution of $\overline{\mathcal{E}^*}$ to the overall reward of Linear_γ. Overloading notation, let $R_{\mathsf{Linear}_\gamma}(\mathcal{D}, \boldsymbol{\sigma}^* \mid \overline{\mathcal{E}^*})$ be the expected reward of Linear_γ when $\overline{\mathcal{E}^*}$occurs. Then, we have

$$\begin{aligned}
&R_{\mathsf{Linear}_\gamma}(\mathcal{D}, \boldsymbol{\sigma}^* \mid \overline{\mathcal{E}^*}) \cdot \Pr[\overline{\mathcal{E}^*}] \le \mathbb{E}[\max_i X_i \mid \overline{\mathcal{E}_1^*} \cup \overline{\mathcal{E}_2^*}] \cdot \Pr[\overline{\mathcal{E}_1^*} \cup \overline{\mathcal{E}_2^*}] \\
&\le \left(\mathbb{E}\left[\max_{i\in[2,c_s+1]} X_i \mid \overline{\mathcal{E}_1^*} \cup \overline{\mathcal{E}_2^*}\right] + \mathbb{E}\left[\max_{i\in[1,n]\setminus[2,c_s+1]} X_i \mid \overline{\mathcal{E}_1^*} \cup \overline{\mathcal{E}_2^*}\right]\right) \cdot \Pr[\overline{\mathcal{E}_1^*} \cup \overline{\mathcal{E}_2^*}] \\
&= \left(\mathbb{E}\left[\max_{i\in[2,c_s+1]} X_i \mid \overline{\mathcal{E}_1^*}\right] + \mathbb{E}\left[\max_{i\in[1,n]\setminus[2,c_s+1]} X_i \mid \overline{\mathcal{E}_2^*}\right]\right) \cdot \Pr[\overline{\mathcal{E}_1^*} \cup \overline{\mathcal{E}_2^*}] \\
&= \left(\mathbb{E}\left[\mathcal{D}_{c_s:c_s} \mid \mathcal{D}_{c_s:c_s} > \alpha_{n^{1/10000}}^{(\mathcal{D}_{c_s:c_s})}\right] + \mathbb{E}\left[\mathcal{D}_{n-c_s:n-c_s} \mid \mathcal{D}_{n-c_s:n-c_s} > \alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})}\right]\right) \cdot \Pr[\overline{\mathcal{E}_1^*} \cup \overline{\mathcal{E}_2^*}] \\
&\le 2\left(\frac{\mathbb{E}\left[\mathcal{D}_{c_s:c_s} \mid \mathcal{D}_{c_s:c_s} > \alpha_{n^{1/10000}}^{(\mathcal{D}_{c_s:c_s})}\right]}{n^{1/10000}} + \frac{\mathbb{E}\left[\mathcal{D}_{n-c_s:n-c_s} \mid \mathcal{D}_{n-c_s:n-c_s} > \alpha_{n^{1/10000}}^{(\mathcal{D}_{n-c_s:n-c_s})}\right]}{n^{1/10000}}\right).
\end{aligned}$$

Note that $\frac{\mathbb{E}\left[\mathcal{D}_{c_s:c_s} | \mathcal{D}_{c_s:c_s} > \alpha^{(\mathcal{D}_{c_s:c_s})}_{n^{1/10000}}\right]}{n^{1/10000}} = \mathbb{E}\left[\mathcal{D}_{c_s:c_s} \mid \mathcal{D}_{c_s:c_s} > \alpha^{(\mathcal{D}_{c_s:c_s})}_{n^{1/10000}}\right] \cdot \Pr\left[\mathcal{D}_{c_s:c_s} > \alpha^{(\mathcal{D}_{c_s:c_s})}_{n^{1/10000}}\right]$, and similarly for the second term. Applying Lemma 38 here (noting that $\mathcal{D}_{a:a}$ is MHR for all $a \geq 1$; see Lemma 31), we have:

$$\frac{\mathbb{E}\left[\mathcal{D}_{c_s:c_s} \mid \mathcal{D}_{c_s:c_s} > \alpha^{(\mathcal{D}_{c_s:c_s})}_{n^{1/10000}}\right]}{n^{1/10000}} \leq \frac{15(\ln(n^{1/10000}) + \ln(c_s) + 1)}{2n^{1/10000}}\mathbb{E}[\mathcal{D}] \leq \frac{\mathbb{E}[\mathcal{D}]}{4}.$$

Similarly, $\frac{\mathbb{E}\left[\mathcal{D}_{n-c_s:n-c_s} | \mathcal{D}_{n-c_s:n-c_s} > \alpha^{(\mathcal{D}_{n-c_s:n-c_s})}_{n^{1/10000}}\right]}{n^{1/10000}} \leq \frac{\mathbb{E}[\mathcal{D}]}{4}$, for an overall bound of $R_{\mathsf{Linear}_\gamma}(\mathcal{D}, \boldsymbol{\sigma}^* \mid \overline{\mathcal{E}^*}) \cdot \Pr[\overline{\mathcal{E}^*}] \leq 2(\frac{\mathbb{E}[\mathcal{D}]}{4} + \frac{\mathbb{E}[\mathcal{D}]}{4}) = \mathbb{E}[\mathcal{D}]$. Putting everything together, we have

$$\begin{aligned} R_{\mathsf{Linear}_\gamma}(\mathcal{D}, \boldsymbol{\sigma}^*) &= R_{\mathsf{Linear}_\gamma}(\mathcal{D}, \boldsymbol{\sigma}^* \mid \mathcal{E}^*) \cdot \Pr[\mathcal{E}^*] + R_{\mathsf{Linear}_\gamma}(\mathcal{D}, \boldsymbol{\sigma}^* \mid \overline{\mathcal{E}^*}) \cdot \Pr[\overline{\mathcal{E}^*}] \\ &\leq^{\text{(Lemma 29)}} 7\mathbb{E}[\mathcal{D}] + \mathbb{E}[\mathcal{D}] \\ &= 8\mathbb{E}[\mathcal{D}]. \end{aligned}$$

■