IID Prophet Inequality with Random Horizon: Going Beyond Increasing Hazard Rates

Giordano Giambartolomei[†], Frederik Mallmann-Trenn[†] and Raimundo Saona[††]

Abstract

Prophet inequalities are a central object of study in optimal stopping theory. In the *iid* model, a gambler sees values in an online fashion, sampled independently from a given distribution. Upon observing each value, the gambler either accepts it as a reward or irrevocably rejects it and proceeds to observe the next value. The goal of the gambler, who cannot see the future, is maximising the expected value of the reward while competing against the expectation of a prophet (the offline maximum). In other words, one seeks to maximise the gambler-to-prophet ratio of the expectations.

This model has been studied with infinite, finite and unknown number of values. When the gambler faces a random number of values, the model is said to have random horizon. We consider the model in which the gambler is given *a priori* knowledge of the horizon's distribution. Alijani et al. (2020) designed a single-threshold algorithm achieving a ratio of $1/2$ when the random horizon has an increasing hazard rate and is independent of the values. We prove that with a single threshold, a ratio of $1/2$ is actually achievable for several larger classes of horizon distributions, with the largest being known as the $\mathcal{G}$ class in reliability theory. Moreover, we show that this does not extend to its dual, the $\overline{\mathcal{G}}$ class (which includes the decreasing hazard rate class), while it can be extended to low-variance horizons. Finally, we construct the first example of a family of horizons, for which multiple thresholds are necessary to achieve a nonzero ratio. We establish that the Secretary Problem optimal stopping rule provides one such algorithm, paving the way towards the study of the model beyond single-threshold algorithms.

[†]King's College London.
[††]Institute of Science and Technology Austria.

Contents

1   Introduction

Prophet inequalities are a central object of study in optimal stopping theory. A gambler sees nonnegative values in an online fashion, sampled from an instance of independent random variables $\{X_i\}$ with known distributions $\{\mathcal{D}_i\}$, in adversarial, random or selected order, depending on the particular model. When observing each value, the gambler either accepts it as a reward, or irrevocably rejects it and proceeds with observing the next value. The goal of the gambler, who cannot see the future, is to maximise the expected value of the reward while competing against the expectation of a prophet (out of metaphor, the offline maximum or supremum, depending on whether the instance is finite or not). In other words, one seeks to maximise the gambler-to-prophet ratio of the expectations.

The gambler represents any online decision maker, such as an algorithm or stopping rule. Probabilistically, we will refer to it as a *stopping time* τ. The term *online* implies that the gambler, unable to see the future, will always stop at a time τ such that the event $\{\tau = i\}$ depends, informally speaking, only on the first i values observed.

Due to the online nature of prophet inequalities, some terminology from *competitive analysis* is usually borrowed. Somewhat informally, it is common to refer to a worst-case gambler-to-prophet ratio of the algorithm (that is a ratio known to be achievable for any given instance by the algorithm) as *competitive ratio*. Let $c \in [0, 1]$ be a real constant. An algorithm is said to be c-competitive (or equivalently, a $1/c$-approximation) if it has a competitive ratio of c. An upper bound on any algorithm's highest possible competitive ratio for a given instance will be called *hardness* of the instance. Saying that a prophet inequality model has a hardness of c (or equivalently is *c-hard*), means that there is a c-hard instance for that problem. A hardness for a model is said to be *tight* when it is matched by the competitive ratio of an algorithm solving it. Similarly, the competitive ratio c of an algorithm solving a model is tight when it is matched by a hardness of c for that model (equivalently, the algorithm is a tight $1/c$-approximation).

1.1   Prophet inequality models

To date, several models and extensions of prophet inequalities are present in the literature. In this section we introduce the classical model and describe two variants of it, briefly reviewing the state-of-the-art concerning their hardness and the competitiveness of known algorithms.

Classical Prophet Inequality.   The very first prophet inequality model, typically referred to as classical (or adversarial) Prophet Inequality (PI), is due to Krengel and Sucheston [31, 32]. The given instance is composed of countably (possibly infinitely) many integrable independent nonnegative random variables $\{X_i\}$ with known distributions $\{\mathcal{D}_i\}$ in a fixed given order, usually referred to as *adversarial* order. As per the general dynamics previously described, the gambler observes in an online fashion the sequence of sampled values in the fixed order. Formally, the goal is to maximise the gambler-to-prophet ratio $\mathbb{E}X_\tau$ over $\mathbb{E}\sup_i X_i$. When working with *infinite* instances, it is customary to consider only finite stopping rules and and variables such that $\mathbb{E}\sup_i X_i < \infty$.[1] In [32], it was shown that the $1/2$-hardness of PI (shown by Garlin [31]) is tight. Later, a single-threshold 2-approximation was constructed in [41].

IID Prophet Inequality.   Denoted as IID PI, it is a specialised model of PI, where the random variables $\{X_i\}$ are further assumed independent and identically distributed (*iid*) according to a given distribution $\mathcal{D}$. The hardness of this problem has been shown to be $1/\beta \approx 0.7451$, where $\beta \in \mathbb{R}$ is the unique solution to $\int_0^1 \frac{dx}{x(1-\log x)+\beta-1} = 1$ in [27]. The β-approximation *quantile strategy* devised in [13] shows that the aforementioned hardness is tight.

---

[1]Formally, we maximise the gambler-to-prophet ratio with respect to all stopping times τ such that $\mathbb{P}(\tau < \infty) = 1$.

IID Prophet Inequality with Random Horizon.   The IID PI with Random Horizon (RH) was first introduced in [22]. Consider a random variable H, which we will call the (random) *horizon*, with given discrete distribution $\mathcal{H}$, that is, supported on an arbitrary subset of $\mathbb{N}$. This assumption comes with no loss of generality (formally, it rules out that the horizon has mass at zero). H will be assumed finite ($\mathbb{P}(H < \infty) = 1$) and integrable ($\mathbb{E}H < \infty$), which we denote $H \in \mathcal{L}^1(\Omega)$. RH is a relaxation of the IID PI, considering integrable *iid* random variables $X_1, \ldots, X_H$, with given distribution $\mathcal{D}$ independent of $\mathcal{H}$. This setup, supported on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, models the game where the gambler, facing an unknown number of values $X_1, \ldots, X_H$, can still use *a priori* knowledge of $\mathcal{H}$ when maximising returns. More precisely, denote $[n] = \{1, \ldots, n\}$. The goal is to maximise the gambler-to-prophet ratio of $\mathbb{E}X_\tau$ over $\mathbb{E}M_H$, where $M_H := \max_{i \in [H]} X_i$ and both expectations run over the randomness of the *iid* copies of $X \sim \mathcal{D}$ and $H \sim \mathcal{H}$. The gambler must select the value in ignorance of whether it is the last or not. If the gambler fails to stop by the time the last value has been inspected, the return is zero. This ignorance, added to the usual nonanticipative constraint, yields a no easier model than the IID PI. On top of the usual disadvantage of playing against a prophet, which knows all future realisations of the values, and chooses the largest; the gambler now competes against a prophet, which can also foresee the random number of values. We leave more in depth measure-theoretic details of the model for Section 2.

1.2   Previous bounds on the IID Prophet Inequality with Random Horizon

No constant-approximations are possible for RH [2, Theorem 1.6]. Distributional knowledge of the horizon is not enough to guarantee, in expectation, that the gambler's return can attain a worst-case constant multiple of the prophet's return. Therefore, we impose additional restrictions on the distribution of the horizon. The largest distributional class for which a constant-approximation has been found so far is that of horizons with increasing hazard rates $\lambda(h)$. For simplicity, we will use *increasing* and *decreasing* instead of *nondecreasing* and *nonincreasing* respectively. Recall that the hazard rate $\lambda(h)$ of a horizon H is defined as follows. Denote the survival function $S(h) = \mathbb{P}(H \geq h)$. If $|\operatorname{supp}(H)| = \infty$, for every $h \in \mathbb{N}$, $\lambda(h) := \mathbb{P}(H = h)/S(h)$. If $|\operatorname{supp}(H)| < \infty$, for every $h \in \mathbb{N}$, $\lambda(h)$ is defined analogously for all $h \leq \sup \operatorname{supp}(H)$, whereas it is set to 1 for all $h > \sup \operatorname{supp}(H)$.

Definition 1.1 (IHR class)**.** *H is* Increasing Hazard Rate (IHR) *if for every $h \in \mathbb{N}$, $\lambda(h) \leq \lambda(h + 1)$.*

The *dual* class is defined by reversing the inequality on the support of the horizon $\operatorname{supp}(H)$.

Definition 1.2 (DHR class)**.** *H is* Decreasing Hazard Rate (DHR) *if for every $h \geq \inf \operatorname{supp}(H)$, $\lambda(h) \geq \lambda(h + 1)$.*

The geometric distribution is the only discrete distribution that has constant hazard rate, and is therefore the only discrete distribution that belongs to both classes.

The state-of-the-art for RH consists in the findings of [2]. In [2, Theorem 3.2], through the use of *second order stochastic dominance*, it is shown that if $H \in$ IHR, $\mathbb{E}X_\tau \geq (2 - 1/\mu)^{-1}\mathbb{E}M_H$, where $\mu := \mathbb{E}H \geq 1$. This ensures a (uniform) 2-approximation on the IHR class.[2] Remarkably, this is achieved by a single-threshold algorithm [2, Theorem 3.1], and it is tight [2, Theorem 3.5], since if H is geometrically distributed, a parametric family of two-point distributions can be shown to make the prophet-to-gambler ratio approach $2 - 1/\mu$, as the parameters vary suitably. The emphasis on monotone hazard rate classes stems from [22], which is the first work that considers RH, to the best of our knowledge.

---

[2]For brevity, outside formal statements, we will often omit the dependence on distributional parameters when referring to competitive ratios. For example, in this sentence 2-*approximation* means, more precisely, $2 - 1/\mu$-approximations.

1.3 Our contributions

Implicitly, the motivation behind the approach that considers monotone hazard rate classes stems from interpreting RH from a reliability theory point of view: an optimal stopping game under evolving risk (or ageing) that it ends by the next step. In reliability theory several classes of distribution are used to model this aspect. However, despite more than 15 years have passed, no progress has been made on classes larger than the IHR class. In this paper we contribute to the study of RH by:

- Extending existing characterisations of the optimal algorithm to unbounded horizons.
- Complementing and extending key ideas from [2] to yield the existence of single-threshold 2-approximations on important superclasses of the IHR class, culminating in the $\mathcal{G}$ class.
- Showing that the $\overline{\mathcal{G}}$ class (dual of the $\mathcal{G}$ class) admits hard horizons for single-threshold algorithms..
- Deriving single-threshold constant-approximations for sufficiently concentrated horizons with finite second moments.
- Showing that RH admits families of horizons for which single-threshold algorithms are not competitive, despite the Secretary Problem (SP) optimal stopping rule providing a constant-approximation (see Section 1.5 for details on SP and its variants).

In this section we give formal statements of the above results, with the exception of Theorem 2.1 in Section 2, due to the overall degree of technicalities involved. Informally speaking, this theorem characterises the optimal algorithm in terms of a discounted infinite optimal stopping problem, with a special focus on unbounded horizons. As a technical tool, the equivalence of stopping rules for RH with stopping rules for the discounted problem is leveraged not only to obtain hardness results for single-threshold algorithms, but also to formalise heuristic arguments involving the optimal algorithm under geometric horizon. More specifically: we exploit the equivalence between stopping rules for RH and corresponding rules for the discounted problem to extend hardness to neighbouring instances; we show that with a geometric horizon, if the value distribution is bounded, the optimal algorithm of RH is single-threshold, pinpointing the structural equation for the threshold.

The next three results involve a stochastic order, which we now introduce. Standard facts regarding stochastic orders are often relied upon throughout this work. They will be recalled as remarks, whose proof can be found in [45]. In the following definitions, H and G always refer to horizons.

Definition 1.3 (Probability Generating Function Order). *Given two horizons H, G, we say that G is dominated by H in the probability generating function (pgf) order, and denote it as $G \prec_{pgf} H$, if for every $t \in (0,1)$, $\mathbb{E}\left(t^G\right) \geq \mathbb{E}\left(t^H\right)$.*

Remark 1.1. *$G \prec_{pgf} H$ if and only if for all $t \in (0,1)$, $\sum_{i=1}^{\infty} S_G(i)t^i \leq \sum_{i=1}^{\infty} S_H(i)t^i$.*

The $\mathcal{G}$ class, consisting of every positive discrete distribution that dominates, in the pgf order, the ($\mathbb{N}$-valued) geometric distribution with the same mean, has been introduced in [30].

Definition 1.4 ($\mathcal{G}$ class). *The $\mathcal{G}$ class consists of every horizon that dominates, in the pgf order, the geometric distribution with the same mean: $\mathcal{G} := \{H : G \prec_{pgf} H, G \sim \text{Geom}(1/\mathbb{E}H)\}$.*

The dual class is obtained by reversing the pgf ordering, and is denoted as $\overline{\mathcal{G}}$. It is well known that IHR $\subset \mathcal{G}$, with the inclusion being strict. Similarly, DHR $\subset \overline{\mathcal{G}}$. In applications of reliability theory, the classes within $\mathcal{G}$ and its dual are a staple of modelling many aspects of ageing and more in general of an evolving risk of failure.

Theorem 1.1. *For every $H \in \mathcal{G}$ with $\mathbb{E}H =: \mu$, a single-threshold $2 - 1/\mu$-approximation exists.*

From [2, Theorem 3.5] it follows that any 2-approximation on the $\mathcal{G}$ class is tight. More specifically, the 2-approximation yielding Theorem 1.1 is essentially[3] the same as the one used on the IHR class in [2, Theorem 3.1]. The key feature of the stopping rule is determining the value p such that $\mathbb{P}(X > p) = 1/\mu$, and then accepting the first value exceeding p. The $\mathcal{G}$ class may be somewhat abstract, compared to its subclasses having immediate intuitive interpretations in terms of ageing concepts, such as the IHR class. Its possible interpretations are quite general and go beyond mere ageing, as can be read in [30]. To gain some intuition of how much more general of a class it is, the reader is referred to Section 3, where we will describe some of its largest most well-known subclasses, as evidence that Theorem 1.1 generalises significantly [2, Theorem 3.1, Theorem 3.2].[4] We conclude by mentioning that statistical tests for the $\mathcal{L}$ class (the continuous equivalent of the $\mathcal{G}$ class) have also been developed [29]. The possibility of adapting them into tests for the $\mathcal{G}$ class adds to the practical significance of the result.

Our next result requires introducing the stopping rule for SP with deterministic horizon m. This consists of rejecting the first $r - 1$ values, and subsequently accepting the first value ranked better than all the previous ones. The waiting time $r_m \sim m/e$ as $m \longrightarrow \infty$ yields the optimal stopping rule (see Section 1.5 for details). In the following, we provide the first example of a family of horizons, for which no single-threshold constant-approximation is possible, and yet an adaptation of the SP stopping rule to random horizons is proven to provide a constant-approximation. The family of horizons we will consider is of the form $\{H_m\}_{m \geq M}$, with pmf's parametrised by $\varepsilon > 0$, $M \in \mathbb{N}$ large enough and $m := \sup \mathrm{supp}(H_m) < \infty$, and will therefore be denoted as $\mathcal{H}_M(\varepsilon)$. Let ζ_{r_m} be the (optimal) stopping rule for SP with deterministic horizon m. Taking the minimum $\tau_m := \zeta_{r_m} \wedge (H_m + 1)$ produces a natural adaptation of the SP stopping rule to the random horizon H_m.

Theorem 1.2. *There exists a family of horizons $\mathcal{H}_M(\varepsilon)$, such that for all fixed ε small enough and arbitrary M, no single-threshold constant-approximation is possible. Nonetheless, for every fixed ε, no matter how small, the random horizon SP rule τ_m is an approximate constant-approximation, as long as $M = M(\varepsilon)$ is fixed large enough.*

For brevity, we say that the family of horizons $\mathcal{H}_M(\varepsilon)$ is *hard* for single-threshold algorithms, but the random horizon SP rule is *competitive* on it.[5] Considering the extensive number of classes of distributions for which single-threshold algorithms are competitive, this result is quite surprising, especially because the SP stopping rule only needs information regarding the relative ranks of the values, not the values themselves.

Since we showed that single-threshold 2-approximations exist for the $\mathcal{G}$ class, finding whether single-threshold constant-approximations exist for the $\overline{\mathcal{G}}$ class turns into an interesting problem.

---

[3]The argument and the algorithm provided in [2], although valuable and insightful, are somewhat informal and require some tweaks and filling some gaps. A more in depth discussion is given in Section 3, where we show that RH is amenable enough to allow for stochastic tie-breaking, a standard techniques which combines well with the aforementioned tweaks.

[4]It may be beneficial to give a concrete example of a distribution that, although elementary, could not be handled without our extension. Let H with $S(h)$ given as $1, 1/2, 1/4, 1/5$ for $h = 1, 2, 3, 4$ respectively and 0 for all $h > 4$. The list of the corresponding hazard rates $\lambda(h)$, $1/2, 1/2, 1/5, 1, \ldots$, is not monotone increasing because $\lambda(3)$ drops. The $\mathcal{G}$-class condition is, however, satisfied, thus our result can ensure a 2-approximation (to be specific, a $2 - 20/39 \approx 1.49$-approximation). It is straightforward to compute $\mathbb{E}H = 39/20$, so Definition 1.4 requires $\mathbb{E}t^G \geq \mathbb{E}t^H$ for all $0 < t < 1$, with $G \sim \mathrm{Geom}(20/39)$, and the two pgf's being respectively $20t/(39 - 19t)$ and $t/2 + t^2/4 + t^3/20 + t^5/5$.

[5]We omit the qualifier *approximate* appearing in the statement of Theorem 1.2, a technicality which we now clarify. *Approximate* refers to the fact that the competitive ratio ensured, which is a positive, increasing function of ε, denoted as $g(\varepsilon)$, is approached as $M \longrightarrow \infty$. At the large ε regime, $g(\varepsilon) \approx 1/e^2$. As ε grows, the family of horizons $\mathcal{H}_M(\varepsilon)$ admits constant-competitive single-threshold algorithms too. However, at small ε regime, these algorithms fail, whereas τ_m guarantees approximately a $g(\varepsilon)$ fraction of $\mathbb{E}M_{H_m}$. Specifically, for all $m \geq M$, a ratio of $g(\varepsilon) - \delta$ is guaranteed, where $\delta = \delta(M) \longrightarrow 0^+$ as $M \longrightarrow \infty$. Hence the dependence $M = M(\varepsilon)$ with $M(\varepsilon)$ growing large as ε vanishes.

The equivalent characterisation of stopping rules for RH in terms of a discounted stopping problem, combined with a perturbation argument on $\mathcal{H}_M(\varepsilon)$, yields a negative answer.

Theorem 1.3. *No single-threshold constant-approximation is possible on the $\overline{\mathcal{G}}$ class.*

In Section 4.2 we will describe more in detail some of the largest most well-known subclasses of the $\overline{\mathcal{G}}$ class. They are no less important than the corresponding subclasses of the $\mathcal{G}$ class. For example, discrete DHR distributions are crucial in several areas: they have been shown to describe well the number of seasons a show is run before it gets cancelled and the number of periods until failure of a device (a system, or a component, that functions until first failure) governed by a continuous DHR life distribution in the grouped data case [33]. Theorem 1.3 motivates our conjecture that single-threshold 2-approximations are not possible on the DHR class either. Since our characterisation result of stopping rules for RH extends to unbounded horizons, it will be a crucial tool in proving this conjecture within our framework, as DHR horizons have support of the form $[a, \infty) \cap \mathbb{N}$. Furthermore, suitably combining the techniques yielding Theorems 1.1 and 1.3 with the fact that the tight geometric instance for 2-approximations belongs to the $\mathcal{G}$ class, motivates our second conjecture that the $\mathcal{G}$ class is essentially optimal for 2-approximations on $\mathcal{L}^1$ horizons.

In our last result, we go back to single-threshold algorithms, but consider only horizons H satisfying th e sole condition $\mathbb{E}(H^2) < \infty$, denoted $H \in \mathcal{L}^2(\Omega)$. Here no particular information is available regarding the classification of the distribution in terms of the $\mathcal{G}$. We show that it is possible to exploit concentration bounds, in order to ensure that the same single-threshold algorithm used for the $\mathcal{G}$ is still a 2-approximation. In the following result, W_0 denotes the principal branch of the Lambert function.

Theorem 1.4. *For every H having both $\mu := \mathbb{E}H$ and $\sigma^2 := \operatorname{Var} H$ finite, such that*

$$\sigma^2 \leq \mu^2 \left[1 - \frac{1}{\mu} + W_0 \left(-\left(2 - \frac{1}{\mu}\right) e^{-\left(2 - \frac{1}{\mu}\right)} \right) \right], \tag{1.1}$$

a single-threshold $2 - 1/\mu$-approximation exists.

The Lambert function is readily simulated in many languages, with well-studied error bounds. Therefore, one can reliably compute numerically the magnitude of concentration required by (1.1), so that a 2-approximation is ensured. To exemplify, consider the large-market limit, that is $\mu \longrightarrow \infty$, which is both interesting for practical applications involving horizons with large expectation, and simplifies computations. Under this hypothesis, taking square-root on both sides of (1.1) yields approximately $\mathrm{CV} := \sigma/\mu \leq 0.770$, where CV is the *coefficient of variation* of the horizon. Variations over this procedure show that both weaker and stronger than 2 constant-approximations are possible for all suitably concentrated horizons. Consider again the upper bound on CV provided by square-rooting (1.1) in the large-market limit: $\mathrm{CV} \leq \sqrt{1 + W_0(-2e^{-2})}$. Replacing 2 with $C > 2$ yields $\mathrm{CV} \leq \sqrt{C - 1 + W_0(-Ce^{-C})}$. This will ensure a large-market limit weaker C-approximation. As $x \geq 1$ grows, $W_0(-xe^{-x})$ is negative strictly increasing (valued -1 at $x = 1$) and vanishing. Thus the upper bound on the CV ensuring a C-approximation for C growing large scales roughly as $\sqrt{C-1}$, meaning that there will be, although progressively weaker, constant-approximations, as long as $\mathrm{CV} < \sqrt{C-1}$. In the other direction, stronger C-approximations are ensured when taking *admissible* $1 < C < 2$. When reducing C, we clearly cannot go past $e/(e-1)$, where the upper bound on the CV vanishes, meaning that the horizon is required to progressively concentrate until it becomes degenerate. At this point our estimates hit the deterministic optimum for single-threshold IID PI, $1 - 1/e$ [17]. Following similar ideas, the following holds without large-market hypothesis.

Corollary 1.1. *For every H having both $\mu := \mathbb{E}H$ and $\sigma^2 := \operatorname{Var} H$ finite, and every constant $C \geq [1 - (1 - 1/\mu)^\mu]^{-1}$, if*

$$CV \leq \sqrt{C - 1 + W_0(-Ce^{-C})},$$

then a single-threshold C-approximation exists.

At a closer look, it is not surprising that stronger approximations are possible within the class of sufficiently concentrated horizons. Although it is true that the geometric horizon is tight for 2-approximations in $\mathcal{L}^2(\Omega)$ (which is the reason for which Theorem 1.4 is formulated in terms of them), the aforementioned properties of $W_0(-xe^{-x})$ readily imply that no geometric horizon satisfies (1.1).[6]

1.4 Our techniques

The toolbox we assembled, can be described in three key-points.

Discounted infinite problem. The optimal algorithm is characterised via a correspondence between stopping rules for RH and stopping rules for a discounted optimal stopping problem, with the discount factors being the survival function of the horizon. More general optimal stopping problems (unbounded random horizon) do not always admit an optimal algorithm [43]. We show that for RH, this is always possible. This is crucial for geometric and DHR horizons, which are unbounded. The generalisation is the product of an original measure-theoretic construction, leveraging trace σ-algebras.[7]

Stochastic orders. The extension of the 2-approximation from the IHR class to the $\mathcal{G}$ class and $\mathcal{L}^2$ horizons is obtained by exploiting stochastic orders, which have never before been applied to prophet inequality problems. Several distributional classes, starting from the IHR class, can be equivalently defined through some stochastic-ordering with respect to the geometric distribution (as previously seen, the largest of such classes exploits the pgf order). Thus our Theorem 1.1 not only extends to the $\mathcal{G}$ class, but simplifies the arguments of [2]. The proof of Theorem 1.4 leverages extremal properties of specific Bernoulli distributions with respect to another stochastic order, the Laplace transform order (see Section 5 for details), for distributions having finite variance.

Hard instance. Our analysis of the hard instance for single-thresholds algorithms is also novel. The literature on prophet inequalities has focused only on bounded discrete instances that are hard for every algorithm. On the other hand, the instance we provide exploits a continuous Pareto distribution, which, combined with a suitable family of horizons, is both hard for single-threshold algorithms, and competitive for the SP stopping rule, against the offline maximum. This requires a more sophisticated and analytic approach. The phenomenon, for which certain algorithms fail on an instance but others do not, is subtle. To the best of our knowledge, a competitive analysis of the SP stopping rule against the offline maximum with a random horizon is also new (see Section 1.5). It is worth noting that this hard instance is designed to live in a close neighbourhood of the $\overline{\mathcal{G}}$-class, so that perturbing it suitably yields, at the same time, a hard family in the $\overline{\mathcal{G}}$ class.

1.5 Related work

The literature on prophet inequalities models with deterministic horizon is vast. The techniques required for RH differ significantly, and the reader interested in the deterministic setting is referred to surveys such as [14], which contains many recent developments; [35, 15], for an

---

[6]It might be beneficial to give another concrete example of a distribution that, although elementary and sufficiently concentrated for our result to ensure a 2-approximation (to be specific, a 1.89-approximation), could not be handled by previous results. Consider H having Zipf's distribution on [20] with shape $a = 0.3$, that is with probability mass function $f(h) := h^{-a}/z$ for all $h \in [20]$, where $Z := \sum_{h=1}^{20} h^{-a} = 10.9$. The list of hazard rates $\lambda(h)$ is approximately $0.0914, 0.0818, 0.0789, 0.0785, 0.0797, 0.0820, 0.0853, 0.0896, 0.0950, \ldots$, so it is not monotone. Yet the distribution satisfies (1.1), since $\operatorname{Var} H \approx 34.6$ is smaller than the right-hand side of (1.1), which evaluates to approximately 37.0 (all values provided are numerical approximations up to machine precision).

