

# Optimal Single-Choice Prophet Inequalities from Samples

**Aviad Rubinstein**

Stanford University, Stanford, CA, USA  
aviad@cs.stanford.edu

**Jack Z. Wang**

Cornell University, Ithaca, NY, USA  
jackzwang@cs.cornell.edu

**S. Matthew Weinberg**

Princeton University, Princeton, NJ, USA  
smweinberg@princeton.edu

---

## Abstract

We study the single-choice Prophet Inequality problem when the gambler is given access to samples. We show that the optimal competitive ratio of  $1/2$  can be achieved with a single sample from each distribution. When the distributions are identical, we show that for any constant  $\varepsilon > 0$ ,  $O(n)$  samples from the distribution suffice to achieve the optimal competitive ratio ( $\approx 0.745$ ) within  $(1 + \varepsilon)$ , resolving an open problem of [9].

**2012 ACM Subject Classification** Theory of computation  $\rightarrow$  Online algorithms

**Keywords and phrases** Online algorithms, Probability, Optimization, Prophet inequalities, Samples, Auctions

**Digital Object Identifier** 10.4230/LIPIcs.ITCS.2020.60

**Funding** *S. Matthew Weinberg*: Supported by NSF CCF-1717899.

## 1 Introduction

Consider the classic single-choice Prophet Inequality problem. Offline, there are  $n$  distributions  $\mathcal{D}_1, \dots, \mathcal{D}_n$  presented to a gambler. For  $i = 1$  to  $n$ , a random variable  $X_i$  is drawn independently from  $\mathcal{D}_i$  and revealed online. The gambler must then decide immediately and irrevocably whether to accept  $X_i$  (and achieve reward  $X_i$ , ending the game), or reject  $X_i$  (continuing the game, but never revisiting  $X_i$  again). The goal of the gambler is to devise a stopping rule which maximizes their expected reward. The performance of potential stopping rules is typically measured by their *competitive ratio* in comparison to a prophet (who knows all  $X_i$  in advance and achieves expected reward  $\mathbb{E}[\max_i \{X_i\}]$ ). Typically, prophet inequalities are designed assuming that the distributions presented offline are fully known. This paper focuses on the setting where the gambler is instead presented with offline samples from the  $\mathcal{D}_i$ , rather than complete knowledge.

In the classic setting, seminal work of Krengel and Sucheston provides a strategy guaranteeing a competitive ratio of  $1/2$ , which is the best possible [22].<sup>1</sup> Samuel-Cahn later proved that simply setting a threshold equal to the *median* of  $\max_i \{X_i\}$  (i.e. a value  $T$  such that  $\Pr[\max_i \{X_i\} > v] = 1/2$ ) also achieves the optimal competitive ratio of  $1/2$  [25], and

---

<sup>1</sup> To see that no better than  $1/2$  is possible, consider an instance where  $X_1$  is deterministically 1, and  $X_2$  is  $1/\varepsilon$  with probability  $\varepsilon$ , and 0 otherwise. The prophet achieves  $2 - \varepsilon$  (taking  $X_2$  when it is large, and  $X_1$  otherwise), while the gambler achieves only 1 (they must decide whether to take  $X_1$  without knowing if  $X_2$  is large).



it was later shown that a threshold of  $\mathbb{E}[\max_i\{X_i\}]/2$  suffices as well [21]. These last two thresholds are remarkably simple, but certainly require a non-trivial number of samples to estimate well.

Our first result establishes that *a single sample* from each  $\mathcal{D}_i$  suffices to achieve the optimal competitive ratio of  $1/2$ . The algorithm is also exceptionally simple: if  $\tilde{X}_1, \dots, \tilde{X}_n$  denote independent samples from  $\mathcal{D}_1, \dots, \mathcal{D}_n$ , simply set  $\max_i\{\tilde{X}_i\}$  as a threshold.

► **Definition 1** (Single Sample Algorithm). *Given as input  $\tilde{X}_1, \dots, \tilde{X}_n$ , set a threshold  $T = \max_i\{\tilde{X}_i\}$  and accept the first random variable exceeding  $T$ .*

► **Theorem 2.** *The Single Sample Algorithm guarantees a competitive ratio of  $1/2$ .*

A subsequent line of works considers the special case where each  $\mathcal{D}_i$  is identical (which we'll refer to as  $\mathcal{D}$ ). Here, work of Hill and Kertz provided the first improved competitive ratio (of  $1 - 1/e$ ) [19], and this was recently improved to the optimal competitive ratio of  $\alpha \approx 0.745$  [10]. Our second result establishes that a *linear number of samples* from  $\mathcal{D}$  suffices to achieve the optimal competitive ratio, up to  $\varepsilon$ . Since our algorithm simply replaces the quantile-based thresholds of [10] with samples, we call it Samples-CFHOV (the five authors of [10]). The algorithm and analysis are fairly simple and we provide a formal description in Section 4.

► **Theorem 3.** *With  $O(n/\varepsilon^6)$  samples, Samples-CFHOV achieves a competitive ratio of  $\alpha - O(\varepsilon)$ .*

## 1.1 Related Work

Over the past decade, prophet inequalities have been studied from numerous angles within the TCS community [7, 3, 21, 17, 15, 23, 16, 24, 12, 14, 6, 2, 11, 4, 18, 13, 8]. All of these works assume explicit knowledge of the given distributions. The limited prior work most related to ours considers sample access to the underlying distributions. On this front, Azar et al. consider prophet inequalities subject to combinatorial constraints, and establish that limited samples suffice to obtain constant competitive ratios in many settings [5]. In comparison to this work, our paper considers only *optimal* competitive ratios, and the simple single-choice setting.

In the i.i.d. model (each value is drawn from the same  $\mathcal{D}$ ), a  $(1 - 1/e)$ -approximation was first shown in [19], and recent work achieved the same guarantee with  $n - 1$  samples from  $\mathcal{D}$  [9]. This ratio was later improved to  $\approx 0.738$  [1], and then to  $\alpha \approx 0.745$  [10], where  $\alpha$  is the optimal achievable competitive ratio [19, 20]. The most related work in this sequence to ours is [9], who establish that a competitive ratio of  $\alpha - \varepsilon$  is achievable with  $O(n^