[7]The reader not familiar with measure-theoretic probability should not be discouraged, since the rest of our main proofs leverage only the intuitive meaning of this, which will be made clear.

economic point of view (including applications to Posted Price Mechanisms and more broadly online auctions); [25], for classical results concerning infinite instances. Besides the classical and IID PI, already reviewed in Section 1.1, the other two main models with deterministic horizon we mention here are the Random Order and the Order Selection PI. In the first the random values arrive in a uniform random order (this model has been shown to be 0.688-competitive [9] and 0.7235-hard [19]), in the second the order can be chosen by the gambler (this model has been shown to be 0.7258-competitive [7] and is $1/\beta$-hard, since the IID PI is a particular case). Very recently also Oracle-augmented PI have been introduced [23]. The existing literature strictly related to RH is limited to [2, 22, 36], which explore the topic from the point of view of online automated mechanism design (see [44] for a survey), with a strong emphasis on the multiple-items setting. Although the setting treated in this work is the single-item one, it is crucial, since a standard machinery has already been developed in [2] to extend results for single-item RH to multiple items, as long as the horizons for each item are uniformly bounded. To the best of our knowledge, [36] is the first paper to consider the problem of unknown market size (that is, random horizon with unknown distribution). As an example of practical applications of RH, the original motivation proposed by the authors was that in search keyword auctions such as Google, the supply (search queries that the users enter, resulting in a search results page on which ads can be displayed) is not known in advance, and arrives in an online fashion. This is a multi-unit auction of unknown size for a perishable item (the search queries, who perish within a fraction of a second), to be allocated to buyers (the ads that will be displayed on the search results page). Their analysis provides constant-approximations of both the online and offline optimal single-price algorithm. These results are limited to unknown horizons. This assumption has later been shown to be hard for RH in [22]. In practice, distributional knowledge is available via the history of past transactions. Therefore, in [22] the focus shifted to admitting known horizons. As to obtaining constant-approximations, however, they focus solely on generalising to multiple items the *additive* prophet inequality of [24] for *bounded* value distributions, whereas our main focus is on multiplicative prophet inequalities, in line with [2], already extensively reviewed (see Section 1.2). We conclude this paragraph by mentioning that, on a technical level, RH is also well-recognised as connected to other online stochastic matching problems with random number of queries, which have recently been shown to admit 2-approximations [3]. A potential ground for applications of our techniques would be improving such guarantees.

In the remaining of this section we will focus on another well-known problem in theoretical computer science, the Secretary Problem (SP), extensively studied in the random horizon setting, partially inspiring our approach. For the history and variations on the problem see [21, 18]. The classical SP is as follows: let n be a known number of secretaries for hire. They are interviewed in an online fashion by the employer, all $n!$ possible orderings being equally likely. The employer is able, at the end of each interview, to rank all the secretaries that have been presented so far (*relative ranks*). Each time, the employer must either hire (and stop with the interview), or irrevocably reject the secretary (and proceed with the next interview, except for the last one: the employer will have to hire the last secretary, if the previous $n-1$ were rejected). The objective is maximising the probability of choosing the best secretary, that is the one having *absolute rank* 1. The problem is solved in [34]: the optimal stopping rule consists in rejecting all the first $r-1$ of the n secretaries, and to accept the first secretary, which is better than all previous ones (relative rank 1) thereafter. The number $r = r_n$ can be characterised such that as $n \longrightarrow \infty$, both r/n and the probability of selecting the best (secretary) tend to $1/e \approx 0.3679$. This result is also obtained in [16] through an alternative study of an underlying Markov chain. Consider the following process $\{x_t\}_{t \in [n]_0}$: if the employer enumerates the secretaries in the order in which they are encountered, and then lets $x_0 = 1$, and for $t \in [n]$, lets x_t be the number of the first secretary, which is preferable to the one denoted by x_{t-1}; then we have that $\{x_t\}$ is a Markov chain with state space $[n]$. Potential-based optimal stopping techniques yield the aforementioned results. The Markovian approach is crucial to the analysis

of random horizon SP.

When the number of secretaries is random, it is denoted by N, and the secretaries are interviewed in uniform random order, conditionally on N. At each interview the employer, who has distributional knowledge of N, makes the decision of stopping in ignorance of whether there will be a secretary to interview next. The employer would obtain no reward, in the eventuality that the secretary just rejected was indeed the last. This model was first studied in [39] through a generalisation of the Markovian approach of [16]. The optimal stopping rule is no longer as simple as in the deterministic case, where we had that it is optimal to stop after a waiting time, as soon as a secretary having relative rank 1 is interviewed. This is called one *stopping island* scenario. Depending on the horizon's distribution, we can have a more complicated set Γ composed of multiple islands, in which it is optimal to stop at a secretary with relative rank 1, and outside which it is not. For uniform, geometric and Poisson horizons there still is only one island yielding non-vanishing asymptotics of the probability of choosing the best, but the value varies. In [1] it was later shown that a hard family of horizons $\{N_m\}$ (with $\text{supp}(N_m) = [m]_0$, for every $m \in \mathbb{N}$) could be constructed, meaning that as $m \longrightarrow \infty$, the probability of choosing the best vanishes. In order to recover a constant nonzero limiting value, more information is needed.

SP has also been studied in its *full-information* variant, that is when the common continuous distribution, from which the quality measurement of each of the n secretaries is sampled independently, is known by the employer, who can proceed to directly inspect, at each interview, the secretary's true quality. A first heuristic solution to this problem was given in [21] via a backward induction argument. A rigorous proof of this form of the optimal stopping rule was successively derived in [5] via the Markovian approach. The full-information variant has also been studied with random horizon by [38], where the Markovian approach of [5] is successfully extended. To the best of our knowledge, a hard instance is not yet known for the full-information problem with random horizon.

Both no- and full-information SP have also been studied under the harder hypothesis of *random freeze*, respectively in [42, 43]. In this case, given n secretaries, interviews may have to stop at an independent random time $N \leq n$, with distribution known to the employer. Finally, we mention a variant of SP studied in [10]: instead of seeking to maximise the probability of choosing the best, the employer seeks to minimise the expected rank of the chosen secretary. This problem was extended to bounded random horizons N in [20], and studied for a parametric family of IHR horizons. To the best of our knowledge this is the first work, which investigates broad distributional classes (note that they are defined in terms of hazard rate properties) rather than families of distributions of a given form. More recent developments extend SP to combinatorial structures and unknown horizons [26, 37].

1.6 Organisation of the paper and notation

The paper is organised as follows. In Section 2 we give preliminaries on the formal approach to random horizon optimal stopping problems, and prove the equivalence with discounted optimal stopping problem. In Section 3 we prove Theorem 1.1. In Section 4 we prove Theorem 1.2. In Section 4.2 we prove Theorem 1.3. In Section 5 we prove Theorem 1.4. The Supplementary materials contain a detailed version of all the proofs.

2 Preliminaries

In this section we set up our notation, give finer details regarding the probabilistic model of RH and characterise the optimal algorithm by deriving a useful equivalence between stopping rules for RH and stopping rules of an associated discounted (possibly infinite) stopping problem.

2.1 Notation

We denote $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, $\overline{\mathbb{N}} := \mathbb{N} \cup \{\infty\}$, $\overline{\mathbb{R}} := \mathbb{R} \cup \{\infty, -\infty\}$, $[n] := \{1, 2, \ldots, n\} \subset \mathbb{N}$, $[0] = \{0\}$, $[n]_0 := \{0\} \cup [n]$, $a \wedge b := \min(a, b)$ and $a \vee b := \max(a, b)$ for all $a, b \in \mathbb{R}$. Given nonnegative functions $f(x)$ and $g(x)$ (or equivalently the corresponding sequences $\{f_n\}$ and $\{g_n\}$ if we consider $x = n \in \mathbb{N}$) and a point $a \in \overline{\mathbb{R}}$, assume, for simplicity, that $g(x)$ is strictly positive in a punctured neighbourhood of a (if $a = \infty$, this is understood to mean for all x large enough). Then we denote:

- $f(x) = \mathcal{O}(g(x))$ as $x \longrightarrow a$ if $\limsup_{x \longrightarrow a} \frac{f(x)}{g(x)} < \infty$.
- $f(x) = o(g(x))$ as $x \longrightarrow a$ if $\lim_{x \longrightarrow a} \frac{f(x)}{g(x)} = 0$.
- $f(x) = \Omega(g(x))$ as $x \longrightarrow a$ if $\liminf_{x \longrightarrow a} \frac{f(x)}{g(x)} > 0$.
- $f(x) \sim g(x)$ as $x \longrightarrow a$ if $\lim_{x \longrightarrow a} \frac{f(x)}{g(x)} = 1$.
- $f(x) \asymp g(x)$ as $x \longrightarrow a$ if $f(x) = \mathcal{O}(g(x))$ and $f(x) = \Omega(g(x))$.

Some of the asymptotic notation above, as customary, can also be extended to a function $f(x)$ possibly taking negative values. When this occurs, it is understood that we consider $|f(x)|$ instead. We will adopt the convention of omitting the dependence on all parameters not involved in the main limiting process (which will be clear from the context), regardless of whether uniformity (when not relevant to the context) holds within the parametric range.[8] In evaluating integrals we denote $f(x)|_a^b := f(b) - f(a)$. *Absolutely continuous* distributions will be referred to as *continuous* for brevity. *Discontinuous* distributions are assumed to satisfy the usual regularity assumption (isolated jumps). Convergence in distribution of random variables is denoted as weak convergence: $\xrightarrow{w}$. *Almost surely* is abbreviated a.s. When random variables H, G are identically distributed, it is denoted by $H \sim G$. The essential supremum of a collection of random variables is denoted with the symbol ess sup. $S(h) := \mathbb{P}(H \geq h)$ is taken as the *survival function* of H (when necessary the notation $S_H(h)$ will be used), while we denote $\overline{F}(h) := S(h + 1) = \mathbb{P}(H > h)$. We follow the standard probabilistic functional notation for powers: for example $\mathbb{P}^2(\cdot) := [\mathbb{P}(\cdot)]^2$ and $\mathbb{E}^2(\cdot) := [\mathbb{E}(\cdot)]^2$.

2.2 Probabilistic model

As aforementioned in this section (and in this section alone) RH is studied from the point of view of optimal stopping, hence horizons admitting mass at zero will be also considered.[9] In Section 1.1 we ruled out mass at zero, since in all later sections, concerned with competitive analysis, this will be the standard assumption. A formal convention in fact restores the equivalence with a scenario admitting mass at zero, from a competitive analysis point of view. Add a value $X_0 := 0$, and set $X_\tau := X_0 =: M_{[0]}$ for all $\omega \in \{H = 0\}$. Informally, when the game does not start, both the gambler and the prophet obtain no return. Adopt the convention $0/0 = 1$. Having set $X_0 = 0$ ensures that the game always starts, unless $\omega \in \{H = 0\}$. Assuming absence of penalty for entering the game even though it does not start yields the equivalence: the ratio being 1 does not affect the worst case competitive ratio.

In RH the gambler must select the value in ignorance of whether it is the last or not: at each step $i \in \mathbb{N}_0$ the gambler only learns whether the inspected value x_i is the last or not (that is, whether $\omega \in \{H = i\}$ or $\omega \in \{H > i\}$) *after* the decision of accepting or rejecting it has been made (that is in the $i + 1$st step). Probabilistically, it is standard to model this as follows. The random list $X_1, \ldots, X_H$ is constructed from the infinite underlying process $\mathbf{X} := \{X_i\}$ of *iid* copies of X, whose natural filtration (intuitively, the information associated with the history of

---

[8]For example, the sentence: as $x \longrightarrow \infty$, for every $0 < \varepsilon < 1$, $f(x, \varepsilon) \asymp g(x, \varepsilon)$; means that $\asymp$ refers to x, for every ε fixed; we avoid $\asymp_\varepsilon$, and there is no intended reference to whether, as ε varies in $(0, 1)$, the asymptotic constants hold uniformly or not.

[9]This has been shown to be preferable[43].

the process up to the present) is the sequence of $\sigma(X_1, \ldots, X_i) \coloneqq \mathcal{B}_i$, for every $i \in \mathbb{N}_0$, where $\mathcal{B}_0 \coloneqq \{\emptyset, \Omega\}$ (no information at the start: this is the implicit starting point of all filtration we will introduce). Let $\nu \coloneqq \mathbb{E}X < \infty$. The nonanticipative condition and the ignorance aforementioned are modelled by requiring that $\{\tau = i\} \in \sigma(\mathbb{1}_{\{H=0\}}, X_1, \ldots, \mathbb{1}_{\{H=i-1\}}, X_i) =: \mathcal{F}_i$ for every $i \in \mathbb{N}_0$. Intuitively, this second filtration represents the information associated with the history of the game, which combines both the history of the process of the values, and the state of the game in all preceding steps (that is, whether it ended or it is still running). We recover the original $X_1, \ldots, X_H$ by imposing that the reward is zero if the gambler fails to stop by time H on the infinite sequence, that is defining the reward sequence $\{Y_i\}$, where $Y_i \coloneqq X_i \mathbb{1}_{\{H \geq i\}}$, to which we add $Y_0 \coloneqq X_0$.

For technical reasons, we consider the slightly less common class of all possibly infinite stopping times for the problem $\mathbf{X}(\mathbf{Y})$, denoted as $\mathcal{T}(\mathcal{T}^*)$. Formally, $\tau \in \mathcal{T}(\mathcal{T}^*)$, means that $\{\tau = i\} \in \mathcal{B}_i(\mathcal{F}_i)$ for all $i \in \mathbb{N}$, admitting also $\mathbb{P}(\tau = \infty) > 0$. This requires that we *compactify* the problem, by prescribing a conventional, yet natural, reward value, in the event that the gambler never stops. We set $Y_\infty \coloneqq \limsup_{i \to \infty} Y_i = 0$, since $\mathbb{P}(H < \infty) = 1$, so if the gambler never stops, a.s. the horizon will have been reached, so there can be no reward. This way infinite stopping rules do not increase the largest possible expected reward and the equivalence with the original formulation is maintained. In conclusion, RH has been reformulated as an infinite optimal stopping problem with reward sequence $\mathbf{Y} \coloneqq \{Y_i\}_{i \in \overline{\mathbb{N}}_0}$, underlying process $\mathbf{X}$ and filtration $\mathcal{F} \coloneqq \{\mathcal{F}_i\}_{i \in \overline{\mathbb{N}}_0}$, where $\mathcal{F}_\infty \coloneqq \sigma\left(\bigcup_{i \in \mathbb{N}_0} \mathcal{F}_i\right)$ (the full information about the history of the process, which is the ending point of all infinite filtration we will introduce), with respect to the class $\mathcal{T}^*$. We seek to find the value of the problem $\mathcal{V}(\mathbf{Y}) \coloneqq \sup_{\tau \in \mathcal{T}^*} \mathbb{E}Y_\tau$ and, if it exists, an optimal stopping rule $\bar{\tau} \in \mathcal{T}^*$. *Optimal* means that $\mathbb{E}Y_{\bar{\tau}} = \mathcal{V}(\mathbf{Y})$. The reference to the horizon H can be made explicit, whenever the necessity arises, via the notation $\mathbf{Y}^{(H)} \coloneqq \{Y_i^{(H)}\}_{i \in \overline{\mathbb{N}}_0}$.

2.3   Optimal algorithm

Under suitable hypotheses, an optimal stopping problem with independent horizon H, having survival function $S(i)$, is equivalent to a finite (if H is bounded) or infinite (if H is unbounded) optimal stopping problem *discounted* by factors $S(i)$. The discounted problem is $\mathbf{Z} \coloneqq \{Z_i\}_{i \in \overline{\mathbb{N}}_0}$, where $Z_0 \coloneqq Y_0$, $Z_i \coloneqq S(i)X_i$ for every $i \in \mathbb{N}$, $Z_\infty \coloneqq \limsup_{i \to \infty} Z_i = 0 = Y_\infty$. That is, we change only the reward function, not the underlying process. Thus, for $\mathbf{Z}$ we use the filtration $\mathcal{B} \coloneqq \{\mathcal{B}_i\}$, having set $\mathcal{B}_0$ and $\mathcal{B}_\infty$ as usual. The latter is used when the problem is infinite. Thus $\mathcal{V}(\mathbf{Z}) \coloneqq \sup_{\zeta \in \mathcal{T}} \mathbb{E}Z_\zeta$, and the equivalence aforementioned means that $\mathcal{V}(\mathbf{Y}) = \mathcal{V}(\mathbf{Z})$.

In [43] the case where the horizon H is bounded (finite support $\mathrm{supp}(H)$) is solved. Let $m \coloneqq \sup \mathrm{supp}(H) < \infty$, and take the discounted problem $Z_0, Z_1, \ldots, Z_m, 0, \ldots$ denoted $\mathbf{Z}^{(m)}$. Then *backward induction*, starting at m, yields the optimal stopping rule $\bar{\zeta} \in \mathcal{T}^m$ attaining $\mathcal{V}(\mathbf{Z})$, where $\mathcal{T}^m$ is the class of all all stopping rules in $\mathcal{T}$ that stop no later than time m. The stopping rule $\bar{\tau} \coloneqq \bar{\zeta} \wedge (H+1)$ then is shown to yield $\mathcal{V}(\mathbf{Y})$. The setting considered in [43] does not allow, in general, for the existence of an optimal stopping rule for unbounded horizons, which require an infinite $\mathbf{Z}$. Only the equivalence with the discounted problem holds in general. The unbounded case is crucial for RH, as an optimal stopping rule is needed explicitly. The main contribution of our Theorem 2.1 is showing that RH with unbounded horizon H admits an optimal stopping rule of the form $\bar{\tau} \coloneqq \bar{\zeta} \wedge (H+1)$ as well, where we will take $\bar{\zeta}$ to be the *Snell stopping rule* for $\mathbf{Z}$, since for an infinite problem, backward induction is not possible. Let us recall a few facts, adapted from [11, 12] to our compactified framework. Let $\mathcal{T}_i$ be the class of all all stopping rules in $\mathcal{T}$ that do not stop before step i.

Definition 2.1 (Snell envelope)**.** *The* Snell envelope *of* $\mathbf{Z}$ *is* $\{V_i\}$, *where* $V_i \coloneqq \mathrm{ess\,sup}_{\zeta \in \mathcal{T}_i} \mathbb{E}_{\mathcal{B}_i} Z_\zeta$.

Note that $\mathcal{V}(\mathbf{Z}) = V_0$. The intuitive interpretation of the stochastic process $\mathbf{V} = \{V_n\}$ is the best expected return available at step i among rules that have reached step i.

Remark 2.1. *Since $\mathbb{E}\sup_{i\in\mathbb{N}} Z_i < \infty$ for every $i \in \mathbb{N}_0$ the Snell envelope satisfies*

$$V_i = Z_i \vee \mathbb{E}_{\mathcal{B}_i} V_{i+1}. \tag{2.1}$$

Informally, (2.1) represents an extension of the dynamic programming principle for the infinite process $\{V_i\}$: at time i it is optimal to stop whenever the currently inspected value Z_i is greater than the expected highest expected future return as per the previous informal characterisation of V_i. We use (2.1) to compute V_0. Note that for finite problems, it coincides with backward induction.

Definition 2.2 (Snell stopping rule)**.** *The* Snell rule *is $\bar\zeta := \inf\{i \in \mathbb{N}_0 : Z_i \geq \mathbb{E}_{\mathcal{B}_i} V_{i+1}\}$.*

Remark 2.2. *If $\mathbb{E}\sup_{i\in\mathbb{N}} Z_i < \infty$ and $Z_\infty := \limsup_{i\longrightarrow\infty} Z_i \in \mathbb{R}$, then the* Snell rule $\bar\zeta$ *is optimal.*

We are now ready to prove the main result. The details of the proof are provided in Appendix A.

Theorem 2.1. *Let H be a finite horizon with $\mu := \mathbb{E}H$ and possibly mass at zero, and let $m := \sup\operatorname{supp}(H)$, $\mathbf{Y}$ and $\mathbf{Z}$ as previously defined. It holds that:*

- *$\mathcal{V}(\mathbf{Y}) = \mathcal{V}(\mathbf{Z})$;*
- *there exists $\bar\zeta \in \mathcal{T}$ such that $\mathbb{E}Z_{\bar\zeta} = \mathcal{V}(\mathbf{Z})$;*
- *$\bar\tau := \bar\zeta \wedge (H+1)$ is such that $\mathbb{E}Y_{\bar\tau} = \mathcal{V}(\mathbf{Y})$.*

Furthermore, $\bar\zeta$ is characterised as:

- *The* backward induction stopping rule *for $\mathbf{Z}^{(m)}$, if $m < \infty$;*
- *The* Snell rule *for $\mathbf{Z}$, if $m = \infty$.*

Idea of the proof. The case $m < \infty$ corresponds to [43, Theorem 2.1]. The most crucial step in the proof of the case $m = \infty$ is showing that for every stopping rule $\sigma \in \mathcal{T}^*$, we can construct a stopping rule $\zeta \in \mathcal{T}$, such that $\mathbb{E}Y_\sigma = \mathbb{E}Y_\tau$ (that is σ and τ are *equivalent*), where $\tau := \zeta \wedge (H+1) \in \mathcal{T}^*$. This requires an inductive measure-theoretic construction. For every $i \in \mathbb{N}$, let $E_i := \{\sigma = i\} \cap \{H \geq i\}$. We have that $E_i \in \mathcal{F}_i \cap \{H \geq i\} = \mathcal{B}_i \cap \{H \geq i\}$, where the intersections denote, as standard, the corresponding *trace σ-algebras*. Next we construct inductively mutually disjoint events $A_i \in \mathcal{B}_i$, such that $E_i = A_i \cap \{H \geq i\}$. These events $\{A_i\}$ are then relied upon to define ζ as, informally speaking, a stopping rule that stops when σ successfully stops by the time the horizon has realised, and does not stop otherwise. Formally, denote $C := \left(\bigcup_{i\in\mathbb{N}} A_i\right)^c \in \mathcal{B}_\infty$, and define for every $i \in \mathbb{N}$,

$$\zeta(\omega) := i, \; \omega \in A_i, \tag{2.2}$$

while setting ζ to ∞ otherwise. Following the standard convention $\mathbb{E}Y_\sigma := \mathbb{E}\sum_{i\in\mathbb{N}} Y_i \mathbb{1}_{\{\sigma=i\}} + \mathbb{E}Y_\infty$, we verify by direct computation that $\mathbb{E}Y_\sigma = \mathbb{E}Y_\tau$ is ensured by $\mathbb{E}Y_\infty = 0$.

In order to show that $\mathcal{V}(\mathbf{Y}) = \mathcal{V}(\mathbf{Z})$, note that since $Y_\infty = Z_\infty = 0$, we can show firstly, that for every $\sigma \in \mathcal{T}^*$, $\mathbb{E}Y_\sigma = \mathbb{E}Z_\zeta$, with ζ as defined in (2.2), and secondly, that for every $\zeta \in \mathcal{T}$,

$$\mathbb{E}Y_\zeta = \mathbb{E}Y_\tau, \tag{2.3}$$

with $\tau := \zeta \wedge (H+1)$. These two facts imply, by contradiction, that $\sup_{\sigma\in\mathcal{T}^*} \mathbb{E}Y_\sigma = \sup_{\zeta\in\mathcal{T}} \mathbb{E}Z_\zeta$.

Given an optimal stopping rule $\bar\zeta \in \mathcal{T}$, we have that $\bar\tau := \bar\zeta \wedge (H+1) \in \mathcal{T}^*$ provides an optimal rule by an argument by contradiction based on (2.3). Furthermore, our assumptions ensure that $\mathbb{E}\sup_{i\in\mathbb{N}} Z_i < \infty$ and $\limsup_{i\longrightarrow\infty} Z_i = 0$. Hence the Snell envelope $\mathbf{V}$ of $\mathbf{Z}$ satisfies (2.1) by Remark 2.1, and yields the optimal stopping rule in Definition 2.2 by Remark 2.2. $\qquad\square$

11

The following result gives the intuitive interpretation of $\mathbb{E}_{\mathcal{B}_i} V_{i+1}$ as the best expected return available at step i among rules that will not stop *by or at* step i. We believe that it is known within the optimal stopping community, but we could not find a reference nor a proof. The proof is provided in Appendix A for ease of the reader. As usual, assume that the process $\mathbf{Z}$, adapted to a filtration $\mathcal{B}$, satisfies $\mathbb{E}\sup_{i\in\mathbb{N}_0} Z_i < \infty$ and $\limsup_{i\longrightarrow\infty} Z_i = Z_\infty = 0 = Z_0$.

Lemma 2.1. *Let* $\mathbf{V}$ *be the Snell envelope of* $\mathbf{Z}$. *Then* $\mathbb{E}_{\mathcal{B}_i} V_{i+1} = \operatorname{ess\,sup}_{\zeta\in\mathcal{T}_{i+1}} \mathbb{E}_{\mathcal{B}_i} Z_\zeta$ *a.s.*

3 Single-threshold 2-approximation on the $\mathcal{G}$ class

We build familiarity with the model by proving Theorem 1.1, which derives a 2-approximation on the $\mathcal{G}$ class, extending the results of [2, Section 3]. Recall that $X \sim \mathcal{D}$ is the distribution of the values and $H \sim \mathcal{H}$ of the horizon (with no mass at zero). The expected return of a single-threshold algorithm is readily computed. The details are included in Appendix B.1 for ease of the reader.

Lemma 3.1. *Let* $\pi > 0$ *and consider the stopping rule* $\tau_\pi := \inf\{i \in \mathbb{N} : Y_i \geq \pi\} \in \mathcal{T}^*$, *where* $Y_i := X_i \mathbb{1}_{\{H \geq i\}}$. *Then* $\mathbb{E}X_{\tau_\pi} = c_\pi \mathbb{E}(X|X \geq \pi)$, *where* $c_\pi = c_\pi(\mathcal{H}, \mathcal{D}) := 1 - \mathbb{E}\left[\mathbb{P}^H(X < \pi)\right]$.

The quantity c_π yields the competitive ratio. With this in mind, we need to upper-bound $\mathbb{E}M_H$. This is done by *ex-ante* relaxation of the prophet, combined with a straightforward variational argument.[10] Further details and the complete proof of Lemma 3.2 are provided in Appendix B.2.

Lemma 3.2. *Let* X *be continuous and* $p > 0$ *such that* $\mathbb{P}(X \geq p) = {}^1\!/\!_\mu$. *Then* $\mathbb{E}M_H \leq \mathbb{E}(X|X \geq p)$.

Idea of the proof. Let $f(x)$ be the probability density function (pdf) of M_H and $v(x)$ the pdf of X. In order to upper-bound $\mathbb{E}M_H = \int_0^\infty xf(x)dx =: I(f)$, we solve the variational problem of maximising the functional $I(f)$ with constraints: $\int_0^\infty f(x)dx = 1$; $0 \leq f(x) \leq \mu v(x)$ for all $x \in [0, \infty)$; f is nonnegative Lebesgue integrable on $[0, \infty)$, the constraint follows from a union bound. The maximal solution is $\bar{f}(x) = \mu v(x)$ for all $x \geq p$, and zero otherwise, with p as in the claim. Then $I(\bar{f}) = \mathbb{E}(X|X \geq p)$. $\qquad\square$

To prove Theorem 1.1 we extend Lemma 3.2 to discontinuous X via *stochastic tie-breaking*. The technique is standard with deterministic horizons, but with random horizons it is not yet clear that the assumptions imposed on RH are robust enough. We will show that an underlying property of the prophet, Uniform Integrability (UI), is the key. First we derive the equivalent of Lemma 3.1 for a randomised algorithm (randomisation does not increase the optimal value of optimal stopping problems). Attach to every X_i a biased coin flip, that is *iid* Bernoulli random variables $B_i \sim B \sim \mathrm{Ber}(q)$ independent of X and H, q denoting the probability of 1 (heads).[11] We denote the class of adapted, and thus randomised, stopping rules as $\overline{\mathcal{T}}$.

Lemma 3.3. *Let* $\pi > 0$ *and consider the stopping rule* $\tau_{\pi,q} := \inf\{i \in \mathbb{N} : Y_i \geq \pi\} \in \overline{\mathcal{T}}$. *Then* $\mathbb{E}X_{\tau_{\pi,q}} = c_{\pi,q}\mathbb{E}(X|X \geq \pi)$, *where* $c_{\pi,q} = c_{\pi,q}(\mathcal{H}, \mathcal{D}) := 1 - \mathbb{E}\left\{[1 - q\mathbb{P}(X \geq \pi)]^H\right\}$.

---

[10]Lemma 3.1 and a slightly incomplete, informal version of Lemma 3.2 leveraging mathematical programming already appeared in [2, Theorem 3.2]. The contribution is valuable and insightful. The alternative argument we provide however is not only simpler, but formalises the claim and fills some gaps, which will demand our attention later.

[11]This extended version of RH, can be treated through the equivalent process $Y_i := X_i B_i \mathbb{1}_{\{H \geq i\}}$, whose history will generate the extended filtration $\overline{\mathcal{F}}$ defined via $\overline{\mathcal{F}}_i := \sigma(\mathbb{1}_{\{H=0\}}, X_1, B_1, \ldots, \mathbb{1}_{\{H=i-1\}}, X_i, B_i)$. Formally, randomised stopping rules are those adapted to $\overline{\mathcal{F}}$.

Rewriting $c_\pi = 1 - \mathbb{E}\left\{[1 - \mathbb{P}(X \geq \pi)]^H\right\}$ in Lemma 3.1 helps seeing the statement of Lemma 3.3 as its natural generalisation. The proof of Lemma 3.3 is combinatorially tedious, but straightforward, and is provided in Appendix B.3. We are now ready to show the main result of this section, Theorem 1.1. The details of the proof are provided in Appendix B.4.

Idea of the proof of Theorem 1.1. If X is continuous, by Lemmas 3.1 and 3.2,

$$\frac{\mathbb{E}X_{\tau_p}}{\mathbb{E}M_H} \geq c_p := 1 - \mathbb{E}\left[\mathbb{P}^H(X < p)\right].$$

For any $H \in \mathcal{G}$, $G \prec_{pgf} H$ with $G \sim \text{Geom}(1/\mu)$, $\mu = \mathbb{E}H$. Let $t := \mathbb{P}(X < p) = 1 - 1/\mu$, then

$$\mathbb{E}t^H \leq \mathbb{E}t^G = \frac{1}{\mu}\sum_{h\in\mathbb{N}} t^h \left(1 - 1/\mu\right)^{h-1} = \frac{1 - 1/\mu}{2 - 1/\mu}.$$

This implies that $c_p \geq (2 - 1/\mu)^{-1}$ and the result follows.

Consider now X discontinuous. Let $V(x)$ be the cumulative distribution function (cdf) of X. Assume that for some p, $\lim_{\varepsilon \to 0^+} V(p - \varepsilon) < 1 - 1/\mu < V(p)$, that is $1 - 1/\mu$ is in correspondence of a jump of the cdf at p.[12] $V(x)$ has at most countably many jump discontinuities at values $D := \{j_k\}$ (increasingly enumerated), one of which is p. Let $\{\varepsilon(l)\}$ be a small enough positive monotonically vanishing sequence. Linearly interpolate $V(x)$ on disjointed intervals of length $\varepsilon(l)$ (thus eliminating all jump discontinuities) and keep it unaltered otherwise. This yields an approximating sequence of continuous and piecewise differentiable cdf's $V_l(x)$ of continuous random variables $X^{(l)}$. Upon showing that the family $\{X^{(l)}\}$ is Uniformly Integrable (UI), define $M_H^{(l)} := \max\{X_1^{(l)}, \ldots, X_H^{(l)}\}$. Then $\{M_H^{(l)}\}$ is UI. Furthermore, since $X^{(l)} \xrightarrow{w} X$, $M_H^{(l)} \xrightarrow{w} M_H$. By Lemma 3.2, for every l there is p_l such that $\mathbb{P}(X^{(l)} \geq p_l) = 1/\mu$ and $\mathbb{E}M_H^{(l)} \leq \mathbb{E}(X^{(l)}|X^{(l)} \geq p_l)$. By construction, $p_l \to p$ as $l \to \infty$, so $\mathbb{E}(X^{(l)}|X^{(l)} \geq p_l) \to \mathbb{E}(X|X \geq p)$.[13] These facts altogether ensure that $\mathbb{E}M_H \leq \mathbb{E}(X|X \geq p)$.

To conclude, for any $H \in \mathcal{G}$, since $G \prec_{pgf} H$, by Lemma 3.3 the single-threshold randomised algorithm $\tau_{p,\bar{q}}$, with $\bar{q} := [\mu\mathbb{P}(X \geq p)]^{-1}$, is such that

$$\frac{\mathbb{E}X_{\tau_{p,\bar{q}}}}{\mathbb{E}M_H} \geq c_{p,\bar{q}} := 1 - \mathbb{E}\left\{[1 - \bar{q}\mathbb{P}(X \geq p)]^H\right\} = 1 - \mathbb{E}\left[\left(1 - \frac{1}{\mu}\right)^H\right] \geq \frac{1}{2 - \frac{1}{\mu}}.$$

$\square$

Remark 3.1. *For every horizon $H \in \mathcal{G}$ the following hold.*

- *If X is continuous, the $2 - 1/\mu$-approximation is a single-threshold algorithm, where the threshold is the value p such that $\mathbb{P}(X \geq p) = 1/\mu$.*
- *If X is discontinuous with $1/\mu$ in correspondence of a discontinuity jump of the survival function of X, the $2 - 1/\mu$-approximation is a single-threshold randomised algorithm, where the threshold is the value p at which the discontinuity jump occurs, and the randomisation parameter is $[\mu\mathbb{P}(X \geq p)]^{-1}$.*
- *As a result if X is discontinuous but $1/\mu$ does not occur in correspondence of a discontinuity jump, the randomisation parameter is 1, meaning that the nonrandomised algorithm for the continuous case is enough to yield a $2 - 1/\mu$-approximation.*

---

[12]This is the most difficult case to handle, when adapting the previous part of the argument. To present the key ideas of how Lemma 3.2 can be ensured for a discontinuous value distribution, we focus solely on this scenario. The extension to easier scenarios is trivial.

[13]A reference for convergence in expectation under UI and weak convergence hypotheses is [4].

Some of the most well-known subclasses of the $\mathcal{G}$ class, which earned considerable interest in applications, are: IHRE (Increasing Hazard Rate in Expectation); HIHRE (Harmonically Increasing Hazard Rate in Expectation); NBU (New Better than Used); NBUE (New Better than Used in Expectation). The following tower of inclusions is well-established [6, 40, 28] and known to be strict: IHR $\subset$ IHRE $\subset$ HIHRE $\subset$ NBU $\subset$ NBUE $\subset$ HNBUE $\subset \mathcal{G}$. Further details on these subclasses are given in Appendix B.5.

The tight instance for the $2-1/\mu$-approximation involves a geometric horizon, and was studied in [2, Theorem 3.5]. Their basic assumptions, concerning a two-point value distribution, are intuitive facts: that an optimal algorithm with geometric horizon should exist, and that it has a single-threshold. We show mathematically that they both hold, with the second extending, more in general, to any bounded value distribution. As a bonus we obtain an equation for the optimal threshold. The details of the proof of Lemma 3.4 are provided in Appendix B.6. Recall that $\text{Geom}(1-q)$ denotes the geometric distribution with failure probability $0 < q < 1$.

Lemma 3.4. *Let* $H \sim \text{Geom}(1-q)$, *and* X *such that* $\text{supp}(X) = [a, b]$, $0 \leq a \leq b$. *Then there exists a threshold* V_0 *such that the corresponding single-threshold algorithm is optimal.*

Idea of the proof. By Theorem 2.1, $\mathcal{V}(\mathbf{X}) = \mathcal{V}(\mathbf{Y}) = \mathcal{V}(\mathbf{Z}) =: V_0$. By (2.1), it holds that $V_0 = \mathbb{E}V_1 = \mathbb{E}(Z_1 \vee \mathbb{E}_{\mathcal{B}_1} V_2)$. Since $Z_i := q^{i-1} X_i$, by time-invariance and Lemma 2.1 it follows that $\mathbb{E}_{\mathcal{B}_1} V_2 = q\mathbb{E}V_1 = qV_0$. Therefore, $V_0 = \mathbb{E}(X \vee qV_0)$; equivalently, if X has cdf $F(x)$,

$$f(V_0) := V_0 + \int_{qV_0}^{b} F(x)dx = b.$$

The function $f(V)$ is a strictly increasing differentiable function on $[0, b]$, and $f(0) \leq b \leq f(b)$. Hence the value V_0 exists. By Remark 2.2 and Definition 2.2 and time-invariance again, it is optimal to stop at the first value greater or equal than V_0 (which is also the expected return). $\square$

The tightness then follows by the straightforward computation of [2][Theorem 3.5]. For self-containedness we provide the short computation in Appendix B.7, with small variations to better fit our framework.

4 A first step beyond single-threshold algorithms

In this section we construct a parametric family of horizons, which is hard for any single-threshold algorithm, and yet it allows for an adaptive competitive stopping rule. We also show that it can be perturbed so that it enters the $\overline{\mathcal{G}}$ class, while remaining hard for single-threshold algorithms.

4.1 The hard instance

In this section we construct the instance, which will provide the hardness result for single-thresholds. Let us start with the instance X for the values, which is a Pareto (Type I) distribution with scale parameter 1 and shape parameter $1 + \varepsilon$, where $\varepsilon > 0$ is fixed small enough. The corresponding cdf, denoted as $V(x)$, vanishes on $(-\infty, 1)$, whereas on $[1, \infty)$

$$V(x) = 1 - x^{-(1+\varepsilon)}. \tag{4.1}$$

Since M_n has cdf $F_n(x) = V^n(x)$, which is differentiable for all $x \neq 1$, we obtain that its density (defined almost everywhere) $f_n(x)$ vanishes on $(-\infty, 1)$, whereas $f_n(x) = nV^{n-1}(x)V'(x)$ on $(1, \infty)$. By direct computation of $\mathbb{E}M_n := \int_{\mathbb{R}} x f_n(x)dx$, exploiting the properties of the Beta function and well-known asymptotics of the Gamma function, the following result holds. The proof is provided in Appendix C.1.

Lemma 4.1. *As $n \longrightarrow \infty$, uniformly in $\varepsilon > 0$, $\mathbb{E}M_n \sim \Gamma\left(1 - \frac{1}{1+\varepsilon}\right) n^{\frac{1}{1+\varepsilon}}$.*

In conjunction with the instance X for the values, consider a family of horizons $\{H_m\}$, with $m \in \mathbb{N}$ large enough, such that for every fixed such m, $\mathrm{supp}(H_m) = [\ell, m] \cap \mathbb{N}$, with the lower limit $1 < \ell \in \mathbb{N}$. For every m fixed, we define the pmf of each horizon by letting, for every $\ell \leq h \leq m$,

$$p_h^{(m)} := \mathbb{P}(H_m = h) = Z_m^{-1} h^{-\frac{1}{1+2\varepsilon}}, \tag{4.2}$$

having defined the normalising constant $Z_m := \sum_{h=\ell}^{m} h^{-\frac{1}{1+2\varepsilon}}$. Through standard integral estimates of summations, it is shown that

$$Z_m \sim [1 + (2\varepsilon)^{-1}] m^{\frac{2\varepsilon}{1+2\varepsilon}}, \tag{4.3}$$

and that the following holds.

Lemma 4.2. *As $m \longrightarrow \infty$, uniformly in $\varepsilon > 0$ small enough, $\mathbb{E}M_{H_m} \sim Nm^{\frac{1}{1+\varepsilon}}$, where we defined $N = N(\varepsilon) := \frac{\Gamma\left(1 - \frac{1}{1+\varepsilon}\right)}{\left[1 + \frac{\varepsilon}{(1+\varepsilon)(1+2\varepsilon)}\right][1 + (2\varepsilon)^{-1}]}$.*

The proof of (4.3) can be found in the proof of Lemma 4.2, which is provided in Appendix C.1. By (4.3) and similar integral estimates as in Lemma 4.2 it is straightforward to compute that for every $n \in \mathbb{N}$, as $m \longrightarrow \infty$,

$$\mathbb{E}H_m^n \sim \frac{\sum_{h=\ell}^{m} h^{n+1-\frac{1}{1+2\varepsilon}}}{[1 + (2\varepsilon)^{-1}] m^{\frac{2\varepsilon}{1+2\varepsilon}}} \sim \frac{m^{n - \frac{1}{1+2\varepsilon}}}{\left(n + 1 - \frac{1}{1+2\varepsilon}\right)[1 + (2\varepsilon)^{-1}] m^{\frac{2\varepsilon}{1+2\varepsilon}}} = \frac{2\varepsilon m^n}{n + 2(n+1)\varepsilon}. \tag{4.4}$$

Note that from the elementary properties of the Gamma function it holds that as $x \longrightarrow 0^+$, $\Gamma(x) \sim x^{-1}$. As a result, as ε vanishes, $N \sim 2\varepsilon\left(1 - (1+\varepsilon)^{-1}\right)^{-1} \longrightarrow 2$ and the following is immediate consequence.

Remark 4.1. *As $m \longrightarrow \infty$, uniformly in $\varepsilon > 0$ small enough, $\mathbb{E}M_{H_m} \asymp m^{\frac{1}{1+\varepsilon}}$.*

We are now ready to prove the hardness of the family of horizons $\{H_m\}$ for single-threshold algorithms. A stopping rule with single-threshold π is denoted as τ_π. A crucial point is that π can be tailored to each horizon in the family, that is, the algorithm facing H_m will set a threshold π_m exploiting all distributional knowledge regarding X and H_m. Yet, this is not enough to obtain a constant-approximation with the family $(X, \{H_m\})$ for suitably small ε. The details of the proof, which involves a quite technical asymptotic analysis, is provided in Appendix C.2. We also include some supporting numerics and heuristic consideration in Appendix C.3, which ease assessing that the family of horizons in question has divergent CV as ε vanish (thus violating the requirement of Theorem 1.4), and admits horizons outside the $\mathcal{G}$ and $\overline{\mathcal{G}}$ class (thus violating the requirements of Theorem 1.1 and Theorem 1.3).

Proposition 4.1. *For fixed $\varepsilon > 0$ small enough, given the instance of copies of X and the family of horizons $\{H_m\}$ of (4.1) and (4.2) respectively, we have that for every sequence of thresholds $\{\pi_m\}$, $\liminf_{m \to \infty} \frac{\mathbb{E}X_{\tau_{\pi_m}}}{\mathbb{E}M_{H_m}} = 0$.*

Idea of the proof. Our strategy will be to characterise the sequence of optimal single thresholds $\{\bar{\pi}_m\}$ (corresponding to horizons $\{H_m\}$ for m large enough and ε small enough), where $\bar{\pi}_m$ maximizes the value obtained $\mathbb{E}X_{\tau_{\pi_m}} := \mathbb{E}Y_{\tau_{\pi_m}}^{(H_m)}$ over all single threshold strategies π_m. Denote with r_m the gambler-to-prophet ratio of $\mathbb{E}Y_{\tau_{\bar{\pi}_m}}^{(H_m)}$ and $\mathbb{E}M_{H_m}$. As a consequence of the characterisation obtained, it will follow that there is always a subsequence $\{m_j\}$, such that as $j \longrightarrow \infty$, r_{m_j} vanishes. This yields the claim, since r_{m_j} is an upper bound on any gambler-to-prophet ratio for arbitrary single-thresholds.

Since X is continuous, by Lemma 3.1 we compute $\mathbb{E}Y_{\tau_\pi}^{(H_m)}$, which is nontrivial only if $\pi > 1$. This is expressed as the product of

$$\mathbb{E}(X|X \geq \pi) = (1 + 1/\varepsilon)\,\pi$$

and

$$c_\pi = 1 - \mathbb{E}\left[V^{H_m}(\pi)\right] = 1 - \mathbb{E}\left[\left(1 - 1/\pi^{1+\varepsilon}\right)^{H_m}\right].$$

Next, we show analytically that the optimal sequence of thresholds $\{\bar{\pi}_m\}$ satisfies the *optimality equation*

$$g_m(\bar{\pi}_m) := \mathbb{E}\left[\left(1 - \frac{1}{\bar{\pi}_m^{1+\varepsilon}}\right)^{H_m}\right] + \frac{1+\varepsilon}{\bar{\pi}_m^{1+\varepsilon}}\mathbb{E}\left[H_m\left(1 - \frac{1}{\bar{\pi}_m^{1+\varepsilon}}\right)^{H_m - 1}\right] = 1.$$

To conclude, by these facts and Lemma 4.2 and Remark 4.1 we have that, as $m \longrightarrow \infty$, uniformly in $\varepsilon > 0$ small enough,

$$r_m = c_{\bar{\pi}_m}\frac{\mathbb{E}(X|X \geq \bar{\pi}_m)}{\mathbb{E}M_{H_m}} \sim \left(1 + \frac{1}{\varepsilon}\right)\frac{c_{\bar{\pi}_m}\bar{\pi}_m}{Nm^{\frac{1}{1+\varepsilon}}} \asymp \frac{c_{\bar{\pi}_m}\bar{\pi}_m}{\varepsilon m^{\frac{1}{1+\varepsilon}}}.$$

Consider that $\{\bar{\pi}_m\}$ is either bounded or unbounded.

- In the bounded case, since $c_{\bar{\pi}_m} < 1$, r_m vanishes as $m \longrightarrow \infty$, for any fixed ε small enough. Note that this case covers also the trivial case for which there exists a subsequence $\{\bar{\pi}_{m_j}\}$ such that $\bar{\pi}_{m_j} \leq 1$, for which the numerator of r_{m_j} becomes upper-bounded by $\nu := \mathbb{E}X$.
- Before moving on to the general unbounded case, we consider the subcase of divergence to infinity with $m^{\frac{1}{1+\varepsilon}} = \mathcal{O}\left(\bar{\pi}_m\right)$. Note that by Bernoulli's inequality we have that

$$c_{\bar{\pi}_m} := 1 - \mathbb{E}\left[\left(1 - 1/\bar{\pi}_m^{1+\varepsilon}\right)^{H_m}\right] \leq \frac{\mathbb{E}H_m}{\bar{\pi}_m^{1+\varepsilon}}.$$

By (4.4) with $n = 1$ it holds that as $m \longrightarrow \infty$, for every fixed ε small enough r_m vanishes, since

$$c_{\bar{\pi}_m}\bar{\pi}_m m^{-\frac{1}{1+\varepsilon}} \leq \mathbb{E}H_m m^{-\frac{1}{1+\varepsilon}}\bar{\pi}_m^{-\varepsilon} \asymp \varepsilon\left(m^{\frac{1}{1+\varepsilon}}\bar{\pi}_m^{-1}\right)^\varepsilon \longrightarrow 0.$$

- More in general, we can rely on the previous point to show that the unbounded case overall only allows for r_m to vanish. Consider that since $0 \leq r_m \leq 1$, if all of its convergent subsequences $\{r_{m_j}\}$ vanish for all ε small enough, then $\{r_m\}$ vanishes for all ε small enough. We show the sufficient condition by contradiction: assume that as $j \longrightarrow \infty$, $r_{m_j} \longrightarrow \rho = \rho(\varepsilon) \in (0, 1]$, with $\rho(\varepsilon)$ being bounded away from 0 no matter how small ε is taken. Then we have that as $j \longrightarrow \infty$, for $\varepsilon > 0$ fixed small enough, $\mu_j := m_j/\bar{\pi}_{m_j}^{1+\varepsilon}$ is asymptotically equivalent to a bounded sequence, so it must be bounded too. Thus, we can consider a convergent subsequence of $\{\mu_j\}$, denoted $\{\mu_{j_k}\}$, such that as $k \longrightarrow \infty$, $\mu_{j_k} \longrightarrow \alpha = \alpha(\varepsilon) \geq 0$. A sharp asymptotic analysis of $c_{\bar{\pi}_{m_{j_k}}}$ reveals that the *optimality equation* satisfied by $\bar{\pi}_{m_{j_k}}$ allows only for α being bounded away from zero, unless it is identically zero. In what follows, it is useful to rewrite the *optimality equation* as

$$g_{m_{j_k}}(\bar{\pi}_{m_{j_k}}) := \mathbb{E}\left[\phi(\bar{\pi}_{m_{j_k}}, H_{m_{j_k}})\right] = 1,$$

having defined

$$\phi(\pi, H) := \left(1 - \pi^{-(1+\varepsilon)}\right)^{H-1}\left[1 - \pi^{-(1+\varepsilon)} + (1+\varepsilon)H\pi^{-(1+\varepsilon)}\right].$$

16

By estimating asymptotically the series expansion of $\phi(\bar{\pi}_{m_{j_k}}, H_{m_{j_k}})$ with respect to the first argument, uniformly in $\alpha > 0$, the following bound is shown to hold as $k \longrightarrow \infty$:

$$g_{m_{j_k}}(\bar{\pi}_{m_{j_k}}) \leq 1 + \varepsilon \left[(2 + \alpha)e^{-\alpha} - 2 \right] + \mathcal{O}(\varepsilon^2) + \mathcal{O}\left(\frac{1}{m_{j_k}} \right).$$

It follows that for $\varepsilon > 0$ small enough, as $k \longrightarrow \infty$, $g_{m_{j_k}}(\bar{\pi}_{m_{j_k}}) < 1$ if the strictly decreasing function $h(\alpha) := (2 + \alpha)e^{-\alpha} - 2$ is bounded away from zero. Note that $h(\alpha)$ is negative and vanishes as $\alpha \geq 0$ vanishes. Since we have previously shown that if $\alpha > 0$ exists, it must be bounded away from zero as ε vanishes (and thus $h(\alpha$ is bounded away from zero too)), in this case the optimality equation cannot be satisfied, as long as ε is taken sufficiently small. Thus only $\alpha = 0$ would not yield a contradiction at this point, meaning that for all ε small enough, $\mu_{j_k} \longrightarrow 0$ as $k \longrightarrow \infty$. By boundedness, also $\mu_j \longrightarrow 0$ as $j \longrightarrow \infty$. This corresponds to the scenario in the previous point (replace, in that argument, m with m_j) and therefore r_{m_j} vanishes. This is in contradiction with the assumption, which ensures $r_{m_j} \longrightarrow \rho > 0$, and therefore we must have $\rho = 0$.

The sequence $\{r_m\}$ is thus shown to always admit a vanishing subsequence. $\square$

4.2 Hardness of the $\overline{\mathcal{G}}$ class

In principle, there is no obvious reason why the main result of this section, Theorem 1.3, should hold, merely based on duality. However, a hint towards the hardness of the $\overline{\mathcal{G}}$ class could be that the duality turns the estimates of Section 3 the wrong way around. It is straightforward to see that the approach adopted on the $\mathcal{G}$ class is mirrored by the duality, meaning that the lower-bound, which provides the competitive ratio ensured, turns into an upper-bound, informally speaking. This mirroring has significance, because the ex-ante relaxation of the prophet is tight on geometric horizons, and is essentially maxed-out by the very definition of the $\mathcal{G}$ class. A similar phenomenon occurs with the argument of [2], if we restrict to the IHR and DHR class. The most notable subclasses of the $\overline{\mathcal{G}}$ class are the dual of those introduced at the end of the previous section: DHR $\subset$ DHRE $\subset$ HDHRE $\subset$ NWU $\subset$ NWUE $\subset$ HNWUE $\subset \overline{\mathcal{G}}$.[14] These facts, combined with Theorem 1.3, motivate our conjecture that no constant-approximation is possible on the DHR class either. The details of the following proof are in Appendix C.4.

Idea of the proof of Theorem 1.3. We show that a perturbation of the hard instance from the previous section prevents any single-threshold constant-approximation on the $\overline{\mathcal{G}}$ class.

Step 1. We start by constructing a family $\tilde{\mathcal{H}}_M(\varepsilon) \subset \overline{\mathcal{G}}$, which is a sequence of horizons $\tilde{H}_m$ arising as a perturbation of horizons H_m in the hard family $\mathcal{H}_M(\varepsilon)$ of (4.2), with fixed $\ell = 2$, as m grows, for all fixed $0 < \varepsilon < 1/4$ small enough. This is done by adding to the pmf of said H_m, a (to be suitably rescaled) mass at one,

$$\delta_m := Cm^{\frac{2\varepsilon}{1+2\varepsilon}}, \quad C = C(\varepsilon) := \frac{3 + 10\varepsilon}{2\varepsilon}.$$

We denote the perturbed horizon $\tilde{H}_m$ for every m, ε considered, and observe that the corresponding normalizing constants for the pmf's $\tilde{Z}_m$ and Z_m; and expectations $\tilde{\mu}_m := \mathbb{E}\tilde{H}_m$ and $\mu_m := \mathbb{E}H_m$; satisfy, as m grows, the following:

$$\tilde{Z}_m \sim \tilde{C} Z_m, \quad \tilde{C} := C + 1 + \frac{1}{2\varepsilon}$$

$$\tilde{\mu}_m \sim \frac{1}{\tilde{C}} \left(C + \frac{1 + 2\varepsilon}{1 + 4\varepsilon} m \right).$$

---

[14]Since these classes are obtained by reversing the inequalities in the corresponding definitions, the corresponding acronyms are obtained by replacing *increasing* with *decreasing* and *better* with *worse*.

Trivially, these follow from (4.3) and (4.4) with $n = 1$ and the definitions, respectively $\tilde{Z}_m :=
Z_m + \delta_m$ and $\tilde{\mu}_m = \delta_m/\tilde{Z}_m + \mu_m Z_m/\tilde{Z}_m$. To show that for all m large enough, $\tilde{H}_m \in \overline{\mathcal{G}}$, we have
to establish that eventually (short for all m large enough from now on), for all $t \in (0, 1)$ (the
implicit range from now on),

$$\frac{1}{\tilde{Z}_m} \left[\delta_m t + \sum_{h=2}^{m} t^h h^{-\frac{1}{1+2\varepsilon}} \right] \geq \frac{1}{\tilde{\mu}_m} \sum_{h=1}^{\infty} (1 - 1/\tilde{\mu}_m)^{h-1} t^h,$$

which we recast as

$$\sum_{h=2}^{m} t^h h^{-\frac{1}{1+2\varepsilon}} \geq -\delta_m t + \frac{\varepsilon_m t}{1 - t(1 - 1/\tilde{\mu}_m)}, \tag{4.5}$$

having set $\varepsilon_m := \tilde{Z}_m/\tilde{\mu}_m$. Let

$$\eta_m := \frac{1 - \varepsilon_m/\delta_m}{1 - 1/\tilde{\mu}_m}.$$

We establish that eventually for all $t \in (0, \eta_m]$

$$-\delta_m t + \frac{\varepsilon_m t}{1 - t(1 - 1/\tilde{\mu}_m)} \leq 0, \tag{4.6}$$

and that eventually for all $t \in [\eta_m, 1)$

$$\sum_{h=2}^{m} h(h-1) t^{h-2} h^{-\frac{1}{1+2\varepsilon}} < \frac{2\varepsilon_m (1 - 1/\tilde{\mu}_m)}{[1 - t(1 - 1/\tilde{\mu}_m)]^3}. \tag{4.7}$$

This implies (4.5) for all $t \in (0, 1)$, since eventually on $(0, \eta_m]$ (4.5) holds with strict inequality
directly by (4.6), while for all $t \in [\eta_m, 1)$ we have that the difference of the left-hand and right-
hand side of (4.5), denoted $f_m(t)$, is strictly concave ((4.7) states precisely that $f_m''(t) < 0$).
Since $f_m(t) > 0$ for all $t \in (0, \eta_m]$ and since it vanishes at both ends of the unit interval, it
cannot vanish on $(\eta_m, 1)$, meaning that there cannot be any crossings of the two sides of (4.5)
on $(\eta_m, 1)$, which therefore holds everywhere.

Step 2. Next, we show that with X being set to be the hard instance of values X of (4.1), with
$\varepsilon > 0$ fixed small enough as in *Step 1*, $\mathbb{E} M_{\tilde{H}_m} = \Omega(\mathbb{E} M_{H_m})$ and $\mathbb{E} X_{\tau_{\pi_m}} := \mathbb{E} Y_{\tau_{\pi_m}}^{(\tilde{H}_m)} \leq \mathbb{E} Y_{\tau_{\pi_m}}^{(H_m)}$
as m grows, for every sequence of thresholds $\{\pi_m\}$. The first fact follows from Remark 4.1 and
observing that the law of total expectation yields

$$\mathbb{E} M_{\tilde{H}_m} = \frac{\delta_m}{\tilde{Z}_m} \mathbb{E} X + \frac{Z_m}{\tilde{Z}_m} \mathbb{E} M_{H_m} > \frac{Z_m}{\tilde{Z}_m} \mathbb{E} M_{H_m} \sim \frac{1}{\tilde{C}} \mathbb{E} M_{H_m}.$$

The second fact follows from considering the survival functions of $\tilde{H}_m$ and H_m, respectively
$\tilde{S}_m(i)$ and $S_m(i)$. Note that for all $i \geq 3$, $\tilde{S}_m(i) = S_m(i) Z_m/\tilde{Z}_m$, whereas $\tilde{S}_m(2) = S_m(2) - \delta_m/\tilde{Z}_m$.
Consider next a single-threshold algorithm τ_π, where $\pi = \pi_m$ on horizon $\tilde{H}_m$. By Theorem 2.1
(note that the notations τ and σ used there are now swapped, as our focus here is on the original
stopping rule) for any horizon $\tilde{H}_m$, there is a stopping rule equivalent to τ_π for the problem
$\mathbf{Y}^{(\tilde{H}_m)}$, denoted as $\sigma_\pi := \zeta_\pi \wedge (\tilde{H}_m + 1)$, where ζ_π is, as per (2.2), an algorithm for the discounted
problem $\mathbf{Z} = \{\tilde{S}_m(i) X_i\}$, which stops only when τ_π successfully stops by the time the horizon
has realised. Intuitively, we can construct ζ_π as the stopping rule with single threshold π on the
events $\{\tau_\pi = i\} \cap \{\tilde{H}_m \geq i\}$. Otherwise it stops at m on $\{\tau_\pi > \tilde{H}_m\}$. The proof of Theorem 2.1
shows that this can be done, so that ζ_π is adapted to $\mathbf{X}$, and as a result, as m grows,

$$\mathbb{E} Y_{\tau_\pi}^{(\tilde{H}_m)} = \mathbb{E} \sum_{i=1}^{m} \tilde{S}_m(i) X_i \mathbb{1}_{\{\zeta_\pi = i\}} \leq \mathbb{E} \sum_{i=1}^{m} S_m(i) X_i \mathbb{1}_{\{\zeta_\pi = i\}} = \mathbb{E} Y_{\tau_\pi}^{(H_m)}.$$

Step 3. Finally, suppose by contradiction that a constant-approximation exists on the $\overline{\mathcal{G}}$ class. Then there exist a sequence of thresholds $\{\pi_m\}$ and $c > 0$ such that τ_{π_m} is a $1/c$-approximation on $\tilde{\mathcal{H}}_M(\varepsilon) \subset \overline{\mathcal{G}}$ by *Step 1*. By Proposition 4.1 for all $\varepsilon > 0$ fixed small enough, there exists a subsequence $\{\pi_{m_j}\}$ such that r_{m_j}, the gambler-to-prophet ratio of $\mathbb{E}Y^{(H_{m_j})}_{\tau_{\pi_{m_j}}}$ over $\mathbb{E}M_{H_{m_j}}$, vanishes. Combining this with *Step 2*'s asymptotics yields the following contradiction as j grows:

$$c \leq \frac{\mathbb{E}Y^{(\tilde{H}_{m_j})}_{\tau_{\pi_{m_j}}}}{\mathbb{E}M_{\tilde{H}_{m_j}}} = \mathcal{O}(r_{m_j}) \longrightarrow 0.$$

$\square$

4.3   The competitive Secretary Problem rule

In this section we show that a straightforward adaptation of the SP stopping rule with deterministic horizon m is competitive on the horizons H_m defined in (4.2). Consider any instance of the values X which is nonnegative, integrable and continuous random variable.[15] Consider the process $\mathbf{X} := (X_1, \ldots, X_m)$, where X_i are *iid* copies of X. We equivalently characterise it in the context of SP via a uniform random permutation $\pi : \Omega \longrightarrow S_m$, where S_m is the set of all permutations of $[m]$ and Ω the supporting probability space. In online notation, $\pi = (\pi_1, \ldots, \pi_m)$. Consider $\mathbf{X}^\pi := (X_{\pi_1}, \ldots, X_{\pi_m})$. Since $\mathbf{X}$ is exchangeable, $\mathbf{X}^\pi \sim \mathbf{X}$. Consider now the waiting time $r_m - 1$ for the SP stopping rule, denoted as ζ_{r_m}, and recall that

$$r_m \sim \frac{m}{e}. \tag{4.8}$$

Consider $\mathbf{X}^\pi$ conditionally on (the σ-algebra generated by) $\mathbf{X}$, denoted as $(\mathbf{X}^\pi | \mathbf{X})$. This is the process of realised values arriving in uniform random order, and constitutes therefore the usual setup for SP. In this framework, the classic result of [34] can be restated as

$$\mathbb{P}\left(X_{\pi_{\zeta_{r_m}}} = M_m | \mathbf{X}\right) \geq \frac{1}{e}. \tag{4.9}$$

Given the rule ζ_{r_m}, denote the event of winning (that is, choosing the best) as $W := \{X_{\pi_{\zeta_{r_m}}} = M_m\}$ and the event of winning at time i as $W_i := \{i = \inf\{j \geq r_m : X_{\pi_j} = M_m\}\}$. Note that by partitioning W conditionally on $\mathbf{X}$, we have that

$$\mathbb{P}(W | \mathbf{X}) = \sum_{i=r_m}^m \mathbb{P}(W_i | \mathbf{X}). \tag{4.10}$$

The proof of (4.9) relies on the following fact, which we will also exploit:

$$\mathbb{P}\left(i = \inf\{j \geq r_m : X_{\pi_j} = M_m\} | \mathbf{X}\right) = \frac{r_m - 1}{m} \frac{1}{i - 1}. \tag{4.11}$$

When considering RH, instead of working with $\mathbf{X}$ as usual, we start from $\mathbf{X}^\pi$, by defining accordingly $Y_i^{(H_m)} := X_{\pi_i} \mathbb{1}_{\{H_m \geq i\}}$. This is equivalent, since for every stopping rule $\tau \in \mathcal{T}^*$, $\mathbb{E}X_\tau = \mathbb{E}X_\tau^\pi = \mathbb{E}Y_\tau^{(H_m)}$. Yet it will be formally advantageous in the following proof of the main result of the section. The details of the argument are provided in Appendix C.5.

---

[15]The hypothesis of absolute continuity is to ensure that there are no ties in the ranking process involved in the SP rule. It is possible to also allow discontinuous random variables, by exploiting the following tie-breaking procedure: consider $(U_1, \ldots, U_m)$, where U_i are *iid* copies of $U \sim \text{Unif}(0, 1)$. Then define $\mathbf{X} = (X_1, U_1, \ldots, X_m, U_m)$, and instead of the usual order relationship that gives rise to the usual ranking for SP, adopt the following ranking order: if $X_i < X_j$, $(X_i, U_i) \preccurlyeq (X_j, U_j)$, whereas if $X_i = X_j$ and $U_i < U_j$, then by breaking ties thanks to the uniform random variables we say that, again, $(X_i, U_i) \preccurlyeq (X_j, U_j)$. Since the uniform random variable are continuous, there are almost surely no ties by ranking via $\preccurlyeq$ instead of $<$. For simplicity, we present the argument only for the continuous case and by using the standard $\leq$ as a ranking order.

Idea of the proof of Theorem 1.2. Consider the stopping rule $\tau_m := \zeta_{r_m} \wedge (H_m + 1) \in \mathcal{T}^*$. By (2.3) in Theorem 2.1 firstly, and the law of total expectation with respect to $\mathbf{X}$ secondly, we have that

$$\mathbb{E}Y_{\tau_m}^{(H_m)} = \mathbb{E}\left[S_m(\zeta_{r_m}) X_{\pi_{\zeta_{r_m}}} \right] = \mathbb{E}\left\{ \mathbb{E}\left[S_m(\zeta_{r_m}) X_{\pi_{\zeta_{r_m}}} | \mathbf{X} \right] \right\}.$$

Since at any given step (greater than $r_m - 1$) stopping at the overall maximum is a subevent of stopping at the relative maximum,

$$\mathbb{E}\left[S_m(i) X_{\pi_i} \mathbb{1}_{\{\zeta_{r_m}=i\}} | \mathbf{X} \right] \geq \mathbb{E}\left[S_m(i) X_{\pi_i} \mathbb{1}_{W_i} | \mathbf{X} \right] = S_m(i) M_m \mathbb{P}(W_i | \mathbf{X}).$$

Therefore,

$$\mathbb{E}\left[S_m(\zeta_{r_m}) X_{\pi_{\zeta_{r_m}}} | \mathbf{X} \right] \geq M_m \sum_{i=r_m}^{m} S_m(i) \mathbb{P}(W_i | \mathbf{X}).$$

By (4.3) and comparison of the summation with the corresponding integral as in Lemma 4.2, we estimate the survival function of H_m as $m \longrightarrow \infty$: recalling (4.8), for all $i > r_m \sim m/e$ we have that

$$S_m(i) > 1 - \left(\frac{i-1}{m+1} \right)^{\frac{2\varepsilon}{1+2\varepsilon}}.$$

Define the sequence of functions

$$f_m(\varepsilon) := (m+1)^{-\frac{2\varepsilon}{1+2\varepsilon}} \frac{r_m}{m} \int_{r_m-2}^{m-1} x^{-\frac{1}{1+2\varepsilon}} \, dx$$

and the function

$$g(\varepsilon) := \frac{1}{e} \left\{ 1 - \left[1 + \frac{1}{2\varepsilon} \right] \left(1 - e^{-\frac{2\varepsilon}{1+2\varepsilon}} \right) \right\}.$$

For every fixed $\varepsilon > 0$,

$$\sum_{i=r_m}^{m} S_m(i) \mathbb{P}(W_i | \mathbf{X}) \geq \sum_{i=r_m}^{m} \mathbb{P}(W_i | \mathbf{X}) - \sum_{i=r_m}^{m} \left(\frac{i-1}{m+1} \right)^{\frac{2\varepsilon}{1+2\varepsilon}} \mathbb{P}(W_i | \mathbf{X}) \geq \frac{1}{e} - f_m(\varepsilon) \longrightarrow g(\varepsilon)$$

as $m \longrightarrow \infty$, where we used (4.9) to (4.11) and the usual integral estimate in the second inequality, whereas the asymptotics follows by (4.8), the fact that

$$f_m(\varepsilon) \longrightarrow f(\varepsilon) := \frac{1}{e} \left[1 + \frac{1}{2\varepsilon} \right] \left(1 - e^{-\frac{2\varepsilon}{1+2\varepsilon}} \right),$$

and $g(\varepsilon) = 1/e - f(\varepsilon)$. It is crucial that $g(\varepsilon)$ is positive and strictly increasing on $(0, \infty)$: even though as $\varepsilon \longrightarrow 0$, $f(\varepsilon) = 1/e + \mathcal{O}(\varepsilon)$, yielding that $g(\varepsilon) = \mathcal{O}(\varepsilon)$, no matter how small, it still provides a constant approximation for any ε fixed.

Putting together all the above yields, as $m \longrightarrow \infty$, $\mathbb{E}[S_m(\zeta_{r_m}) X_{\pi_{\zeta_{r_m}}} | \mathbf{X}] \geq (g(\varepsilon) - \delta) M_m$, where $\delta \longrightarrow 0^+$ as $m \longrightarrow \infty$. This implies the final estimate

$$\mathbb{E}X_{\tau_m} := \mathbb{E}Y_{\tau_m}^{(H_m)} \geq (g(\varepsilon) - \delta) \mathbb{E}M_m.$$

Fix $\varepsilon > 0$ small, and consider horizons $\{H_m\}_{m \geq M}$ as defined in (4.2), where $M = M(\varepsilon) \in \mathbb{N}$ is large enough. By Proposition 4.1, the RH problem under this family does not admit single-threshold constant-approximations for the fixed ε, assuming it is small enough (the hard instance for the values being the Pareto distribution defined in (4.1)). On the other hand, by our final estimate, τ_m provides an approximate $g(\varepsilon)$-approximation: for every (small) $\varepsilon > 0$ fixed, setting M large enough, using $H_m \leq m$ yields that for every instance X,

$$\frac{\mathbb{E}X_{\tau_m}}{\mathbb{E}M_{H_m}} \geq g(\varepsilon) - \delta,$$

where $\delta = \delta(M) \longrightarrow 0$ as $M \longrightarrow \infty$. $\qquad \square$

5   A 2-approximation for concentrated $\mathcal{L}^2$-horizons

In this final section we prove Theorem 1.4, thus extending the 2-approximation for the $\mathcal{G}$ class to sufficiently concentrated horizons. Besides its practical value, this will reveal interesting connections with deterministic prophet inequalities. We will make use of the following stochastic order.

Definition 5.1 (Laplace transform order)**.** *Given nonnegative random variables H and G, G is dominated by H in the* Laplace transform (Lt) *order, denoted as $G \prec_{Lt} H$, if $\forall \, s > 0$, $\mathbb{E}e^{-sG} \geq \mathbb{E}e^{-sH}$.*

Remark 5.1. *Consider a horizon H, having finite expectation μ and variance σ^2, and random variables $B \sim \mathrm{Ber}\left(\frac{\mu^2}{\mu^2+\sigma^2}\right)$ and $G := \frac{\mu^2+\sigma^2}{\mu}B$. Then $G \prec_{Lt} H$.*

We are now ready to prove Theorem 1.4. The details of the proof are provided in Appendix D.

Idea of the proof of Theorem 1.4. Recall that the single-threshold 2-approximation of Theorem 1.1 obtains at least competitive ratio $c_{p,\bar{q}} = 1 - \mathbb{E}\left[(1 - 1/\mu)^H\right] = 1 - \mathbb{E}\left(e^{-\bar{s}H}\right)$ with $\bar{s} := -\log(1 - 1/\mu)$. Since $G \prec_{Lt} H$, where G is the two-point distribution of Remark 5.1, it follows that

$$c_{p,\bar{q}} \geq 1 - \mathbb{E}\left(e^{-\bar{s}G}\right) = \frac{\mu^2}{\mu^2 + \sigma^2}\left[1 - \left(1 - \frac{1}{\mu}\right)^{\frac{\mu^2+\sigma^2}{\mu}}\right]. \tag{5.1}$$

As a result, an inequality that yields a single-threshold $2 - 1/\mu$-approximation is

$$\frac{\mu^2}{\mu^2 + \sigma^2}\left[1 - \left(1 - \frac{1}{\mu}\right)^{\frac{\mu^2+\sigma^2}{\mu}}\right] \geq \frac{1}{2 - \frac{1}{\mu}}, \tag{5.2}$$

and straightforward estimates show that this is ensured by $e^{-x} \leq 1 - \frac{x}{2 - 1/\mu}$, having defined $x := 1 + \sigma^2/\mu^2$. Defining constraints $y := x - (2 - 1/\mu) > -1$ and $-2 < \bar{y} := -(2 - 1/\mu) < -1$, we can rewrite the condition above as $ye^y \leq \bar{y}e^{\bar{y}}$. Under the constraints given, this is more explicitly rewritten as $y \leq W_0(\bar{y}e^{\bar{y}})$, where W_0 denotes the principal branch of the Lambert function [16] . Recalling the definition of y and x, this yields (1.1). $\qquad\square$

On a practical note, the Lambert function is readily simulated in various languages.[17] Thus it is possible to reliably implement the substitution of suitable values of $C > 1$ for $2 - 1/\mu$ in (1.1), even without large-market hypothesis (the latter was used to heuristically exemplify this fact in Section 1.3). The only accuracy issue would arise near the branching point of the Lambert function, that is, when $\bar{y} \approx -1$, which is an uninteresting range for RH, since it only concerns horizons with $2 - 1/\mu \approx 1$, that is $\mu \approx 1$, which approaches a degenerate case at which the gambler and the prophet perform equally, since there is only one item to allocate. With this caveat, we conclude by discussing how to derive Corollary 1.1.

As direct consequence of substituting *admissible* values of $C > 1$ for $2 - 1/\mu$ in (5.2) we obtain the upper-bound

$$\mathrm{CV} \leq \sqrt{1 - 1/\mu + W_0(-Ce^{-C})} = \sqrt{C - 1 + W_0(-Ce^{-C})},$$

---

[16] Recall that the Lambert function is the solution to the equation $ye^y = z$ over the complex numbers, and thus a function with branches. Over $\mathbb{R}$ the Lambert function is described through a principal branch, denoted as $W_0(z)$, and a secondary branch, denoted as $W_{-1}(z)$, since, as evident from the graph of the function ye^y on $(-\infty, 0]$, $y = -1$ is a branching point for the solution. $W_{-1}(z)$ takes care of the branch corresponding to $y \leq -1$, $W_0(z)$ of the one corresponding to $y \geq -1$.

[17] For example in Python, with standard error bounds of 10^{-8}, via scipy.special.lambertw. Higher precision alternatives also exist (for example through the bisection method).

which ensures a C-approximation. This highlights interesting degenerate cases, and establishes an important connection with the existing literature. Note that the deterministic case ($\mathrm{CV} = 0$) can be trivially recovered directly from (5.1), which can be reformulated as a guarantee that the competitive ratio (given by $1/C$ here) is always greater than

$$\frac{1 - \left(1 - \mu^{-1}\right)^{\mu(1+\mathrm{CV}^2)}}{1 + \mathrm{CV}^2} \geq \frac{1 - e^{-(1+\mathrm{CV}^2)}}{1 + \mathrm{CV}^2}.$$

Therefore if $\mathrm{CV} = 0$, the optimal single-threshold algorithm exploited ensures a competitive ratio of $1 - 1/e$. This value is the global maximum of the function on the right-hand side, and coincides with the optimal competitive ratio for single-threshold algorithms for (deterministic) IID PI [17][Theorem 21]. Note that the left-hand side tends to the right-hand side in the large-market limit, at which regime the admissible values of C aforementioned are easily computed to be constrained to be larger than $e/(e-1)$, and therefore a competitive ratio of $1 - 1/e$ cannot be exceeded even in a best-case analysis (with respect to the horizon). More in general, dropping the large-market limit, the left-hand side also achieves global maximum $1 - (1 - 1/\mu)^\mu$ at $\mathrm{CV} = 0$ (this is not surprising, since the deterministic horizon is no harder than the random horizon for the gambler). Hence C will always be constrained to be greater than $[1 - (1 - 1/\mu)^\mu]^{-1}$. This proves Corollary 1.1 and justifies the admissibility condition.

6   Conclusion

The contributions of this work can be divided into two main groups. The first group of results (in Sections 3 and 5) extend the distributional classes to which single-threshold algorithms can be applied, which impacts areas such as pricing and mechanism design of online auctions for the following reasons. The classes of distributions in $\mathcal{G}$ are well established across several areas, in particular in reliability theory. Furthermore, horizons classified in terms of concentration measures such as the CV can be of wide use in applications, since their classification only requires the first two moments. Single-threshold algorithms are simple to implement, and thus proving their competitiveness and optimality for as many distributional classes as possible is advantageous. Although our main focus is on results for single item RH, they can be generalised to any number of items via the mechanism designed in [2]. Furthermore, it is likely that similar results related to optimal stopping problems with IHR random horizons may benefit from a similar extension to the $\mathcal{G}$ class. The key to this extension is clear from the fact that our proof, in essence, relies only on certain stochastic ordering properties of the IHR class, which is popular not only in reliability theory, but in economics and computer science. RH has several connections to interesting online stochastic matching problems [3]. A potential implementation of the techniques we established in this work would be extending the use of stochastic orders to these models and see if they improve known guarantees.

The second group of results (in Sections 2 and 4) provide a more theoretical contribution. On one end, extending the characterisation of the optimal algorithm to unbounded horizons is of help in formalising hardness results. On the other hand, it has also helped significantly in showing that there are instances which are hard for single-thresholds, but for which a multiple-thresholds algorithm such as the SP rule is competitive. This last fact, combined with the fact that single-threshold algorithms already fail on the $\overline{\mathcal{G}}$ class, suggests that the study of RH under (most likely adaptive) multiple-threshold algorithms is a new and promising area of research.

Acknowledgements.

References

[1] A. R. Abdel-Hamid, J. A. Bather, and G. B. Trustrum. The secretary problem with an unknown number of candidates. *J. Appl. Probab.*, 19(3):619–630, 1982.

[2] R. Alijani, S. Banerjee, S. Gollapudi, K. Munagala, and K. Wang. Predict and match: prophet inequalities with uncertain supply. *Proc. ACM Meas. Anal. Comput. Syst.*, 4(1):Article 4, 2020.

[3] A. Aouad and W. Ma. A nonparametric framework for online stochastic matching with correlated arrivals. In *EC*, page 114. ACM, 2023.

[4] P. Billingsley. *Weak convergence of measures: applications in probability.* Society for Industrial Mathematics, 1987.

[5] T. Bojdecki. On optimal stopping of a sequence of independent random variables - probability maximizing approach. *Stoch. Process. Their Appl.*, 6:153–163, 1978.

[6] C. Bracquemond, D. Roy, and M. Xie. On some discrete notions of aging. In *System and bayesian reliability. Essays in honor of Prof. E. Barlow on his 70th birthday*, pages 185–197. World Scientific, 2001.

[7] A. Bubna and A. Chiplunkar. Prophet inequality: order selection beats random order. In *EC*, pages 302–336. ACM, 2023.

[8] Y. Cai and J. Wu. On the optimal fixed-price mechanism in bilateral trade. In *STOC*, pages 737–750, 2023.

[9] Z. Chen, Z. Huang, D. Li, and Z. G. Tang. Prophet secretary and matching: the significance of the largest item. In *SODA*, pages 1371–1401, 2025.

[10] Y. S. Chow, S. Moriguti, H. Robbins, and S. M. Samuels. Optimal selection based on relative ranks (the "secretary problem"). *Isr. J. Math.*, 2:81–90, 1964.

[11] Y. S. Chow and H. Robbins. On optimal stopping rules. *Z. Wahrscheinlichkeitstheorie*, 2:33–49, 1963.

[12] Y. S. Chow, H. Robbins, and D. Siegmund. *Great expectations: the theory of optimal stopping.* Houghton Mifflin, 1971.

[13] J. Correa, P. Foncea, R. Hoeksma, R. Oosterwijk, and T. Vredeveld. Posted price mechanisms for a random stream of customers. In *EC*, pages 169–189. ACM, 2017.

[14] J. Correa, P. Foncea, R. Hoeksma, R. Oosterwijk, and T. Vredeveld. Recent developments in prophet inequalities. *SIGecom Exch.*, 17(1):61–70, 2018.

[15] J. Correa, P. Foncea, D. Pizarro, and V. Verdugo. From pricing to prophets, and back! *Oper. Res. Lett.*, 47:25–29, 2019.

[16] E. B. Dynkin. The optimal choice of the stopping moment for a Markov process. *Dokl. Akad. Nauk. SSSR*, 150:238–240, 1963.

[17] S. Eshani, M. Hajiaghayi, M. Kesselheim, and S. Singla. Prophet secretary for combinatorial auctions and matroids. In *SODA*, pages 700–714. SIAM, 2018.

[18] P. R. Freeman. The secretary problem and its extensions: a review. *Int. Stat. Rev.*, 51(2):189–206, 1983.

[19] G. Giambartolomei, F. Mallmann-Trenn, and S. Saona. Prophet inequalities: Separating Random Order from Order Selection. arXiv:2304.04024 [cs.DS].

[20] J. Gianini-Pettitt. Optimal selection based on relative ranks with a random number of individuals. *Adv. Appl. Probab.*, 11(4):720–736, 1979.

[21] J. P. Gilbert and F. Mosteller. Recognizing the maximum of a sequence. *J. Am. Stat. Assoc.*, 61(313):35–73, 1966.

[22] M. T. Hajiaghayi, R. D. Kleinberg, and T. Sandholm. Automated online mechanism design and prophet inequalities. In *AAAI*, pages 58–65, 2007.

[23] S. Har-Peled, E. Harb, and V. Livanos. Oracle-augmented prophet inequalities. In *ICALP*, pages 81:1–81:19, 2024.

[24] T. P. Hill and R. P. Kertz. Stop rule inequalities for uniformly bounded sequences of random variables. *Trans. Am. Math. Soc.*, 278(1):197–207, 1983.

[25] T. P. Hill and R. P. Kertz. A survey of prophet inequalities in optimal stopping theory. *Contemp. Math.*, 125:191–207, 1992.

[26] M. Hoefer and B. Kodric. Combinatorial secretary problems with ordinal information. In *ICALP*, pages 133:1–133:14, 2017.

[27] R. P. Kertz. Stop rule and supremum expectations of i.i.d. random variables: a complete comparison by conjugate duality. *J. Multivar. Anal.*, 19(1):88–112, 1986.

[28] B. Klefsjö. The hnbue and hnwue classes of life distributions. *Naval Res. Logist. Quart.*, 29:331–344, 1982.

[29] B. Klefsjö. Testing exponentiality against hnbue. *Scand. J. Stat.*, 10(2):65–75, 1983.

[30] B. Klefsjö. A useful ageing property based on the laplace transform. *J. Appl. Prob.*, 20:615–626, 1983.

[31] U. Krengel and L. Sucheston. Semiamarts and finite values. *Bull. Amer. Math. Soc.*, 83(4):745–747, 1977.

[32] U. Krengel and L. Sucheston. On semiamarts, amarts, and processes with finite value. *Adv. Probab. Related Topics*, 4:197–266, 1978.

[33] N. A. Langberg, R. V. Léon, J. Lynch, and F. Proschan. Extreme points of the class of discrete decreasing failure life distributions. *Math. Oper. Res.*, 5(1):35–42, 1980.

[34] D. V. Lindley. Dynamic programming and decision theory. *J. R. Stat. Soc. Ser. C Appl. Stat.*, 10(1):39–51, 1961.

[35] B. Lucier. An economic view of prophet inequalities. *SIGecom Exch.*, 16(1):24–47, 2017.

[36] M. Mahdian and A. Saberi. Multi-unit auctions with unknown supply. In *CEC*, pages 243–249. ACM, 2006.

[37] S. Oveis Gharan and J. Vondrák. On variants of the matroid secretary problem. *Algorithmica*, 67:472,497, 2013.

[38] Z. Poroziński. The full-information best choice problem with a random number of observations. *Stoch. Process. Their Appl.*, 24:293–307, 1987.

[39] E. L. Presman and I. M. Sonin. The best choice problem for a random number of objects. *Theory Probab. its Appl.*, 17(4):657–668, 1972.

[40] T. Rolski. Mean residual life. *Bull. Int. Statist. Inst.*, 46:220–266, 1975.

[41] E. Samuel-Cahn. Comparisons of threshold stop rule and maximum for independent non-negative random variables. *Ann. Probab.*, 12(4):1213–1216, 1984.

[42] E. Samuel-Cahn. The best-choice secretary problem with random freeze on jobs. *Stoch. Process. Their Appl.*, 55(2):315–327, 1995.

[43] E. Samuel-Cahn. Optimal stopping with random horizon with application to the full-information best-choice problem with random freeze. *J. Am. Stat. Assoc.*, 91(433):357–364, 1996.

[44] T. Sandholm. Automated mechanism design: A new application area for search algorithms. In *CP*, pages 19–36, 2003.

[45] M. Shaked and J. G. Shanthikumar. *Stochastic orders*. Springer, 2007.

Supplementary materials

A Proofs omitted from Section 2

A.1 Proof of Theorem 2.1

Proof of Theorem 2.1. If $m < \infty$, without loss of generality we can trivially reduce the discounted problem $\mathbf{Z}^{(m)}$, that is $Z_0, Z_1, \ldots, Z_m, 0, \ldots$, to the finite sequence $\mathbf{Z}^m$, that is $Z_0, Z_1, \ldots, Z_m$ with filtration $\mathcal{B}^m := \{\mathcal{B}_0, \mathcal{B}_1, \ldots, \mathcal{B}_m\}$. Thus by [43, Theorem 2.1], which uses backward induction [12, Theorem 3.2] to construct $\bar{\zeta}$, it follows that the natural candidate $\bar{\tau} := \bar{\zeta} \wedge (H+1)$ yields $\mathcal{V}(\mathbf{Y})$.

Assume $m = \infty$. We start by constructing, for every stopping rule σ for the problem relative to $\mathbf{Y}$, a stopping rule ζ for the problem relative to $\mathbf{Z}$, such that $\mathbb{E}Y_\sigma = \mathbb{E}Y_\tau$, where $\tau := \zeta \wedge (H+1)$. It will be useful to rewrite $\mathcal{F}_i = \sigma(\mathcal{B}_i, \mathcal{H}_i)$, where $\mathcal{H}_i := \sigma(\{H = 0\}, \ldots, \{H = i-1\})$. The construction will require a suitable decomposition of the probability space Ω. For every $i \in \mathbb{N}$, let $E_i := \{\sigma = i\} \cap \{H \geq i\}$. Since $\{\sigma = i\} \in \mathcal{F}_i$, we have that $E_i \in \mathcal{F}_i \cap \{H \geq i\} = \mathcal{B}_i \cap \{H \geq i\}$, where the intersections denote the corresponding *trace σ-algebras*. The equality is seen as follows: since $\mathcal{F}_i = \sigma(\mathcal{B}_i, \mathcal{H}_i) := \sigma(A \cap B, A \in \mathcal{B}_i, B \in \mathcal{H}_i)$, we have that $\mathcal{F}_i \cap \{H \geq i\} := \sigma(A \cap B \cap \{H \geq i\}, A \in \mathcal{B}_i, B \in \mathcal{H}_i)$. By definition, for every $B \in \mathcal{H}_i$, $B \cap \{H \geq i\}$ is either $\emptyset$ or $\{H \geq i\}$. Thus the only nonempty intersections are $A \cap B \cap \{H \geq i\} = A \cap \{H \geq i\}$.

Since $E_i \in \mathcal{B}_i \cap \{H \geq i\}$, by definition there exists $A_i \in \mathcal{B}_i$ such that $E_i = A_i \cap \{H \geq i\}$. Note that the events $\{\sigma = i\}$ constitute a partitioning of Ω, while $\{H \geq i\} \supseteq \{H \geq i+1\}$. Furthermore, without loss of generality, the events $\{A_i\}$ can be assumed mutually disjoint. This can be seen by induction: suppose we have $E_i = A_i \cap \{H \geq i\}$ and $E_{i+1} = A'_{i+1} \cap \{H \geq i+1\}$, with $A'_{i+1} \cap A_i \neq \emptyset$. Since $E_i \subseteq \{\sigma = i\}$ and $E_{i+1} \subseteq \{\sigma = i+1\}$, $E_i \cap E_{i+1} = \emptyset$. Thus $\emptyset = (A_i \cap \{H \geq i\}) \cap (A'_{i+1} \cap \{H \geq i+1\}) \supseteq (A_i \cap \{H \geq i+1\}) \cap (A'_{i+1} \cap \{H \geq i+1\})$. This means that the parts of A'_{i+1} and A_i, which intersect, lie outside $\{H \geq i+1\}$. Define $A_{i+1} := A'_{i+1} \setminus A_i$. Since A_{i+1} is only missing a part of A'_{i+1} that lies outside $\{H \geq i+1\}$, we also have $E_{i+1} = A_{i+1} \cap \{H \geq i+1\}$. Furthermore, $A_{i+1} = A'_{i+1} \cap A_i^c \in \mathcal{B}_{i+1}$, because $A'_{i+1} \in \mathcal{B}_{i+1}$ and $A_i \in \mathcal{B}_i \subseteq \mathcal{B}_{i+1}$. Thus we replace A'_{i+1} with A_{i+1} and the induction step is complete. Starting with $i = 1$, and performing inductively this substitution whenever necessary, yields the mutually disjoint $\{A_i\}$.

We can finally construct ζ as, informally speaking, a stopping rule that stops when σ successfully stops by the time the horizon has realised, and does not stop otherwise. The precise measure theoretic construction is as follows: given the mutually disjoint $\{A_i\}$ as constructed above, recall that since $A_i \in \mathcal{B}_i$, we can denote $C := \bigcap_{i \in \mathbb{N}} A_i^c = \left(\bigcup_{i \in \mathbb{N}} A_i\right)^c \in \mathcal{B}_\infty$, therefore we have that

$$\zeta(\omega) := \begin{cases} i, & \omega \in A_i, \forall i \in \mathbb{N} \\ \infty, & \omega \in C \end{cases} \tag{A.1}$$

is a valid stopping rule in $\mathcal{T}$. Moreover, $\tau := \zeta \wedge (H+1)$ is a valid finite stopping rule in $\mathcal{T}^*$, since for every $i \in \mathbb{N}$, both $\{\zeta = i\}$ and $\{H+1 = i\} = \{H = i-1\}$ are adapted to $\mathcal{F}_i = \sigma(\mathcal{B}_i, \mathcal{H}_i)$, that is both ζ and $H+1$ are valid stopping rules for the problem, and the minimum of stopping rules is a stopping rule.

The stopping rules σ and τ are equivalent, that is $\mathbb{E}Y_\sigma = \mathbb{E}Y_\tau$. In fact, for every $i \in \mathbb{N}$, $\mathbb{E}Y_i \mathbb{1}_{\{\sigma=i, H<i\}} = 0 = \mathbb{E}Y_i \mathbb{1}_{\{\tau=i, H<i\}}$ by definition of Y_i, which is null on $\{H < i\}$. We also have that $\mathbb{E}Y_\infty = 0$. Following the standard convention that

$$\mathbb{E}Y_\sigma := \mathbb{E}\sum_{i \in \mathbb{N}} Y_i \mathbb{1}_{\{\sigma=i\}} + \mathbb{E}Y_\infty,$$

the equivalence follows from showing that for every $i \in \mathbb{N}$, $\mathbb{E}Y_i \mathbb{1}_{\{\sigma=i, H \geq i\}} = \mathbb{E}Y_i \mathbb{1}_{\{\tau=i, H \geq i\}}$, since the expectation can be exchanged with the summation in our setting (as it can be seen in the

next paragraph, it is enough to use *Fubini's theorem*). By construction of the events $\{A_i\}$ and the stopping rules ζ and τ, we have that, indeed,

$$\mathbb{E}Y_i\mathbb{1}_{\{\sigma=i,\,H\geq i\}} = \mathbb{E}Y_i\mathbb{1}_{A_i\cap\{H\geq i\}} = \mathbb{E}Y_i\mathbb{1}_{\{\zeta=i,\,H\geq i\}} = \mathbb{E}Y_i\mathbb{1}_{\{\tau=i,\,H\geq i\}}.$$

We show first that $\mathcal{V}(\mathbf{Y}) = \mathcal{V}(\mathbf{Z})$. First note that, by the construction of the stopping rule $\zeta \in \mathcal{T}$ in (A.1) in terms of the stopping rule $\sigma \in \mathcal{T}^*$, we have that, since $Y_\infty = 0$,

$$Y_\sigma = \sum_{i\in\mathbb{N}} X_i\mathbb{1}_{\{H\geq i\}}\mathbb{1}_{\{\sigma=i\}} = \sum_{i\in\mathbb{N}} X_i\mathbb{1}_{E_i} = \sum_{i\in\mathbb{N}} X_i\mathbb{1}_{A_i\cap\{H\geq i\}} = \sum_{i\in\mathbb{N}} X_i\mathbb{1}_{\{H\geq i\}}\mathbb{1}_{\{\zeta=i\}},$$

thus yielding

$$\begin{aligned}
\mathbb{E}Y_\sigma &= \sum_{i\in\mathbb{N}} \mathbb{E}X_i\mathbb{1}_{\{H\geq i\}}\mathbb{1}_{\{\zeta=i\}} = \sum_{i\in\mathbb{N}} \mathbb{E}X_i\mathbb{1}_{\{\zeta=i\}}\mathbb{E}_{\mathcal{B}_i}\mathbb{1}_{\{H\geq i\}} \\
&= \sum_{i\in\mathbb{N}} S(i)\mathbb{E}X_i\mathbb{1}_{\{\zeta=i\}} = \mathbb{E}\sum_{i\in\mathbb{N}} S(i)X_i\mathbb{1}_{\{\zeta=i\}} = \mathbb{E}Z_\zeta.
\end{aligned} \tag{A.2}$$

Here the expectation can be exchanged with the summation since

$$\sum_{i\in\mathbb{N}} S(i)\mathbb{E}X_i\mathbb{1}_{\{\zeta=i\}} \leq \nu\sum_{i\in\mathbb{N}} S(i) = \nu\mu < \infty.$$

Secondly we note that, by a similar computation, for any stopping rule $\zeta \in \mathcal{T}$, $\tau := \zeta \wedge (H+1)$ is such that

$$\begin{aligned}
\mathbb{E}Z_\zeta &= \mathbb{E}\sum_{i\in\mathbb{N}} S(i)X_i\mathbb{1}_{\{\zeta=i\}} = \sum_{i\in\mathbb{N}} \mathbb{E}X_i\mathbb{1}_{\{\zeta=i\}}\mathbb{E}_{\mathcal{B}_i}\mathbb{1}_{\{H\geq i\}} = \mathbb{E}\sum_{i\in\mathbb{N}} X_i\mathbb{1}_{\{\zeta=i\}}\mathbb{1}_{\{H\geq i\}} \\
&= \mathbb{E}\sum_{i\in\mathbb{N}} X_i\mathbb{1}_{\{\tau=i\}}\mathbb{1}_{\{H\geq i\}} = \mathbb{E}\sum_{i\in\mathbb{N}} Y_i\mathbb{1}_{\{\tau=i\}} = \mathbb{E}Y_\tau.
\end{aligned} \tag{A.3}$$

These two facts together imply that it is not possible to have $\sup_{\sigma\in\mathcal{T}^*}\mathbb{E}Y_\sigma \neq \sup_{\zeta\in\mathcal{T}}\mathbb{E}Z_\zeta$. Indeed, we cannot have that $\sup_{\sigma\in\mathcal{T}^*}\mathbb{E}Y_\sigma > \sup_{\zeta\in\mathcal{T}}\mathbb{E}Z_\zeta$, because otherwise there would be some $\sigma \in \mathcal{T}^*$ such that $\mathbb{E}Y_\sigma > \mathbb{E}Z_\zeta$ for all $\zeta \in \mathcal{T}$, but this is not possible since by (A.2) for every $\sigma \in \mathcal{T}^*$ there is a $\zeta \in \mathcal{T}$ such that $\mathbb{E}Y_\sigma = \mathbb{E}Z_\zeta$. We cannot have that $\sup_{\sigma\in\mathcal{T}^*}\mathbb{E}Y_\sigma < \sup_{\zeta\in\mathcal{T}}\mathbb{E}Z_\zeta$ either, because otherwise there would be some $\zeta \in \mathcal{T}$ such that $\mathbb{E}Y_\sigma < \mathbb{E}Z_\zeta$ for all $\sigma \in \mathcal{T}^*$, but this is not possible since by (A.3) for every $\zeta \in \mathcal{T}$, $\zeta\wedge(H+1) := \tau \in \mathcal{T}^*$ is such that $\mathbb{E}Y_\tau = \mathbb{E}Z_\zeta$. Therefore $\mathcal{V}(\mathbf{Y}) = \mathcal{V}(\mathbf{Z})$ follows by definition.

Given an optimal stopping rule $\bar{\zeta} \in \mathcal{T}$, we have that $\bar{\tau} := \bar{\zeta} \wedge (H+1)$ provides an optimal stopping rule in $\mathcal{T}^*$ by (A.3) with $\zeta = \bar{\zeta}$ and $\tau = \bar{\tau}$. Since

$$\mathbb{E}\sup_{i\in\mathbb{N}} Z_i \leq \mathbb{E}\sum_{i\in\mathbb{N}} Z_i = \sum_{i\in\mathbb{N}} S(i)\mathbb{E}X_i = \mu\nu < \infty$$

and $\limsup_{i\to\infty} Z_i = Z_\infty = 0$, the Snell envelope of $\mathbf{Z}$, denoted $\mathbf{V} := \{V_i\}_{i\in\mathbb{N}_0}$, satisfies (2.1) by Remark 2.1 and yields the optimal stopping rule

$$\bar{\zeta} = \inf\{i \in \mathbb{N}_0 : Z_i \geq \mathbb{E}_{\mathcal{B}_i}V_{i+1}\}$$

by Remark 2.2. $\qquad\square$

A.2 Proof of Lemma 2.1

Proof of Lemma 2.1. Note that $\operatorname{ess\,sup}_{\zeta\in\mathcal{T}_{i+1}}\mathbb{E}_{\mathcal{B}_i}Z_\zeta \leq \mathbb{E}_{\mathcal{B}_i}V_{i+1}$ is straightforward. By definition, $\mathbb{E}_{\mathcal{B}_{i+1}}Z_\zeta \leq V_{i+1}$ a.s. for every $\zeta \in \mathcal{T}_{i+1}$. Thus, taking expectations, $\mathbb{E}_{\mathcal{B}_i}Z_\zeta \leq \mathbb{E}_{\mathcal{B}_i}V_{i+1}$ a.s. for every $\zeta \in \mathcal{T}_{i+1}$.

The argument consists in finding a lower bound to $\operatorname{ess\,sup}_{\zeta\in\mathcal{T}_{i+1}}\mathbb{E}_{\mathcal{B}_i}Z_\zeta$ which we then show to converge to $\mathbb{E}_{\mathcal{B}_i}V_{i+1}$, yielding the claim by the *squeeze theorem*. This is done by exploiting *regular stopping rules*. We briefly recall a few standard results regarding these, which will be used. Let $i \in \mathbb{N}_0$.

i) A stopping rule $\zeta \in \mathcal{T}_i$ is *regular from i on*, if for every $k \geq i$, $\mathbb{E}_{\mathcal{B}_k} Z_\zeta > Z_k$ a.s. on $\{\zeta > k\}$.

ii) If $\mathbb{E} \sup_{i \in \mathbb{N}_0} Z_i < \infty$, for any stopping rule $\zeta \in \mathcal{T}_i$ there exists a stopping rule $\zeta' \in \mathcal{T}_i$ regular from i on, such that $\mathbb{E}_{\mathcal{B}_i} Z_{\zeta'} \geq \mathbb{E}_{\mathcal{B}_i} Z_\zeta$ a.s.

iii) If $\mathbb{E} \sup_{i \in \mathbb{N}_0} Z_i < \infty$ and both stopping rules $\zeta \in \mathcal{T}_i$ and $\zeta' \in \mathcal{T}_i$ are regular from i on, then so is $\zeta'' := \zeta \vee \zeta'$ and $\mathbb{E}_{\mathcal{B}_i} Z_{\zeta''} \geq \mathbb{E}_{\mathcal{B}_i} Z_\zeta \vee \mathbb{E}_{\mathcal{B}_i} Z_{\zeta'}$ a.s.

By a standard property of the essential supremum, there is a sequence of stopping rules in $\mathcal{T}_{i+1}$, $\{\zeta_k\}_{k \in \mathbb{N}}$, such that

$$V_{i+1} = \sup_{k \in \mathbb{N}} \mathbb{E}_{\mathcal{B}_{i+1}} Z_{\zeta_k} \uparrow \max_{i \leq j \leq k} \mathbb{E}_{\mathcal{B}_{i+1}} Z_{\zeta_j} \tag{A.4}$$

as k grows. By (ii), to every stopping rule ζ_j it corresponds a ζ'_j regular from $i+1$ on, such that $\mathbb{E}_{\mathcal{B}_i} Z_{\zeta'} \geq \mathbb{E}_{\mathcal{B}_i} Z_\zeta$. Let $\zeta''_k := \max_{j \in [k]} \zeta'_j \in \mathcal{T}_{i+1}$. By (iii), $\mathbb{E}_{\mathcal{B}_{i+1}} Z_{\zeta''_k} \geq \max_{j \in [k]} \mathbb{E}_{\mathcal{B}_{i+1}} Z_{\zeta'_j}$. Then

$$\operatorname*{ess\,sup}_{\zeta \in \mathcal{T}_{i+1}} \mathbb{E}_{\mathcal{B}_i} Z_\zeta \geq \mathbb{E}_{\mathcal{B}_i} Z_{\zeta''_k} = \mathbb{E}_{\mathcal{B}_i} \mathbb{E}_{\mathcal{B}_{i+1}} Z_{\zeta''_k} \geq \mathbb{E}_{\mathcal{B}_i} \max_{j \in [k]} \mathbb{E}_{\mathcal{B}_{i+1}} Z_{\zeta'_j} \geq \mathbb{E}_{\mathcal{B}_i} \max_{j \in [k]} \mathbb{E}_{\mathcal{B}_{i+1}} Z_{\zeta_j} \longrightarrow \mathbb{E}_{\mathcal{B}_i} V_{i+1} \text{a.s.}$$

as k grows, where the limit follows by (A.4) and the *monotone convergence theorem*. $\qquad\square$

B Proofs omitted from Section 3

B.1 Proof of Lemma 3.1

Proof of Lemma 3.1. Consider the following decomposition. For every $h \in \mathbb{N}$, let

$$E_\pi^h := \{H = h\} \cap \bigcup_{i \in [h]} \{i = \inf\{j \in [h] : X_j \geq \pi\}\},$$

$$E_0^h := \{H = h\} \cap \bigcap_{i \in [h]} \{X_i < \pi\}.$$

Then $\{H = h\} = E_\pi^h \cup E_0^h$. Intuitively, E_π^h is the part of the event $\{H = h\}$ on which the algorithm τ_π stops, achieving at least π, while E_0^h is the part on which it fails to stop, achieving zero. By partitioning the probability space $\Omega = \bigcup_{h \in \mathbb{N}}(E_\pi^h \cup E_0^h)$ we can rewrite, exploiting that on E_0^h, $Y_{\tau_\pi} = Y_\infty = 0$,

$$\mathbb{E} Y_{\tau_\pi} = \sum_{h \in \mathbb{N}} \mathbb{E}\left(Y_{\tau_\pi} \mathbb{1}_{E_\pi^h}\right).$$

The exchange of expectation and summation holds *a posteriori*, since we will obtain a finite quantity for the right-hand side. By the independence of H and X,

$$\mathbb{E}\left(Y_{\tau_\pi} \mathbb{1}_{E_\pi^h}\right) = \mathbb{E}\left(Y_{\tau_\pi} \mathbb{1}_{\{H=h\}} \sum_{i \in [h]} \mathbb{1}_{\{i=\inf\{j \in [h]: X_j \geq \pi\}\}}\right) = \sum_{i \in [h]} \mathbb{E}\left(X_i \mathbb{1}_{\{H=h\}} \mathbb{1}_{\{i=\inf\{j \in [h]: X_j \geq \pi\}\}}\right)$$

$$= \mathbb{P}(H = h) \sum_{i \in [h]} \mathbb{E}\left(X_i \mathbb{1}_{\{X_i \geq \pi\}} \prod_{j \in [i-1]} \mathbb{1}_{\{X_j < \pi\}}\right)$$

$$= \mathbb{P}(H = h) \sum_{i \in [h]} \mathbb{P}^{i-1}(X < \pi) \mathbb{E}\left(X_i \mathbb{1}_{\{X_i \geq \pi\}}\right).$$

Since $\mathbb{E}\left(X_i \mathbb{1}_{\{X_i \geq \pi\}}\right) = \mathbb{E}(X | X \geq \pi) \mathbb{P}(X \geq \pi)$ for all $i \in \mathbb{N}$ and

$$\sum_{i \in [h]} \mathbb{P}^{i-1}(X < \pi) = \frac{1 - \mathbb{P}^h(X < \pi)}{\mathbb{P}(X \geq \pi)},$$

it follows that $\mathbb{E}\left(X_{\tau_\pi}\mathbb{1}_{E_\pi^h}\right) = \mathbb{E}(X|X \geq \pi)[1 - \mathbb{P}^h(X < \pi)]\mathbb{P}(H = h)$. Hence

$$\mathbb{E}X_{\tau_\pi} = \mathbb{E}Y_{\tau_\pi} = \mathbb{E}(X|X \geq \pi)\sum_{h\in\mathbb{N}}[1 - \mathbb{P}^h(X < \pi)]\mathbb{P}(H = h) = \mathbb{E}(X|X \geq \pi)\mathbb{E}[1 - \mathbb{P}^H(X < \pi)],$$

which yields the result. □

B.2 Proof of Lemma 3.2 and discussion on amendments to [2]

Before giving the details of the proof of Lemma 3.2, we further discuss the need for a variational argument to upper-bound $\mathbb{E}M_H$ for any continuous X. The key idea at the base of this argument is similar to that of [2, Theorem 3.2], but they do not state what restrictions their argument imposes on X. As a matter of fact, they achieve the upper bound only for X with discrete[18] distribution and admitting a value p such that $\mathbb{P}(X \geq p) = 1/\mu$, where $\mu = \mathbb{E}H$. If their claim was true as stated, the resulting algorithm would not have a consistently defined threshold for more general distributions. It is not clear how these results should be extended from this special class to a more general one, nor it is mentioned by the authors. Perhaps, they have in mind discretised continuous distributions. Yet, this would mean that by some approximation or bounding argument, their result only applies to continuous distributions. We could not find a straightforward way, which might suggest this type of argument as trivial. Examples from the literature on other stochastic optimisation problems, such as the bilateral trade problem, adopt a similar approach, of discretising and then bounding, but the proofs of the bounding step are given and usually nontrivial. See for example [8].

Proof of Lemma 3.2. Let $f(x)$ be the pdf and $F(x)$ the cdf of M_H. The pdf exists by absolute continuity of X and $f = F'$. Let $v(x)$ be the pdf and $V(x)$ the cdf of the value X. We are seeking a upper bound on $\mathbb{E}M_H = \int_0^\infty xf(x)dx =: I(f)$, which will therefore be the functional to be maximised over $f \in \mathrm{L}_+^1([0,\infty))$, the set of nonnegative Lebesgue-integrable functions on $[0,\infty)$. Since f is a density supported on the nonnegative reals, we also have a first constraint $\int_0^\infty f(x)dx = 1$ for the maximisation problem. Furthermore the law of total probability holds. Thus for every $x \in [0,\infty)$, as $\varepsilon \longrightarrow 0$,

$$f(x) \longleftarrow \frac{\mathbb{P}(x < M_H \leq x + \varepsilon)}{\varepsilon} = \sum_{h\in\mathbb{N}}\frac{\mathbb{P}(x < M_h \leq x + \varepsilon)}{\varepsilon}\mathbb{P}(H = h)$$

$$\leq \sum_{h\in\mathbb{N}}h\frac{\mathbb{P}(x < X \leq x + \varepsilon)}{\varepsilon}\mathbb{P}(H = h) \longrightarrow \mu v(x),$$

which implies that a second constraint holds: $0 \leq f(x) \leq \mu v(x)$ for all $x \in [0,\infty)$. The inequality above can be shown as follows. Denote by $F_h(x)$ the conditional cdf of M_H, given the event $\{H = h\}$, for all $h \in \mathbb{N}$ such that $\mathbb{P}(H = h) > 0$. Then for all $x \in (0,\infty)$,

$$\mathbb{P}(x < M_h \leq x + \varepsilon) = F_h(x + \varepsilon) - F_h(x) = V^h(x + \varepsilon) - V^h(x) = [V(x + \varepsilon) - V(x)]\cdot$$

$$\sum_{i=0}^{h-1}V^{h-i}(x + \varepsilon)V^i(x) \leq h[V(x + \varepsilon) - V(x)] = h\mathbb{P}(x < X < x + \varepsilon).$$

For $x = 0$, it follows trivially from $\mathbb{P}(0 < M_h \leq \varepsilon) = F_h(\varepsilon) = V^h(\varepsilon) \leq hV(\varepsilon) = h\mathbb{P}(0 < X < \varepsilon)$.

---

[18]For the value distribution, unlike for horizons, we use *discrete* in the standard, more general, sense.

In conclusion we seek to solve the constrained maximisation variational problem

$$I(f) := \int_0^\infty x f(x) dx$$
$$\int_0^\infty f(x) dx = 1$$
$$0 \le f(x) \le \mu v(x), \ \forall \, x \in [0, \infty)$$
$$f \in L^1_+([0, \infty)). \tag{B.1}$$

The functional $I(f)$ is trivially maximised by choosing a function $\bar{f}$ that is as large as possible on as many large values as possible, since they contribute the most to the integral, that is

$$\bar{f}(x) = \begin{cases} \mu v(x), & x \ge p \\ 0, & 0 \le x < p, \end{cases}$$

where $p = \inf \left\{ y \in [0, \infty) : \int_y^\infty \mu v(x) dx = 1 \right\}$. Note that the condition is equivalent to

$$p = \inf \left\{ y \in [0, \infty) : \mathbb{P}(X \ge y) = 1/\mu \right\},$$

and since the survival function of X is assumed continuous, there is only one value y yielding equality, meaning that p is the only value such that $\mathbb{P}(X \ge p) = 1/\mu$. The maximal value of the variational problem is therefore

$$I(\bar{f}) = \int_p^\infty x \mu v(x) dx = \int_p^\infty x \frac{v(x)}{\mathbb{P}(X \ge p)} dx = \frac{\mathbb{E}\left(X \mathbb{1}_{\{X \ge p\}}\right)}{\mathbb{P}(X \ge p)} = \mathbb{E}(X | X \ge p).$$

$\square$

B.3 Proof of Lemma 3.3

Proof of Lemma 3.3. Consider the following decomposition. For every $h \ge 1$, let

$$E_\pi^h := \{H = h\} \cap \bigcup_{i \in [h]} \{i = \inf\{j \in [h] : X_j \ge \pi, \ B_j = 1\}\},$$
$$E_0^h := \{H = h\} \cap \bigcap_{i \in [h]} \{i = \inf\{j \in [h] : X_j \ge \pi, \ B_j = 1\}\}^c.$$

Then $\{H = h\} = E_\pi^h \cup E_0^h$. Intuitively, E_π^h is the part of the event $\{H = h\}$ on which the randomised algorithm τ_π stops, achieving at least π, while E_0^h is the part on which it fails to stop, achieving zero. By partitioning the probability space $\Omega = \bigcup_{h \in \mathbb{N}} (E_\pi^h \cup E_0^h)$ we can rewrite, exploiting that on E_0^h, $Y_{\tau_\pi} = Y_\infty = 0$,

$$\mathbb{E} Y_{\tau_\pi} = \sum_{h \in \mathbb{N}} \mathbb{E} Y_{\tau_\pi} \mathbb{1}_{E_\pi^h}.$$

The exchange of expectation and summation holds *a posteriori*, since we will obtain a finite quantity for the right-hand side.

From a combinatorial point of view, note that for every $i \in \mathbb{N}$,

$$\{i = \inf\{j \in [h] : X_j \ge \pi, \ B_j = 1\}\} = \bigcup_{k=1}^{2^{i-1}} E_k^i$$

30

having defined, informally,

$$E_1^i := \{X_1 < \pi, \ \ldots, \ X_{i-1} < \pi, \ X_i \geq \pi, \ B_i = 1\}$$
$$E_2^i := \{X_1 \geq \pi, \ B_1 = 0, \ X_2 < \pi \ \ldots, \ X_{i-1} < \pi, \ X_i \geq \pi, \ B_i = 1\}$$
$$\vdots$$
$$E_{2^{i-1}}^i := \{X_1 \geq \pi, \ B_1 = 0, \ \ldots, \ X_{i-1} \geq \pi, \ B_{i-1} = 0, \ X_i \geq \pi, \ B_i = 1\},$$

where the 2^{i-1} events run through all possible combinations of the first $i-1$ steps not allowing the randomised stopping rule to stop. For example, assuming $i \geq 3$ and denoting ${}^nC_k := \binom{n}{k}$,

- $E_1^i = E_{i-1}^i{}_{C_0}$ is the event where none of the first $i-1$ steps has a random variable surpassing the threshold π;
- $E_{i-1}^i{}_{C_0+1}, \ \ldots, \ E_{i-1}^i{}_{C_0+i-1}{}_{C_1}$ are all those events having only one of the first $i-1$ steps surpassing the threshold π and yet the corresponding coin flip represented by the Bernoulli variable not allowing the randomised rule to stop at the value;
- $E_{i-1}^i{}_{C_0+i-1}{}_{C_1+1}, \ \ldots, \ E_{i-1}^i{}_{C_0+i-1}{}_{C_1+i-1}{}_{C_2}$ are all those events having only two of the first $i-1$ steps surpassing the threshold π and yet the corresponding coin flips not allowing the randomised rule to stop at the values and so on.

This yields a total of $\sum_{j=0}^{i-1} {}^{i-1}C_j = (1+1)^{i-1}$ disjoint events. Denote $E_k^i := \overline{E}_k^i \cap \{X_i \geq \pi, \ B_i = 1\}$, where the overline denotes the complement. By the independence of X, B and H,

$$\mathbb{E}\left(Y_{\tau_{\pi,q}} \mathbb{1}_{E_\pi^h}\right) = \mathbb{E}\left(Y_{\tau_{\pi,q}} \mathbb{1}_{\{H=h\}} \sum_{i \in [h]} \mathbb{1}_{\{i=\inf\{j\in[h]:\, X_j \geq \pi,\, B_j=1\}\}}\right)$$

$$= \sum_{i\in[h]} \mathbb{E}\left(X_i B_i \mathbb{1}_{\{H=h\}} \mathbb{1}_{\{i=\inf\{j\in[h]:\, X_j \geq \pi\}\}}\right)$$

$$= \mathbb{P}(H=h) \sum_{i\in[h]} \mathbb{E}\left(X_i B_i \sum_{k=1}^{2^{i-1}} \mathbb{1}_{E_k^i}\right) = \mathbb{P}(H=h) q \mathbb{E}\left[X \mathbb{1}_{\{X \geq \pi\}} \sum_{i\in[h]} \sum_{k=1}^{2^{i-1}} \mathbb{P}(\overline{E}_k^i)\right].$$

Since $\mathbb{E}\left(X \mathbb{1}_{\{X \geq \pi\}}\right) = \mathbb{E}(X|X \geq \pi)\mathbb{P}(X \geq \pi)$ and

$$\sum_{k=1}^{2^{i-1}} \mathbb{P}(\overline{E}_k^i) = \sum_{l=0}^{i-1} {}^{i-1}C_l \mathbb{P}^{i-1-l}(X < \pi)[(1-q)\mathbb{P}(X \geq \pi)]^l = [1 - q\mathbb{P}(X \geq \pi)]^{i-1},$$

it follows that $\mathbb{E}\left(X_{\tau_\pi} \mathbb{1}_{E_\pi^h}\right) = \mathbb{E}(X|X \geq \pi)\{1 - [1 - q\mathbb{P}(X \geq \pi)]^h\}\mathbb{P}(H=h)$. Hence

$$\mathbb{E}X_{\tau_{\pi,q}} = \mathbb{E}Y_{\tau_{\pi,q}} = \mathbb{E}(X|X \geq \pi) \sum_{h\in\mathbb{N}} \{1 - [1 - q\mathbb{P}(X \geq \pi)]^h\}\mathbb{P}(H=h)$$

$$= \mathbb{E}(X|X \geq \pi)\mathbb{E}\{1 - [1 - q\mathbb{P}(X \geq \pi)]^H\},$$

which yields the result. □

B.4 Proof of Theorem 1.1

Proof of Theorem 1.1. For X continuous, by Lemmas 3.1 and 3.2,

$$\frac{\mathbb{E}X_{\tau_p}}{\mathbb{E}M_H} \geq c_p := 1 - \mathbb{E}\mathbb{P}^H(X < p).$$

For every $H \in \mathcal{G}$, it holds that $G \prec_{pgf} H$, with $G \sim \mathrm{Geom}(1/\mu)$, which denotes the geometric distribution with mean $\mu = \mathbb{E}H$. Thus, denoting $t := \mathbb{P}(X < p) = 1 - 1/\mu$,

$$\mathbb{E}\left(t^H\right) \leq \mathbb{E}\left(t^G\right) = \frac{1}{\mu}\sum_{h \in \mathbb{N}} t^h \left(1 - \frac{1}{\mu}\right)^{h-1} = \frac{t}{\mu}\frac{1}{1 - t\left(1 - \frac{1}{\mu}\right)} = \frac{1 - \frac{1}{\mu}}{2 - \frac{1}{\mu}}.$$

Thus the result follows since this implies that

$$c_p \geq 1 - \frac{1 - \frac{1}{\mu}}{2 - \frac{1}{\mu}} = \frac{1}{2 - \frac{1}{\mu}}.$$

For X discontinuous, let $V(x)$ be the cdf of X. The case that is most difficult to handle by adapting the previous part of the argument, which assumes continuity, is when for some p,

$$\mathbb{P}(X \geq p) > \frac{1}{\mu} > \lim_{\varepsilon \longrightarrow 0^+} \mathbb{P}(X \geq p + \varepsilon),$$

or equivalently,

$$\mathbb{P}(X < p) < 1 - \frac{1}{\mu} < \lim_{\varepsilon \longrightarrow 0^+} \mathbb{P}(X < p + \varepsilon).$$

Using the continuity from above of the cdf, this is equivalent to

$$\lim_{\varepsilon \longrightarrow 0^+} V(p - \varepsilon) < 1 - \frac{1}{\mu} < V(p),$$

that is to having $1 - 1/\mu$ in correspondence of a jump of the cdf at p. We will therefore explicitly treat this scenario.[19] By monotonicity, $V(x)$ has at most countably many jump discontinuities at values $\{j_k\}$, where k ranges either in $\mathbb{N}$ (infinitely many jumps) or in $[n]$ for some $n \in \mathbb{N}$ (finitely many jumps). In the following we adopt the terminology corresponding to the infinite case, replacing *sequence* with *list* yields the corresponding statements for the finite case; furthermore, $\{j_k\}$ is understood as an increasing enumeration of all the discontinuity points, and denoted as D. As customary in the literature concerning the analysis of discontinuous distributions, we will assume that we are not in the pathological case of D dense or having accumulation points. Finally, recall that under our assumptions, there is a value $\bar{k}$ such that $p = j_{\bar{k}}$. Let $\varepsilon(l)$ be a small enough positive monotonically vanishing sequence. Since all jumps are isolated, for every $l \in \mathbb{N}$ there exists an increasing subsequence of the jump discontinuities, denoted, with slight abuse of notation, $\{j_{k_\ell(l)}\} \subseteq D$, such that

$$D \subseteq \bigcup_\ell [j_{k_\ell(l)} - \varepsilon(l), j_{k_\ell(l)}] =: U_l,$$

where the union is disjoint. Note that, for some $l \in \mathbb{N}$, it could happen that $j_{k_1(l)} - \varepsilon(l) < 0$. We now linearly interpolate $V(x)$ on the intervals $[j_{k_\ell(l)} - \varepsilon(l), j_{k_\ell(l)}]$, and keep it unaltered otherwise, so as to construct an approximating sequence of cdf's $V_l(x)$. We define, for every fixed $l \in \mathbb{N}$,

$$V_l(x) := \begin{cases} V(j_{k_\ell(l)} - \varepsilon(l)) + (x - j_{k_\ell(l)} + \varepsilon(l))\frac{V(j_{k_\ell(l)}) - V(j_{k_\ell(l)} - \varepsilon(l))}{\varepsilon(l)} \\ \forall x \in [j_{k_\ell(l)} - \varepsilon(l), j_{k_\ell(l)}], \ \forall \ell : \ j_{k_\ell(l)} - \varepsilon(l) \geq 0, \\ x\frac{V(j_{k_1(l)})}{j_{k_1(l)}} \\ \forall x \in [0, j_{k_1(l)}], \text{ if } j_{k_1(l)} - \varepsilon(l) < 0, \\ V(x) \\ \text{otherwise.} \end{cases}$$

---

[19] If $1 - 1/\mu$ is not in correspondence of a jump discontinuity but the distribution is nonetheless discontinuous, similar ideas can be put in place, without the issue, which will lead to the necessity of randomisation, being there.

Denote $\mathcal{D}_l$ the distribution having cdf $V_l(x)$ and denote the corresponding continuous (since $V_l(x)$ is continuous and piecewise differentiable) random variables $X^{(l)} \sim \mathcal{D}_l$. The family $\{X^{(l)}\}$ is dominated by $\varepsilon(1) + X \in \mathcal{L}^1(\Omega)$ by construction, since

$$\mathbb{E}X^{(l)} \leq \int_0^\infty 1 - V(x)dx + \varepsilon(l)\sum_\ell V(j_{k_\ell(l)}) - V(j_{k_\ell(l)} - \varepsilon(l)) \leq \mathbb{E}X + \varepsilon(1),$$

having used

$$1 - V(j_{k_\ell(l)} - \varepsilon(l)) - [1 - V(j_{k_\ell(l)})] = V(j_{k_\ell(l)}) - V(j_{k_\ell(l)} - \varepsilon(l)) \leq \mathbb{P}(j_{k_\ell(l)} - \varepsilon(l) \leq X \leq j_{k_\ell(l)}),$$

$\mathbb{P}(X \in U_l) \leq 1$ and $\varepsilon(l) \leq \varepsilon(1)$. Therefore, the family $\{X^{(l)}\}$ is Uniformly Integrable (UI). Note also that by construction $X^{(l)} \xrightarrow{w} X$. Therefore, defining $M_H^{(l)} := \max\{X_1^{(l)}, \ldots, X_H^{(l)}\}$, where $X_i^{(l)} \sim \mathcal{D}_l$ are iid, we have that $\{M_H^{(l)}\}$ is UI and $M_H^{(l)} \xrightarrow{w} M_H$. Using [4, Corollary 5] it follows that $\mathbb{E}M_H^{(l)} \longrightarrow \mathbb{E}M_H$ as $l \longrightarrow \infty$. Furthermore, by Lemma 3.2, for every l there is p_l such that $\mathbb{P}(X^{(l)} \geq p_l) = 1/\mu$ and $\mathbb{E}M_H^{(l)} \leq \mathbb{E}(X^{(l)}|X^{(l)} \geq p_l)$. Since by construction $p_l \longrightarrow p$ as $l \longrightarrow \infty$, the fact that $\{X^{(l)}\}$ is UI and $X^{(l)} \xrightarrow{w} X$ implies also that $\mathbb{E}(X^{(l)}|X^{(l)} \geq p_l) \longrightarrow \mathbb{E}(X|X \geq p)$ as $l \longrightarrow \infty$. These facts altogether imply that also in this case $\mathbb{E}M_H \leq \mathbb{E}(X|X \geq p)$.

To conclude, by Lemma 3.3 the single-threshold randomised algorithm $\tau_{p,\bar{q}}$, where

$$\bar{q} := \frac{1}{\mu\mathbb{P}(X \geq p)},$$

is such that

$$\frac{\mathbb{E}X_{\tau_{p,\bar{q}}}}{\mathbb{E}M_H} \geq c_{p,\bar{q}} := 1 - \mathbb{E}\left\{[1 - \bar{q}\mathbb{P}(X \geq p)]^H\right\} = 1 - \mathbb{E}\left[\left(1 - \frac{1}{\mu}\right)^H\right] \geq \frac{1}{2 - \frac{1}{\mu}},$$

where the last inequality follows as in the continuous case, by exploiting that for any $H \in \mathcal{G}$, $G \prec_{pgf} H$. $\qquad\square$

B.5 Further details on the $\mathcal{G}$ class

For a better intuition of how much more general the somewhat abstract $\mathcal{G}$ class is, compared with the IHR class, we mention some of its most well-known subclasses, which earned considerable interest in applications. H always denotes a horizon.

Definition B.1. H *is* New Better than Used (NBU) *if for every $h, k \geq 2$, $S(h+k) \leq S(h)S(k)$.*

Definition B.2. *For every $h \in \mathbb{N}$ the* mean residual life *of H is defined as*

$$m(h) := \begin{cases} \mathbb{E}(H - h|H \geq h), & S(h) > 0 \\ 0, & S(h) = 0. \end{cases}$$

Note that $m(1) = \mu := \mathbb{E}H$.

Definition B.3. H *is* Harmonically New Better than Used in Expectation (HNBUE) *if for all $n \in \mathbb{N}$,*

$$\frac{n}{\sum_{h=1}^n \frac{1}{m(h)}} \leq \mu.$$

Furthermore, consider all the other classes in the tower of inclusions introduced after Remark 3.1 in Section 3. For simplicity we omit introducing further definitions, as the denominations are suggestive enough to convey the intuitive meaning, and can be easily found in the literature (see, for example, [6] for definitions and some of the inclusion properties we will recall). The HNBUE class, introduced in [40], is roughly speaking the largest known, which is a subclass of the $\mathcal{G}$ class [28], with intuitive meaning in terms of ageing properties.

B.6   Proof of Lemma 3.4

Proof of Lemma 3.4. Consider the RH for $X_i \sim X$ for all $i \in \mathbb{N}$, to which we add $X_0 = 0$ as per the framework of Section 2. By Theorem 2.1, and keeping the same notations and constructions, $\mathcal{V}(\mathbf{X}) = \mathcal{V}(\mathbf{Y}) = \mathcal{V}(\mathbf{Z}) = V_0$. By (2.1), $V_0 = \mathbb{E}V_1 = \mathbb{E}(Z_1 \vee \mathbb{E}_{\mathcal{B}_1} V_2)$. Note that $Z_i := q^{i-1} X_i$, thus $\mathbb{E}_{\mathcal{B}_1} V_2 = q\mathbb{E}V_1 = qV_0$. Indeed,

$$\mathbb{E}_{\mathcal{B}_1} V_2 = \operatorname*{ess\,sup}_{\zeta \in \mathcal{T}_2} \mathbb{E}_{\mathcal{B}_1} Z_\zeta = \operatorname*{ess\,sup}_{\zeta \in \mathcal{T}_2} \mathbb{E} Z_\zeta = q \sup_{\zeta \in \mathcal{T}_2} \mathbb{E}\left(q^{\zeta-2} X_\zeta\right) = q \operatorname*{ess\,sup}_{\zeta \in \mathcal{T}_1} \mathbb{E}\left(q^{\zeta-1} X_\zeta\right) = q\mathbb{E}V_1,$$

where we used:

- Lemma 2.1 in the first and last equality;
- that any $\zeta \in \mathcal{T}_2$ is independent of $\mathcal{B}_1$ and $\{X_i\}$ are *iid*, implying that X_ζ is also independent of $\mathcal{B}_1$, in the second equality;
- that $\{X_i\}$ are *iid* in the second to last equality, implying that any stopping rule that has not stopped by step 2 and undergoes discount factors shifted backwards by one, has the same expected future reward as a stopping rule that has not stopped by step 1 and undergoes unshifted discount factors. We call this property of geometric discounts *time-invariance*.

We have obtained the equation

$$V_0 = \mathbb{E}(X \vee qV_0). \tag{B.2}$$

Denoting the cdf of X as $F(x)$, it is straightforward to compute that

$$\mathbb{E}(X \vee qV_0) = qV_0 F(qV_0) + \int_{qV_0}^{b} x \, dF(x) = b - \int_{qV_0}^{b} F(x) dx,$$

where we used integration by parts in the second equality. By plugging this in (B.2) we obtain the equation

$$f(V_0) := V_0 + \int_{qV_0}^{b} F(x) dx = b. \tag{B.3}$$

Note that, since V_0 is the expected return, $0 \le V_0 \le b$, and thus $f(V)$ is a strictly increasing differentiable function on $[0, b]$, since by the *Fundamental Theorem of Calculus* $f'(V) = 1 - F(qV) > 0$ as $qV < b$. Observe that, by using $F(x) \le 1$, we have $f(0) \le b$, and by using the increasing monotonicity of $F(x)$, we have that $f(b) \ge b + F(qb)b(1-q) \ge b$. In conclusion by the *Intermediate Value Theorem* a value V_0 satisfying (B.3) exists. By Remark 2.2 and Definition 2.2 and the time-invariance aforementioned, it is optimal to stop at the first value greater or equal than V_0. $\qquad\square$

B.7   Tightness of 2-approximations

The following argument appears in [2, Theorem 3.5]. We include it for ease of reference, and because of a slight difference in the more rigorous setup we established with Lemma 3.4.

Theorem B.1 ([2][Theorem 3.5]). *RH is $2 - 1/\mu$-hard.*

Proof. Fix a geometric horizon having failure probability $0 < q < 1$, denoted as $H \sim \text{Geom}(1 - q)$, and an instance for the values $X \sim \mathcal{D}_p$, where $\mathcal{D}_p$ denotes a two-point distribution defined as $p := \mathbb{P}(X = x_2) = 1 - \mathbb{P}(X = x_1)$, with $x_2 > x_1 > 0$ and, as we vary $p \in (0, 1)$,

$$x_1 \le \frac{qpx_2}{1 - q(1-p)}. \tag{B.4}$$

Then by Lemma 3.4 it is optimal to adopt a single-threshold

$$V_0 \leq \frac{px_2}{1 - q(1 - p)}, \tag{B.5}$$

coinciding with the expected return. The upper bound follows from the fact that since $x_1 < V_0 < x_2$, we need to explore only two cases:

- If $x_1 < qV_0$, $\mathbb{E}(X \vee qV_0) = px_2 + (1 - p)qV_0$. Plugging this into (B.2) yields $V_0 = px_2 + (1 - p)qV_0$. Solving for V_0, we obtain

$$V_0 = \frac{px_2}{1 - q(1 - p)}.$$

- If $qV_0 \leq x_1$, $\mathbb{E}(X \vee qV_0) = \mathbb{E}X = (1 - p)x_1 + px_2$. Plugging this into (B.2) yields $V_0 = (1 - p)x_1 + px_2$.

Note that by (B.4)

$$px_2 + (1 - p)x_1 \leq \left[p + \frac{(1 - p)pq}{1 - q(1 - p)} \right] x_2 = \frac{px_2}{1 - q(1 - p)},$$

which implies (B.5).

As we vary $p \in (0, 1)$, for simplicity, require (B.4) with equality, that is

$$x_1 = \frac{qpx_2}{1 - q(1 - p)}. \tag{B.6}$$

By the law of total expectation we can compute straightforwardly that

$$\mathbb{E}M_H = (1 - q) \sum_{i \in \mathbb{N}} \mathbb{E}M_i q^{i-1} = (1 - q) \left[x_1 \frac{1 - p}{1 - (1 - p)q} + x_2 \left(\frac{1}{1 - q} - \frac{1 - p}{1 - (1 - p)q} \right) \right]$$

$$= x_1 \frac{(1 - p)(1 - q)}{1 - (1 - p)q} + x_2 \frac{p}{1 - (1 - p)q} = x_1 \left(\frac{(1 - p)(1 - q)}{1 - (1 - p)q} + \frac{1}{q} \right),$$

since $\mathbb{E}M_i = x_1(1 - p)^i + x_2[1 - (1 - p)^i]$ and having used (B.6) in the last equality.

Finally, we consider the gambler-to-prophet ratio. For any $\tau \in \mathcal{T}^*$, by (B.6) again, we obtain

$$\frac{\mathbb{E}X_\tau}{\mathbb{E}M_H} \leq \frac{x_2}{x_1} \frac{\frac{p}{1 - q(1 - p)}}{\frac{(1 - p)(1 - q)}{1 - (1 - p)q} + \frac{1}{q}} = \frac{1 - q(1 - p)}{qp} \frac{\frac{p}{1 - q(1 - p)}}{\frac{(1 - p)(1 - q)}{1 - (1 - p)q} + \frac{1}{q}} = \frac{1}{\frac{(1 - p)(1 - q)q}{1 - (1 - p)q} + 1} \longrightarrow \frac{1}{q + 1} = \frac{1}{2 - \frac{1}{\mu}}$$

as $p \longrightarrow 0$, since $1 - q = 1/\mu$. Thus no online algorithm can be more than $2 - 1/\mu$-competitive. $\quad\square$

C Proofs omitted from Section 4

C.1 Proof of Lemmas 4.1 and 4.2

Proofs of Lemma 4.1. Consider

$$\mathbb{E}M_n = \int_{\mathbb{R}} x f_n(x) dx = \int_1^\infty x n V^{n-1}(x) V'(x) dx. \tag{C.1}$$

Performing the substitution $u = V(x)$ into (C.1) and plugging in the expression $V^{-1}(u) := (1 - u)^{-\frac{1}{1+\varepsilon}}$ yields

$$\mathbb{E}M_n = \int_0^1 V^{-1}(u) n u^{n-1} du = n \, \mathrm{B} \left(n, 1 - \frac{1}{1 + \varepsilon} \right),$$

where
$$B(\alpha, \beta) := \int_0^1 u^{\alpha-1}(1-u)^{\beta-1}du$$

is the Beta function. The claim follows from the fact that

$$B\left(n, 1 - \frac{1}{1+\varepsilon}\right) = \frac{\Gamma(n)\Gamma\left(1 - \frac{1}{1+\varepsilon}\right)}{\Gamma\left(n+1-\frac{1}{1+\varepsilon}\right)} \sim \Gamma\left(1 - \frac{1}{1+\varepsilon}\right)n^{-\frac{\varepsilon}{1+\varepsilon}}. \tag{C.2}$$

This is shown by using the fact that the Gamma function

$$\Gamma(x) := \int_0^\infty t^{x-1}e^{-t}dt \sim (x-1)^{x-1}e^{-(x-1)}\sqrt{2\pi(x-1)}$$

as $x \longrightarrow \infty$, by a well-known asymptotics, derived through the Laplace method. In particular, one applies the result for $x = n$ and $x = n + 1 - \frac{1}{1+\varepsilon}$, and then simplifies the ratio of Gamma functions' asymptotics as $n \longrightarrow \infty$:

$$\frac{\Gamma(n)}{\Gamma\left(n+1-\frac{1}{1+\varepsilon}\right)} \sim \frac{(n-1)^{n-1}\sqrt{n-1}e^{n-\frac{1}{1+\varepsilon}}}{\left(n-\frac{1}{1+\varepsilon}\right)^{n-\frac{1}{1+\varepsilon}}\sqrt{n-\frac{1}{1+\varepsilon}}e^{n-1}} \sim e^{\frac{\varepsilon}{1+\varepsilon}}\frac{(n-1)^{n-1}}{\left(n-\frac{1}{1+\varepsilon}\right)^{n-\frac{1}{1+\varepsilon}}}$$

$$= \left(1 - \frac{1-\frac{1}{1+\varepsilon}}{n-\frac{1}{1+\varepsilon}}\right)^{n-\frac{1}{1+\varepsilon}}\frac{e^{\frac{\varepsilon}{1+\varepsilon}}}{(n-1)^{1-\frac{1}{1+\varepsilon}}} \sim \left(1 - \frac{\frac{\varepsilon}{1+\varepsilon}}{n-\frac{1}{1+\varepsilon}}\right)^{n-\frac{1}{1+\varepsilon}}\frac{e^{\frac{\varepsilon}{1+\varepsilon}}}{n^{\frac{\varepsilon}{1+\varepsilon}}} \sim n^{-\frac{\varepsilon}{1+\varepsilon}}.$$

$\square$

Proof of Lemma 4.2. Since the sequence $\left\{h^{-\frac{1}{1+2\varepsilon}}\right\}_h$ is monotone decreasing, the summation defining Z_m is asymptotically equivalent to the corresponding integral through the standard estimate

$$\int_\ell^m x^{-\frac{1}{1+2\varepsilon}}dx \leq \sum_{h=\ell}^m h^{-\frac{1}{1+2\varepsilon}} \leq (\ell-1)^{-\frac{1}{1+2\varepsilon}} + \int_\ell^m x^{-\frac{1}{1+2\varepsilon}}dx.$$

This yields that as $m \longrightarrow \infty$, (4.3) holds the extra term outside the integral is constant, while the following asymptotics of the integral holds:

$$\int_\ell^m x^{-\frac{1}{1+2\varepsilon}}dx = \frac{x^{1-\frac{1}{1+2\varepsilon}}}{1-\frac{1}{1+2\varepsilon}}\bigg|_\ell^m \sim \frac{1+2\varepsilon}{2\varepsilon}m^{\frac{2\varepsilon}{1+2\varepsilon}}.$$

Consider that by the law of total expectation, Lemma 4.1 and (4.3) and the increasing monotonicity of $\mathbb{E}M_n$, as $m \longrightarrow \infty$,

$$\mathbb{E}M_{H_m} = \sum_{n=\ell}^m \mathbb{E}M_n p_n^{(m)} \sim \sum_{h=\ell}^{\lfloor\sqrt{m}\rfloor-1} \mathbb{E}M_n p_n^{(m)} + \frac{\Gamma\left(1-\frac{1}{1+\varepsilon}\right)}{Z_m}\sum_{n=\lfloor\sqrt{m}\rfloor}^m n^{\frac{1}{1+\varepsilon}-\frac{1}{1+2\varepsilon}}$$

$$\sim \frac{\Gamma\left(1-\frac{1}{1+\varepsilon}\right)}{\frac{1+2\varepsilon}{2\varepsilon}}\frac{\sum_{n=\lfloor\sqrt{m}\rfloor}^m n^{\frac{\varepsilon}{(1+\varepsilon)(1+2\varepsilon)}}}{m^{\frac{2\varepsilon}{1+2\varepsilon}}}. \tag{C.3}$$

where the last asymptotic relation holds since we will now show that as $m \longrightarrow \infty$ the last series diverges with a higher order than

$$\sum_{h=\ell}^{\lfloor\sqrt{m}\rfloor-1} \mathbb{E}M_n p_n^{(m)} \leq \mathbb{E}M_{\lfloor\sqrt{m}\rfloor} \sim \sqrt{m}^{\frac{1}{1+\varepsilon}}.$$

36

Indeed, since the sequence $\left\{ n^{\frac{\varepsilon}{(1+\varepsilon)(1+2\varepsilon)}} \right\}_n$ is monotone increasing, the last series of (C.3) is asymptotically equivalent to the corresponding integral through the standard estimate

$$(\lfloor\sqrt{m}\rfloor - 1)^{\frac{\varepsilon}{(1+\varepsilon)(1+2\varepsilon)}} + \int_{\lfloor\sqrt{m}\rfloor}^m x^{\frac{\varepsilon}{(1+\varepsilon)(1+2\varepsilon)}} dx \le \sum_{n=\lfloor\sqrt{m}\rfloor}^m n^{\frac{\varepsilon}{(1+\varepsilon)(1+2\varepsilon)}}$$

$$\le (m+1)^{\frac{\varepsilon}{(1+\varepsilon)(1+2\varepsilon)}} + \int_{\lfloor\sqrt{m}\rfloor}^m x^{\frac{\varepsilon}{(1+\varepsilon)(1+2\varepsilon)}} dx.$$

As $m \longrightarrow \infty$,

$$\int_{\lfloor\sqrt{m}\rfloor}^m x^{\frac{\varepsilon}{(1+\varepsilon)(1+2\varepsilon)}} dx = \frac{x^{1+\frac{\varepsilon}{(1+\varepsilon)(1+2\varepsilon)}}}{1+\frac{\varepsilon}{(1+\varepsilon)(1+2\varepsilon)}}\Bigg|_{\lfloor\sqrt{m}\rfloor}^m \sim \frac{m^{1+\frac{\varepsilon}{(1+\varepsilon)(1+2\varepsilon)}}}{1+\frac{\varepsilon}{(1+\varepsilon)(1+2\varepsilon)}}, \tag{C.4}$$

and the other terms in $\lfloor\sqrt{m}\rfloor - 1$, $\lfloor\sqrt{m}\rfloor^{\frac{1}{1+\varepsilon}}$ and $m+1$ are negligible, being of lower order. Plugging this asymptotics for the summation in the last term of (C.3) yields the result sought, since

$$1 + \frac{\varepsilon}{(1+\varepsilon)(1+2\varepsilon)} - \frac{2\varepsilon}{1+2\varepsilon} = \frac{1}{1+\varepsilon}.$$

$\square$

C.2   Proof of Proposition 4.1

Proof of Proposition 4.1. Our strategy will be to show the boundedness of the sequence of optimal single thresholds $\{\bar{\pi}_m\}$ (corresponding to horizons $\{H_m\}$ for m large enough and ε small enough), which maximize the sequence $\mathbb{E}X_{\tau_{\pi_m}} := \mathbb{E}Y_{\tau_{\pi_m}}^{(H_m)}$, where $\{\pi_m\}$ is an arbitrary sequence of single-thresholds and $Y_i^{(H_m)} := X_i \mathbb{1}_{\{H_m \ge i\}}$. Then, we show that there is a subsequence $\{m_j\}$, such that as $j \longrightarrow \infty$,

$$r_{m_j} := \frac{\mathbb{E}Y_{\tau_{\bar{\pi}_{m_j}}}^{(H_{m_j})}}{\mathbb{E}M_{H_{m_j}}} \longrightarrow 0, \tag{C.5}$$

thus yielding the claim, since this ratio is an upper bound on any ratio defined for arbitrary single-thresholds.

Since X is a continuous random variable, we can apply Lemma 3.1 to compute $\mathbb{E}Y_{\tau_\pi}^{(H_m)}$, which is nontrivial only if $\pi > 1$. Consider first that by direct computation,

$$\mathbb{E}(X|X \ge \pi) = \pi + \mathbb{E}(X - \pi | X \ge \pi) = \pi + \frac{1}{\overline{V}(\pi)} \int_\pi^\infty \overline{V}(x)dx = \pi + \pi^{1+\varepsilon} \int_\pi^\infty \frac{1}{x^{1+\varepsilon}}dx = \left(1 + \frac{1}{\varepsilon}\right)\pi \tag{C.6}$$

and

$$c_\pi = 1 - \mathbb{E}V^{H_m}(\pi) = 1 - \mathbb{E}\left(1 - \frac{1}{\pi^{1+\varepsilon}}\right)^{H_m}. \tag{C.7}$$

For every fixed m large enough, consider that by (C.6) and (C.7) the function

$$f_m(\pi) := \mathbb{E}Y_{\tau_\pi}^{(H_m)} = c_\pi \mathbb{E}(X|X \ge \pi) = \left(1 + \frac{1}{\varepsilon}\right)\pi\left[1 - \mathbb{E}\left(1 - \frac{1}{\pi^{1+\varepsilon}}\right)^{H_m}\right],$$

and compute

$$f'_m(\pi) = \left(1 + \frac{1}{\varepsilon}\right)\left[1 - \mathbb{E}\left(1 - \frac{1}{\pi^{1+\varepsilon}}\right)^{H_m} - \frac{1+\varepsilon}{\pi^{1+\varepsilon}}\mathbb{E}H_m\left(1 - \frac{1}{\pi^{1+\varepsilon}}\right)^{H_m-1}\right].$$

Note that for any fixed m and $\varepsilon > 0$, assuming that $\ell > 1$, we have the following:

- $f_m(\pi)$ increases when $\pi \approx 1$, since $f'_m(1) = 1 + 1/\varepsilon > 0$.
- $f_m(\pi)$ decreases when π is large enough, since, by Bernoulli's inequality

$$\left(1 - \frac{1}{\pi^{1+\varepsilon}}\right)^h \geq 1 - \frac{h}{\pi^{1+\varepsilon}},$$

we have that, as $\pi \longrightarrow \infty$,

$$f'_m(\pi) \leq \left(1 + \frac{1}{\varepsilon}\right)\left\{\frac{\mathbb{E}H_m}{\pi^{1+\varepsilon}} - \frac{1+\varepsilon}{\pi^{1+\varepsilon}}\left(\mathbb{E}H_m - \frac{\mathbb{E}H_m(H_m - 1)}{\pi^{1+\varepsilon}}\right)\right\}$$

$$= \frac{1}{\pi^{1+\varepsilon}}\left(1 + \frac{1}{\varepsilon}\right)\left(-\varepsilon\mathbb{E}H_m + \frac{1+\varepsilon}{\pi^{1+\varepsilon}}\mathbb{E}H_m(H_m - 1)\right),$$

with the last term being eventually negative as π grows large. Thus, there exists $\pi^* = \pi^*_m(\varepsilon)$ such that $f'_m(\pi) < 0$ for every $\pi > \pi^*$.

The fact that the function $f_m(\pi)$ is strictly increasing at the left end of the interval $[1, \infty)$ and strictly decreasing on the right end, that is on (π^*, ∞), means that, by continuity, that is by the *Intermediate Value Theorem* applied to the compact interval $[1, \pi^*]$, a global maximum exists and is attained at $\bar{\pi}_m = \bar{\pi}_m(\varepsilon) \in (1, \pi^*]$. Furthermore, it is characterised by $f'_m(\bar{\pi}_m) = 0$, which we rewrite as

$$g_m(\bar{\pi}_m) := \mathbb{E}\left(1 - \frac{1}{\bar{\pi}_m^{1+\varepsilon}}\right)^{H_m} + \frac{1+\varepsilon}{\bar{\pi}_m^{1+\varepsilon}}\mathbb{E}H_m\left(1 - \frac{1}{\bar{\pi}_m^{1+\varepsilon}}\right)^{H_m - 1} = 1. \tag{C.8}$$

To conclude, consider that by (C.5) to (C.7), Lemma 4.2, and Remark 4.1 we have that, as $m \longrightarrow \infty$, uniformly in $\varepsilon > 0$ small enough,

$$r_m = c_{\bar{\pi}_m}\frac{\mathbb{E}(X | X \geq \bar{\pi}_m)}{\mathbb{E}M_{H_m}} \sim \left(1 + \frac{1}{\varepsilon}\right)\frac{c_{\bar{\pi}_m}\bar{\pi}_m}{Nm^{\frac{1}{1+\varepsilon}}} \asymp \frac{c_{\bar{\pi}_m}\bar{\pi}_m}{\varepsilon m^{\frac{1}{1+\varepsilon}}}. \tag{C.9}$$

Consider that $\{\bar{\pi}_m\}$ is either bounded or unbounded. Then by considering the following cases, we can always conclude that the optimal sequence of the gambler-to-prophet ratios, for fixed threshold strategies, admits a vanishing $\{r_{m_j}\}$, proving (C.5).

Bounded and slowly increasing. Assume that $\bar{\pi}_m = \mathcal{O}\left(m^{\frac{1}{1+\varepsilon}}\right)$. Since $c_{\bar{\pi}_m} < 1$, by (C.9), we have that r_m trivially vanishes as $m \longrightarrow \infty$, for any fixed ε small enough. Note that this case covers also the trivial case for which there exists a subsequence $\{\bar{\pi}_{m_j}\}$ such that $\bar{\pi}_{m_j} \leq 1$, for which as $j \longrightarrow \infty$,

$$r_{m_j} = \frac{\mathbb{E}X}{\mathbb{E}M_{H_{m_j}}} \asymp \frac{1}{\varepsilon m_j^{\frac{1}{1+\varepsilon}}} \longrightarrow 0.$$

Unbounded and fast increasing. Assume that $m^{\frac{1}{1+\varepsilon}} = \mathcal{O}(\bar{\pi}_m)$. Note that, by Bernoulli's inequality,

$$\mathbb{E}\left(1 - \frac{1}{\bar{\pi}_m^{1+\varepsilon}}\right)^{H_m} \geq 1 - \frac{\mathbb{E}H_m}{\bar{\pi}_m^{1+\varepsilon}},$$

yielding that

$$c_{\bar{\pi}_m} := 1 - \mathbb{E}\left(1 - \frac{1}{\bar{\pi}_m^{1+\varepsilon}}\right)^{H_m} \leq \frac{\mathbb{E}H_m}{\bar{\pi}_m^{1+\varepsilon}}.$$

In conclusion, by (4.4) with $n = 1$ it holds that as $m \longrightarrow \infty$, uniformly in ε small enough, as $m \longrightarrow \infty$,

$$\frac{c_{\bar{\pi}_m}\bar{\pi}_m}{m^{\frac{1}{1+\varepsilon}}} \leq \frac{\mathbb{E}H_m}{m^{\frac{1}{1+\varepsilon}}\bar{\pi}_m^\varepsilon} \asymp \varepsilon\left(\frac{m^{\frac{1}{1+\varepsilon}}}{\bar{\pi}_m}\right)^\varepsilon \longrightarrow 0.$$

Therefore, by (C.9), we have that r_m vanishes.

Unbounded (and moderately increasing). By relying on the result above for unbounded and fast increasing behaviour, we show that if $\{\bar{\pi}_m\}$ is assumed unbounded, with no additional properties, r_m vanishes. This is an argument by contradiction and will essentially boil down to analysing, hypothetically, the moderately increasing regime, that is the one in which $\bar{\pi}_m \asymp m^{\frac{1}{1+\varepsilon}}$. We will show that it is not possible, and draw our conclusion from there.

Recall that $\{r_m\} \subseteq [0, 1]$ is a bounded sequence. Therefore, to show that $\{r_m\}$ vanishes as long as we choose ε small enough, it is sufficient to show that, under the same assumptions, all of its convergent subsequences $\{r_{m_j}\}$ vanish. By contradiction, assume that there is a convergent subsequence $\{r_{m_j}\}$ that does not vanish, no matter how small $\varepsilon > 0$ is taken, that is, as $j \longrightarrow \infty$, we have that $r_{m_j} \longrightarrow \rho = \rho(\varepsilon) \in (0, 1]$, with $\rho(\varepsilon)$ being bounded away from 0 no matter how small $\varepsilon > 0$ is taken. Then, by (C.9), we have that as $j \longrightarrow \infty$, for $\varepsilon > 0$ fixed small enough,

$$\mu_j := \frac{m_j}{\bar{\pi}_{m_j}^{1+\varepsilon}} \sim \left[\left(1 + \frac{1}{\varepsilon}\right)\frac{c_{\bar{\pi}_{m_j}}}{\rho N}\right]^{1+\varepsilon} \leq \left(\frac{1+\varepsilon}{\varepsilon \rho N}\right)^{1+\varepsilon}. \tag{C.10}$$

Since $\{\mu_j\}$ is asymptotically equivalent to a bounded sequence, it must be bounded too. By boundedness, we can consider a convergent subsequence such that for some fixed $\alpha = \alpha(\varepsilon) \geq 0$ satisfying the bound in (C.10), as $k \longrightarrow \infty$, $\mu_{j_k} \longrightarrow \alpha$.

Consider the function $\varepsilon \mapsto \alpha(\varepsilon)$. We show that, unless α is eventually (that is, for all ε small enough) identically 0, $\liminf_{\varepsilon \to 0} \alpha(\varepsilon) > 0$, that is, it is bounded away from zero uniformly for all ε small enough. Indeed, by contradiction, assume that α is not identically 0 for all ε small enough, and that $\liminf_{\varepsilon \to 0} \alpha(\varepsilon) = 0$. Recall that, as $k \longrightarrow \infty$,

$$\frac{1}{\bar{\pi}_{m_{j_k}}^{1+\varepsilon}} \sim \frac{\alpha}{m_{j_k}} \longrightarrow 0. \tag{C.11}$$

By (C.8) and (C.11) and (4.4) with $n = 1$

$$c_{\bar{\pi}_{m_{j_k}}} = \frac{1+\varepsilon}{\bar{\pi}_{m_{j_k}}} \mathbb{E}H_{m_{j_k}} \left(1 - \frac{1}{\bar{\pi}_{m_{j_k}}^{1+\varepsilon}}\right)^{H_{m_{j_k}}-1} \geq \frac{1+\varepsilon}{\bar{\pi}_{m_{j_k}}} \mathbb{E}H_{m_{j_k}} \left(1 - \frac{1}{\bar{\pi}_{m_{j_k}}^{1+\varepsilon}}\right)^{m_{j_k}}$$

$$\sim (1+\varepsilon)\alpha \frac{\mathbb{E}H_{m_{j_k}}}{m_{j_k}} \left(1 - \frac{\alpha}{m_{j_k}}\right)^{m_{j_k}} \sim \frac{1+\varepsilon}{1+4\varepsilon} 2\varepsilon\alpha e^{-\alpha}$$

On the other hand, by definition, Jensen's inequality (C.11) and (4.4) with $n = 1$

$$c_{\bar{\pi}_{m_{j_k}}} := 1 - \mathbb{E}\left(1 - \frac{1}{\bar{\pi}_{m_{j_k}}^{1+\varepsilon}}\right)^{H_{m_{j_k}}} \leq 1 - \left(1 - \frac{1}{\bar{\pi}_{m_{j_k}}^{1+\varepsilon}}\right)^{\mathbb{E}H_{m_{j_k}}}$$

$$\sim 1 - \left(1 - \frac{\alpha}{m_{j_k}}\right)^{\frac{2\varepsilon m_{j_k}}{1+4\varepsilon}} \sim 1 - e^{-\frac{2\alpha\varepsilon}{1+4\varepsilon}}.$$

Then, combining both estimates, we have that, taking the limit as $k \longrightarrow \infty$,

$$1 - e^{-\frac{2\alpha\varepsilon}{1+4\varepsilon}} \geq \frac{1+\varepsilon}{1+4\varepsilon} 2\varepsilon\alpha e^{-\alpha}. \tag{C.12}$$

Since $\liminf_{\varepsilon \to 0} \alpha(\varepsilon) = 0$, we have that, for infinitely many ε arbitrarily small,[20]

$$1 - e^{-\frac{2\alpha\varepsilon}{1+4\varepsilon}} = \frac{2\alpha\varepsilon}{1+4\varepsilon} + \mathcal{O}(\alpha^2\varepsilon^2),$$

---

[20] In what follows all asymptotic notation is understood as ε vanishes along a subsequence. We refrain from using a sequence ε_l for mere simplicity of notation, to avoid working with multiple indices (k and l), which might result confusing.
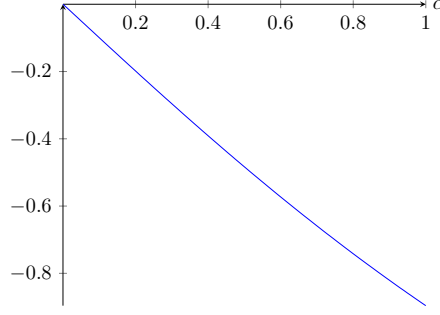
Figure 1: The function $h(\alpha)$.

and

$$\frac{1+\varepsilon}{1+4\varepsilon}2\varepsilon\alpha e^{-\alpha} = \frac{1+\varepsilon}{1+4\varepsilon}2\varepsilon\alpha + \mathcal{O}(\varepsilon\alpha^2).$$

Replacing back in (C.12), we obtain

$$\frac{2\alpha\varepsilon}{1+4\varepsilon} + \mathcal{O}(\alpha^2\varepsilon^2) \geq \frac{1+\varepsilon}{1+4\varepsilon}2\varepsilon\alpha + \mathcal{O}(\varepsilon\alpha^2),$$

which is equivalent to

$$\frac{1}{1+4\varepsilon} + \mathcal{O}(\alpha\varepsilon) \geq \frac{1+\varepsilon}{1+4\varepsilon} + \mathcal{O}(\alpha),$$

which is trivially not possible as ε vanishes, regardless of the rate at which α vanishes. We conclude that $\liminf_{\varepsilon\to 0}\alpha(\varepsilon) > 0$.

Recall also that the optimality condition (C.8) must be satisfied, which we recast as

$$g_{m_{j_k}}(\bar{\pi}_{m_{j_k}}) := \mathbb{E}\left[\phi(\bar{\pi}_{m_{j_k}}, H_{m_{j_k}})\right] = 1, \tag{C.13}$$

where

$$\phi(\pi, H) := \left(1 - \frac{1}{\pi^{1+\varepsilon}}\right)^{H-1}\left[1 - \frac{1}{\pi^{1+\varepsilon}} + (1+\varepsilon)\frac{H}{\pi^{1+\varepsilon}}\right].$$

To reach a contradiction, we derive the following asymptotic upper bound, holding uniformly in $\alpha > 0$ and as $k \longrightarrow \infty$,

$$g_{m_{j_k}}(\bar{\pi}_{m_{j_k}}) \leq 1 + \varepsilon\left[(2+\alpha)e^{-\alpha} - 2\right] + \mathcal{O}(\varepsilon^2) + \mathcal{O}\left(\frac{1}{m_{j_k}}\right). \tag{C.14}$$

We first clinch the argument and then show how to derive this asymptotic upper bound.

- Assume that it is the case that $\liminf_{\varepsilon\to 0}\alpha(\varepsilon) > 0$. Then for all $\varepsilon > 0$ small enough, (C.14) yields that $g_{m_{j_k}}(\bar{\pi}_{m_{j_k}}) < 1$, in contradiction with (C.13). Indeed, under the assumption, the strictly decreasing function $h(\alpha) := (2+\alpha)e^{-\alpha} - 2$ is negative and bounded away from zero for all $\varepsilon > 0$ small enough, since in this range of values of ε, $h(\alpha)$ vanishes only as α vanishes (see Figure 1).
- As a result of the contradiction reached, α must be identically 0 for all $\varepsilon > 0$ small enough, meaning that for all such ε, $\mu_{j_k} \longrightarrow 0$ as $k \longrightarrow \infty$. By boundedness, this ensures that $\mu_j \longrightarrow 0$ as $j \longrightarrow \infty$. That is, we are, subsequentially, in the fast increasing scenario of the previous point, which shows (upon replacing m with m_j)that r_{m_j} vanishes. This is in contradiction with the initial assumption, that $r_{m_j} \longrightarrow \rho > 0$, and therefore we have shown that $\rho = 0$.

40

Proof of (C.14). We are only left with proving the bound (C.14). To show (C.14), we compute the series expansion of $\phi(\bar{\pi}_{m_{j_k}}, H_{m_{j_k}})$ with respect to the first argument. By (C.11) and the fact that $H_{m_{j_k}} \leq m_{j_k}$, we have that, as $k \longrightarrow \infty$,

$$
\left(1 - \frac{1}{\bar{\pi}_{m_{j_k}}^{1+\varepsilon}}\right)^{H_{m_{j_k}}-1} = \exp\left(\left(H_{m_{j_k}} - 1\right)\log\left(1 - \frac{1}{\bar{\pi}_{m_{j_k}}^{1+\varepsilon}}\right)\right) \sim \exp\left(-\frac{H_{m_{j_k}}}{\bar{\pi}_{m_{j_k}}^{1+\varepsilon}}\left(1 + \mathcal{O}\left(\frac{1}{\bar{\pi}_{m_{j_k}}^{1+\varepsilon}}\right)\right)\right)
$$

$$
\sim \exp\left(-\alpha\frac{H_{m_{j_k}}}{m_{j_k}}\left(1 + \mathcal{O}\left(\frac{1}{m_{j_k}}\right)\right)\right)
$$

$$
= \sum_{n\geq 0}\frac{(-1)^n}{n!}\alpha^n\left(\frac{H_{m_{j_k}}}{m_{j_k}}\right)^n\left(1 + \mathcal{O}\left(\frac{1}{m_{j_k}}\right)\right)^n.
$$

Thus

$$
\phi(\bar{\pi}_{m_{j_k}}, H_{m_{j_k}}) \sim \sum_{n\geq 0}\frac{(-1)^n}{n!}\alpha^n\left(\frac{H_{m_{j_k}}}{m_{j_k}}\right)^n\left(1 + n\mathcal{O}\left(\frac{1}{m_{j_k}}\right)\right)\left[1 - \frac{\alpha}{m_{j_k}} + (1+\varepsilon)\alpha\frac{H_{m_{j_k}}}{m_{j_k}}\right]
$$

$$
= \sum_{n\geq 0}\frac{(-1)^n}{n!}\alpha^n\left(\frac{H_{m_{j_k}}}{m_{j_k}}\right)^n\left[1 + (1+\varepsilon)\alpha\frac{H_{m_{j_k}}}{m_{j_k}}\right] + \mathcal{O}\left(\frac{1}{m_{j_k}}\right)
$$

since both series

$$
\sum_{n\geq 0}\frac{(-1)^n}{n!}\alpha^n\left(\frac{H_{m_{j_k}}}{m_{j_k}}\right)^n, \quad \sum_{n\geq 0}\frac{(-1)^n}{n!}n\alpha^n\left(\frac{H_{m_{j_k}}}{m_{j_k}}\right)^n
$$

converge uniformly on the probability space. Due to the uniformity in the probability space holding for all asymptotic estimates so far performed, we can finally take expectation and obtain, by *Fubini Theorem* and (4.4), that

$$
g_{m_{j_k}}(\bar{\pi}_{m_{j_k}}) \sim \sum_{n\geq 0}\frac{(-1)^n}{n!}\alpha^n\frac{\mathbb{E}H_{m_{j_k}}^n}{m_{j_k}^n} + (1+\varepsilon)\alpha\sum_{n\geq 0}\frac{(-1)^n}{n!}\alpha^n\frac{\mathbb{E}H_{m_{j_k}}^{n+1}}{m_{j_k}^{n+1}} + \mathcal{O}\left(\frac{1}{m_{j_k}}\right)
$$

$$
\sim \sum_{n\geq 0}\frac{(-1)^n}{n!}\alpha^n\frac{2\varepsilon}{n+2(n+1)\varepsilon} + (1+\varepsilon)\alpha\sum_{n\geq 0}\frac{(-1)^n}{n!}\alpha^n\frac{2\varepsilon}{n+1+2(n+2)\varepsilon} + \mathcal{O}\left(\frac{1}{m_{j_k}}\right)
$$

$$
< 1 + 2\varepsilon\sum_{n\geq 1}\frac{(-1)^n}{n!}\alpha^n + \varepsilon(1+\varepsilon)\alpha\sum_{n\geq 0}\frac{(-1)^n}{n!}\alpha^n + \mathcal{O}\left(\frac{1}{m_{j_k}}\right)
$$

$$
= 1 + \varepsilon\left[2(e^{-\alpha} - 1) + (1+\varepsilon)\alpha e^{-\alpha}\right] + \mathcal{O}\left(\frac{1}{m_{j_k}}\right).
$$

Recall that globally on the nonnegative reals, $\alpha e^{-\alpha} \leq 1/e$, and therefore as $\varepsilon \longrightarrow 0^+$,

$$
\varepsilon\left[2(e^{-\alpha} - 1) + (1+\varepsilon)\alpha e^{-\alpha}\right] \leq \varepsilon\left[(2+\alpha)e^{-\alpha} - 2\right] + \mathcal{O}(\varepsilon^2)
$$

which implies (C.14). $\qquad\square$

C.3 Heuristics for the horizons in the hard instance of Proposition 4.1

We now comment heuristically regarding why single-threshold approximations should be expected to fail for this instance, perhaps even in higher capacity than what proven in Proposition 4.1 (which restricts the instance to the value distribution X, in order to show hardness). Note that by exploiting (4.4) with $n = 1, 2$ we have that as $m \longrightarrow \infty$,

$$
\frac{\operatorname{Var} H_m}{\mathbb{E}^2 H_m} = \frac{\mathbb{E}H_m^2}{\mathbb{E}^2 H_m} - 1 \longrightarrow \frac{(1+4\varepsilon)^2}{4\varepsilon(1+3\varepsilon)} - 1.
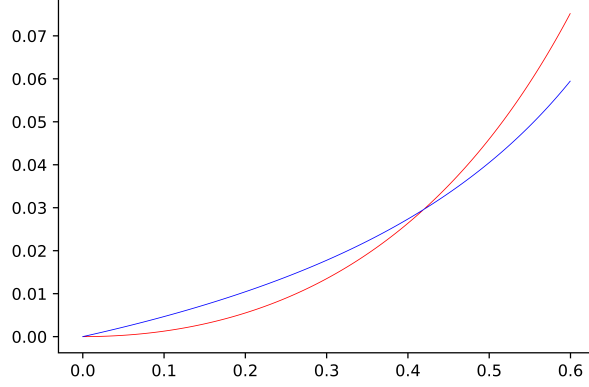$$

Figure 2: Zoom on $(0, 0.6)$ of the pgf of H_{100}, with $\ell = 2$, $\varepsilon = 0.001$ (red) against the pgf of a geometric horizon with the same mean (blue).

Since the ratio in the limit diverges as ε vanishes, this provides an intuitive reason why the single-threshold 2-approximation of Section 5 should fail. In this sense, the hypothesis that ε is small is essential. It is straightforward to compute the derivative of the ratio in the limit and show that it is negative for any ε. Thus this limit decreases as ε grows and tends to $1/3 < 0.594 \approx 1 + W_0(-2e^{-2})$. This means that for ε large enough the single-threshold 2-approximation of Section 5 is still valid, as the distribution H_m has enough concentration in this case. The horizon H_m in Figure 2 instead has CV ≈ 1.074 due to ε being sufficiently small and m large enough. In this case not only the single-threshold 2-approximation of Section 5 is not valid, but no analogous single-threshold C-approximation is possible, as $m \longrightarrow \infty$, since no matter how large C is taken, since $1 + W_0(-Ce^{-C}) \longrightarrow 1 < 1.074$ from below, as $C \longrightarrow \infty$. Furthermore, plotting the pgf of H_m, for ε small and m large, against the pgf of a geometric horizon with the same mean as in Figure 2, shows that there is no ordering of the two (due to the crossing point of the graphs), and therefore this family of horizons is outside the $\mathcal{G}$ (and $PGFd$) class, which explains heuristically why the 2-approximation of Section 3 fails too.

C.4   Proof of Theorem 1.3

Proof of Theorem 1.3.

Step 1.   We start by constructing a family $\tilde{\mathcal{H}}_M(\varepsilon) \subset \overline{\mathcal{G}}$, which is a sequence of horizons $\tilde{H}_m$ arising as a perturbation of horizons H_m in the hard family $\mathcal{H}_M(\varepsilon)$ of (4.2), with fixed $\ell = 2$, as m grows, for all fixed $0 < \varepsilon < 1/4$ small enough. This is done by adding to the pmf of said H_m, a (to be suitably rescaled) mass at one,

$$\delta_m := Cm^{\frac{2\varepsilon}{1+2\varepsilon}}, \quad C = C(\varepsilon) := \frac{3 + 10\varepsilon}{2\varepsilon}. \tag{C.15}$$

We denote the perturbed horizon $\tilde{H}_m$ for every m, ε considered, and observe that the corresponding normalizing constants for the pmf's $\tilde{Z}_m$ and Z_m; and expectations $\tilde{\mu}_m := \mathbb{E}\tilde{H}_m$ and $\mu_m := \mathbb{E}H_m$; satisfy, as m grows, the following:

$$\tilde{Z}_m \sim \tilde{C}Z_m, \quad \tilde{C} := C + 1 + \frac{1}{2\varepsilon} \tag{C.16}$$

$$\tilde{\mu}_m \sim \frac{1}{\tilde{C}}\left(C + \frac{1 + 2\varepsilon}{1 + 4\varepsilon}m\right). \tag{C.17}$$

Trivially, (C.16) and (C.17) follow from (4.3) and (4.4) and the definitions, respectively $\tilde{Z}_m := Z_m + \delta_m$ and $\tilde{\mu}_m = \delta_m/\tilde{Z}_m + \mu_m Z_m/\tilde{Z}_m$. To show that for all m large enough, $\tilde{H}_m \in \overline{\mathcal{G}}$, we have

42

to establish that eventually (short for all m large enough from now on), for all $t \in (0, 1)$ (the implicit range from now on),

$$\frac{1}{\tilde{Z}_m}\left[\delta_m t + \sum_{h=2}^{m} t^h h^{-\frac{1}{1+2\varepsilon}}\right] \geq \frac{1}{\tilde{\mu}_m}\sum_{h=1}^{\infty}(1 - 1/\tilde{\mu}_m)^{h-1} t^h,$$

which we recast as

$$\sum_{h=2}^{m} t^h h^{-\frac{1}{1+2\varepsilon}} \geq -\delta_m t + \frac{\varepsilon_m t}{1 - t(1 - 1/\tilde{\mu}_m)}, \tag{C.18}$$

having set $\varepsilon_m := \tilde{Z}_m/\tilde{\mu}_m$. Let

$$\eta_m := \frac{1 - \varepsilon_m/\delta_m}{1 - 1/\tilde{\mu}_m}.$$

We establish that eventually for all $t \in (0, \eta_m]$

$$-\delta_m t + \frac{\varepsilon_m t}{1 - t(1 - 1/\tilde{\mu}_m)} \leq 0, \tag{C.19}$$

and that eventually for all $t \in [\eta_m, 1)$

$$\sum_{h=2}^{m} h(h-1)t^{h-2}h^{-\frac{1}{1+2\varepsilon}} < \frac{2\varepsilon_m(1 - 1/\tilde{\mu}_m)}{[1 - t(1 - 1/\tilde{\mu}_m)]^3}. \tag{C.20}$$

This implies (C.18) for all $t \in (0, 1)$, since eventually on $(0, \eta_m]$ (C.18) holds with strict inequality directly by (C.19), while for all $t \in [\eta_m, 1)$ we have that the difference of the left-hand and right-hand side of (C.18), denoted $f_m(t)$, is strictly concave ((C.20) states precisely that $f''_m(t) < 0$). Since $f_m(t) > 0$ for all $t \in (0, \eta_m]$ and since it vanishes at both ends of the unit interval, it cannot vanish on $(\eta_m, 1)$, meaning that there cannot be any crossings of the two sides of (C.18) on $(\eta_m, 1)$, which therefore holds everywhere. The rest of this step is dedicated to show (C.19) and (C.20).

In order to establish (C.19), it is enough to note that, upon simplifying the factor t, we can recast it as

$$t \leq \frac{\delta_m - \varepsilon_m}{\delta_m(1 - 1/\tilde{\mu}_m)} = \eta_m. \tag{C.21}$$

Note that

$$\varepsilon_m \sim \frac{\tilde{C}^2 m^{\frac{2\varepsilon}{1+2\varepsilon}}}{C + \frac{1+2\varepsilon}{1+4\varepsilon}m},$$

and therefore we have that

$$\eta_m > 1 - \frac{\varepsilon_m}{\delta_m} = 1 - \frac{\tilde{C}^2}{C\left(C + \frac{1+2\varepsilon}{1+4\varepsilon}m\right)} \longrightarrow 1.$$

In order to establish (C.20), note that a standard integral bound yields that

$$\sum_{h=2}^{m} h(h-1)t^{h-2}h^{-\frac{1}{1+2\varepsilon}} < \sum_{h=2}^{m} h^{2-\frac{1}{1+2\varepsilon}} \leq \int_{2}^{m+1} x^{2-\frac{1}{1+2\varepsilon}}\,dx < \frac{(m+1)^{3-\frac{1}{1+2\varepsilon}}}{3 - \frac{1}{1+2\varepsilon}}$$

on the unit interval considered, and that on $[\eta_m, 1)$

$$\frac{2\varepsilon_m(1 - 1/\tilde{\mu}_m)}{[1 - t(1 - 1/\tilde{\mu}_m)]^3} \geq \frac{2\varepsilon_m(1 - 1/\tilde{\mu}_m)}{[1 - \eta_m(1 - 1/\tilde{\mu}_m)]^3} = 2\frac{\delta_m^3}{\varepsilon_m^2}\left(1 - \frac{1}{\tilde{\mu}_m}\right) \sim \frac{2C^3\frac{1+2\varepsilon}{1+4\varepsilon}}{\tilde{C}^2}m^{2+\frac{2\varepsilon}{1+2\varepsilon}},$$

and therefore it is sufficient to establish that eventually on $[\eta_m, 1)$ we have that

$$\frac{(m+1)^{3-\frac{1}{1+2\varepsilon}}}{3 - \frac{1}{1+2\varepsilon}} < \frac{2C^3 \frac{1+2\varepsilon}{1+4\varepsilon}}{\tilde{C}^2} m^{2+\frac{2\varepsilon}{1+2\varepsilon}}. \tag{C.22}$$

This follows from noting that the powers of the m terms are exactly the same and that the constants of the leading term of the left-hand side is strictly less than that of the right-hand side, that is

$$\frac{1}{3 - \frac{1}{1+2\varepsilon}} < \frac{2C^3 \frac{1+2\varepsilon}{1+4\varepsilon}}{\tilde{C}^2},$$

by the very definition of C and $\varepsilon < 1/4$.

Step 2. Next, we show that with X being set to be the hard instance of values X of (4.1), with $\varepsilon > 0$ fixed small enough as in *Step 1*, $\mathbb{E}M_{\tilde{H}_m} = \Omega(\mathbb{E}M_{H_m})$ and $\mathbb{E}X_{\tau_{\pi_m}} := \mathbb{E}Y^{(\tilde{H}_m)}_{\tau_{\pi_m}} \leq \mathbb{E}Y^{(H_m)}_{\tau_{\pi_m}}$ as m grows, for every sequence of thresholds $\{\pi_m\}$.

The first fact follows from Remark 4.1 and (C.16) and observing that the law of total expectation yields

$$\mathbb{E}M_{\tilde{H}_m} = \frac{\delta_m}{\tilde{Z}_m}\mathbb{E}X + \frac{Z_m}{\tilde{Z}_m}\mathbb{E}M_{H_m} > \frac{Z_m}{\tilde{Z}_m}\mathbb{E}M_{H_m} \sim \frac{1}{\tilde{C}}\mathbb{E}M_{H_m}.$$

The second fact follows from considering the survival functions of $\tilde{H}_m$ and H_m, respectively $\tilde{S}_m(i)$ and $S_m(i)$ and applying Theorem 2.1. We have that

$$S_m(1) = 1, \quad S_m(2) = 1, \qquad S_m(i) = \frac{1}{Z_m}\sum_{h=i}^{m} h^{-\frac{1}{1+2\varepsilon}} \; \forall i \geq 3$$

$$\tilde{S}_m(1) = 1, \quad \tilde{S}_m(2) = 1 - \frac{\delta_m}{\tilde{Z}_m}, \quad \tilde{S}_m(i) = \frac{Z_m}{\tilde{Z}_m}S_m(i) \; \forall i \geq 3. \tag{C.23}$$

Starting with a single-threshold algorithm τ_π, where $\pi = \pi_m$ on horizon $\tilde{H}_m$. By Theorem 2.1 (note that the notations τ and σ used there are now swapped, as our focus here is on the original stopping rule) for any horizon $\tilde{H}_m$, there is a stopping rule equivalent to τ_π for the problem $\mathbf{Y}^{(\tilde{H}_m)}$, denoted as $\sigma_\pi := \zeta_\pi \wedge (\tilde{H}_m + 1)$, where ζ_π is, as per (2.2), an algorithm for the discounted problem $\mathbf{Z} = \{\tilde{S}_m(i)X_i\}$, which stops only when τ_π successfully stops by the time that the horizon $\tilde{H}_m$ has realised. Intuitively, we can construct ζ_π as the stopping rule with single threshold π on the events $\{\tau_\pi = i\} \cap \{\tilde{H}_m \geq i\}$. Otherwise it stops at m on $\{\tau_\pi > \tilde{H}_m\}$. The proof of Theorem 2.1 shows that this can be done, so that ζ_π is adapted to $\mathbf{X}$, and as a result, as m grows,

$$\mathbb{E}Y^{(\tilde{H}_m)}_{\tau_\pi} = \mathbb{E}\sum_{i=1}^{m} \tilde{S}_m(i)X_i \mathbb{1}_{\{\zeta_\pi = i\}} \leq \mathbb{E}\sum_{i=1}^{m} S_m(i)X_i \mathbb{1}_{\{\zeta_\pi = i\}} = \mathbb{E}Y^{(H_m)}_{\tau_\pi}.$$

Indeed, by (C.16) and (C.23) we have that

$$\mathbb{E}Y^{(\tilde{H}_m)}_{\tau_\pi} = -\frac{\delta_m}{\tilde{Z}_m}\mathbb{E}X_2 \mathbb{1}_{\{\zeta_\pi = 2\}} + \mathbb{E}X_1 \mathbb{1}_{\{\zeta_\pi = 1\}} + \mathbb{E}X_2 \mathbb{1}_{\{\zeta_\pi = 2\}} + \frac{Z_m}{\tilde{Z}_m}\mathbb{E}\sum_{i=3}^{m} S_m(i)X_i \mathbb{1}_{\{\zeta_\pi = i\}}$$

$$= \mathbb{E}Y^{(H_m)}_{\tau_\pi}\left(1 - \frac{\delta_m}{\tilde{Z}_m}\right) \leq \mathbb{E}Y^{(H_m)}_{\tau_\pi}.$$

Clearly we also used $\sigma'_\pi := \zeta_\pi \wedge (H_m + 1)$ to obtain the second to last equality, similarly to how we used σ_π in the first equality.

Step 3. Finally, suppose by contradiction that a constant-approximation exists on the $\overline{\mathcal{G}}$ class. Then there exist a sequence of thresholds $\{\pi_m\}$ and $c > 0$ such that τ_{π_m} is a $1/c$-approximation on $\tilde{\mathcal{H}}_M(\varepsilon) \subset \overline{\mathcal{G}}$. By Proposition 4.1 for all $\varepsilon > 0$ fixed small enough, there exists a subsequence $\{\pi_{m_j}\}$ such that r_{m_j}, the gambler-to-prophet ratio of $\mathbb{E}Y_{\tau_{\pi_{m_j}}}^{(\tilde{H}_{m_j})}$ over $\mathbb{E}M_{\tilde{H}_{m_j}}$, vanishes. Combining this with the previous step's asymptotics yields the following contradiction as j grows:

$$c \leq \frac{\mathbb{E}Y_{\tau_{\pi_{m_j}}}^{(\tilde{H}_{m_j})}}{\mathbb{E}M_{\tilde{H}_{m_j}}} = \mathcal{O}(r_{m_j}) \longrightarrow 0.$$

$\square$

C.5 Proof of Theorem 1.2

Proof of Theorem 1.2. Consider the stopping rule $\tau_m := \zeta_{r_m} \wedge (H_m + 1) \in \mathcal{T}^*$. By (2.3) in Theorem 2.1 firstly, and the law of total expectation with respect to $\mathbf{X}$ secondly,

$$\mathbb{E}Y_{\tau_m}^{(H_m)} = \mathbb{E}\left(S_m(\zeta_{r_m})X_{\pi_{\zeta_{r_m}}}\right) = \mathbb{E}\left\{\mathbb{E}\left[S_m(\zeta_{r_m})X_{\pi_{\zeta_{r_m}}}|\mathbf{X}\right]\right\}. \tag{C.24}$$

Having exploited the equivalent discounted deterministic horizon (m) problem spares us from conditioning on H_m, which is now out of the picture and only appears through the survival discount factor of H_m, $S_m(i)$. The only randomness left in the conditional expectation is that of π and the stopping rule ζ_{r_m}.

Consider that

$$\mathbb{E}\left[S_m(\zeta_{r_m})X_{\pi_{\zeta_{r_m}}}|\mathbf{X}\right] = \sum_{i=r_m}^{m}\mathbb{E}\left[S_m(i)X_{\pi_i}\mathbb{1}_{\{\zeta_{r_m}=i\}}|\mathbf{X}\right]. \tag{C.25}$$

Furthermore, denoting $\mathcal{M}_j := \max_{i\in[j]}\{X_{\pi_i}\}$ and noting that $\mathcal{M}_m = M_m$, conditionally on $\mathbf{X}$ we have that

$$\mathbb{1}_{\{\zeta_{r_m}=i\}} = \mathbb{1}_{\{i=\inf\{j\geq r_m:\, X_{\pi_j}=\mathcal{M}_j\}\}} \geq \mathbb{1}_{\{i=\inf\{j\geq r_m:\, X_{\pi_j}=M_m\}\}} := \mathbb{1}_{W_i}.$$

It follows that

$$\mathbb{E}\left[S_m(i)X_{\pi_i}\mathbb{1}_{\{\zeta_{r_m}=i\}}|\mathbf{X}\right] \geq \mathbb{E}\left[S_m(i)X_{\pi_i}\mathbb{1}_{W_i}|\mathbf{X}\right] = S_m(i)M_m\mathbb{P}(W_i|\mathbf{X}). \tag{C.26}$$

Plugging (C.26) into (C.25) yields

$$\mathbb{E}\left[S_m(\zeta_{r_m})X_{\pi_{\zeta_{r_m}}}|\mathbf{X}\right] \geq M_m\sum_{i=r_m}^{m}S_m(i)\mathbb{P}(W_i|\mathbf{X}). \tag{C.27}$$

By (4.3) and comparison of the sum with the corresponding integral as in Lemma 4.2, we can estimate the cdf of H_m

$$F_m(i-1) = \frac{\sum_{h=\ell}^{i-1}h^{-\frac{1}{1+2\varepsilon}}}{\sum_{h=\ell}^{m}h^{-\frac{1}{1+2\varepsilon}}}$$

as $m \longrightarrow \infty$, since for all $i > r_m \sim m/e$ (the lower bound's asymptotics follows from (4.8)) we have that

$$F_m(i-1) \leq \frac{\int_{\ell-1}^{i-1}x^{-\frac{1}{1+2\varepsilon}}dx}{\int_{\ell-1}^{m+1}x^{-\frac{1}{1+2\varepsilon}}dx} = \frac{x^{\frac{2\varepsilon}{1+2\varepsilon}}\big|_{\ell-1}^{i-1}}{x^{\frac{2\varepsilon}{1+2\varepsilon}}\big|_{\ell-1}^{m+1}} = \left(\frac{i-1}{m+1}\right)^{\frac{2\varepsilon}{1+2\varepsilon}}\frac{1-\left(\frac{\ell-1}{i-1}\right)^{\frac{2\varepsilon}{1+2\varepsilon}}}{1-\left(\frac{\ell-1}{m+1}\right)^{\frac{2\varepsilon}{1+2\varepsilon}}} < \left(\frac{i-1}{m+1}\right)^{\frac{2\varepsilon}{1+2\varepsilon}}.$$

$$\tag{C.28}$$

Thus as $m \longrightarrow \infty$,

$$S_m(i) = 1 - F_m(i-1) > 1 - \left(\frac{i-1}{m+1}\right)^{\frac{2\varepsilon}{1+2\varepsilon}}.$$

In conclusion, as $m \longrightarrow \infty$, as $\varepsilon \longrightarrow 0$,

$$\sum_{i=r_m}^{m} S_m(i)\mathbb{P}(W_i|\mathbf{X}) \geq \sum_{i=r_m}^{m} \mathbb{P}(W_i|\mathbf{X}) - \sum_{i=r_m}^{m} \left(\frac{i-1}{m+1}\right)^{\frac{2\varepsilon}{1+2\varepsilon}} \mathbb{P}(W_i|\mathbf{X})$$

$$= \mathbb{P}(W|\mathbf{X}) - \frac{1}{(m+1)^{\frac{2\varepsilon}{1+2\varepsilon}}} \frac{r_m-1}{m} \sum_{i=r_m}^{m} (i-1)^{-\frac{1}{1+2\varepsilon}}$$

$$\geq \frac{1}{e} - \frac{1}{(m+1)^{\frac{2\varepsilon}{1+2\varepsilon}}} \frac{r_m-1}{m} \int_{r_m-2}^{m-1} x^{-\frac{1}{1+2\varepsilon}}\,dx \longrightarrow g(\varepsilon),$$

where we used (4.10) and (4.11) in the equality and (4.9) and the usual integral estimate in the second last inequality, whereas the limiting behaviour as $m \longrightarrow \infty$ follows by (4.8), having defined

$$f_m(\varepsilon) := \frac{1}{(m+1)^{\frac{2\varepsilon}{1+2\varepsilon}}} \frac{r_m}{m} \int_{r_m-2}^{m-1} x^{-\frac{1}{1+2\varepsilon}}\,dx = \frac{1+2\varepsilon}{2\varepsilon} \frac{1}{(m+1)^{\frac{2\varepsilon}{1+2\varepsilon}}} \frac{r_m}{m} x^{\frac{2\varepsilon}{1+2\varepsilon}} \Big|_{r_m-2}^{m-1} \longrightarrow f(\varepsilon)$$

$$:= \frac{1+2\varepsilon}{2\varepsilon} \frac{1}{e} \left(1 - \frac{1}{e^{\frac{2\varepsilon}{1+2\varepsilon}}}\right) = \begin{cases} \frac{1}{e} - \frac{1}{e^2} + \mathcal{O}\left(\frac{1}{\varepsilon}\right), & \varepsilon \longrightarrow \infty \\ \frac{1}{e} + \mathcal{O}(\varepsilon), & \varepsilon \longrightarrow 0, \end{cases}$$

and

$$g(\varepsilon) := \frac{1}{e} - \frac{1+2\varepsilon}{2\varepsilon} \frac{1}{e} \left(1 - \frac{1}{e^{\frac{2\varepsilon}{1+2\varepsilon}}}\right) := \frac{1}{e} - f(\varepsilon) = \mathcal{O}(\varepsilon) > 0$$

as $\varepsilon \longrightarrow 0^+$. It is crucial to note that $g(\varepsilon)$ it is positive and strictly increasing on $(0, \infty)$, since it can be rewritten as $g(\varepsilon) = {}^1\!/e\,(1 - h(t(\varepsilon)))$, having defined

$$t(\varepsilon) := \frac{2\varepsilon}{1+2\varepsilon}, \quad h(t) := \frac{1-e^{-t}}{t}.$$

The monotonicity then follows from the fact that $t(\varepsilon)$ is monotone increasing and upper-bounded by 1, which ensures that $h'(t) < 0$.[21] Plugging the bound obtained into (C.27) yields, as $m \longrightarrow \infty$,

$$\mathbb{E}\left[S_m(\zeta_{r_m})X_{\pi_{\zeta_{r_m}}}\Big|\mathbf{X}\right] \geq (g(\varepsilon)-\delta)M_m,$$

where $\delta \longrightarrow 0^+$ as $m \longrightarrow \infty$. Plugging this into (C.24) yields

$$\mathbb{E}X_{\tau_m} := \mathbb{E}Y_{\tau_m}^{(H_m)} \geq (g(\varepsilon)-\delta)\mathbb{E}M_m.$$

In conclusion, consider the family thus structured. For a fixed $\varepsilon > 0$ small enough, we allow any instance for the values X to undergo the family of horizons $\{H_m\}_{m \geq M}$ as defined in (4.2), where $M = M(\varepsilon) \in \mathbb{N}$ is large enough. By Proposition 4.1, the RH problem under this family does not admit single-threshold constant-approximations for the fixed $\varepsilon > 0$, since for any sequence of thresholds $\{\pi_m\}_{m \geq M}$, we can assume ε small enough, so that there exists a vanishing subsequence

$$r_{m_j} := \frac{\mathbb{E}X_{\tau_{\pi_{m_j}}}}{\mathbb{E}M_{H_{m_j}}},$$

---

[21] As ε grows large, $g(\varepsilon) \longrightarrow e^{-2}$. However, the regime ε large is not relevant to the hardness of single-thresholds, which is only ensured by ε small enough.

as long as we consider the value distribution X defined in (4.1) (recall that here τ_{π_m} refers to the single-threshold algorithm facing horizon H_m and value X with threshold π_m). However, by the above, τ_m provides a constant-approximation. More specifically, regardless of how small ε is fixed, if we set M large enough, τ_m constitutes an approximate $g(\varepsilon)$-approximation, since using $H_m \leq m$ we obtain that for any random variable X

$$\frac{\mathbb{E}X_{\tau_m}}{\mathbb{E}M_{H_m}} \geq (g(\varepsilon) - \delta)\frac{\mathbb{E}M_m}{\mathbb{E}M_m} = g(\varepsilon) - \delta,$$

where $\delta \longrightarrow 0$ as $M \longrightarrow \infty$. $\qquad\square$

D   Proofs omitted from Section 5

Proof of Theorem 1.4. In the argument of Theorem 1.1 we showed that for general distributions the competitive ratio of the single-threshold algorithm used there is

$$c_{p,\bar{q}} = 1 - \mathbb{E}\left[\left(1 - \frac{1}{\mu}\right)^H\right] = 1 - \mathbb{E}\left(e^{-\bar{s}H}\right)$$

with $\bar{s} := -\log(1 - 1/\mu)$. Since $G \prec_{Lt} H$, where G is the two-point distribution of Remark 5.1, that is

$$G \sim \begin{cases} 0, & \text{w.p. } \frac{\sigma^2}{\mu^2+\sigma^2} \\ \frac{\mu^2+\sigma^2}{\mu}, & \text{w.p. } \frac{\mu^2}{\mu^2+\sigma^2}, \end{cases}$$

it follows that

$$c_{p,\bar{q}} \geq 1 - \mathbb{E}\left(e^{-\bar{s}G}\right) = \frac{\mu^2}{\mu^2 + \sigma^2}\left[1 - \left(1 - \frac{1}{\mu}\right)^{\frac{\mu^2+\sigma^2}{\mu}}\right].$$

As a result an inequality that yields a single-threshold $2 - 1/\mu$-approximation is

$$\frac{\mu^2}{\mu^2 + \sigma^2}\left[1 - \left(1 - \frac{1}{\mu}\right)^{\frac{\mu^2+\sigma^2}{\mu}}\right] \geq \frac{1}{2 - \frac{1}{\mu}},$$

and it is equivalent to

$$\left(e^{\mu\log\left(1-\frac{1}{\mu}\right)}\right)^x \leq 1 - \frac{x}{2 - \frac{1}{\mu}}, \tag{D.1}$$

having set $x := 1 + \sigma^2/\mu^2$. Since $\log(1 - 1/\mu) \leq -1/\mu$,

$$e^{\mu\log\left(1-\frac{1}{\mu}\right)} \leq e^{-1}.$$

As a result an even simpler inequality that yields a single-threshold $2 - 1/\mu$-approximation is

$$e^{-x} \leq 1 - \frac{x}{2 - \frac{1}{\mu}}, \tag{D.2}$$

which is equivalent to

$$ye^y \leq \bar{y}e^{\bar{y}}, \tag{D.3}$$

with constraints $y := x - (2 - 1/\mu) > -1$ and $-2 < \bar{y} := -(2 - 1/\mu) < -1$. The solution to the equation $ye^y = z$ is the Lambert function. Over the reals the Lambert function is described through a principal branch, denoted as $W_0(z)$, and a secondary branch, denoted as $W_{-1}(z)$, since $y = -1$ is a branching point for the solution. Since we constrained (D.3) on $y > -1$, $ye^y > -e^{-1}$ is monotone increasing, and for $-2 < \bar{y} < -1$, $-1e^{-1} < \bar{y}e^{\bar{y}} < -2e^{-2}$, we can equivalently formulate (D.3) in terms of W_0, namely

$$y \leq W_0(\bar{y}e^{\bar{y}}).$$

Recalling the definition of y and x, this yields (1.1). $\qquad\square$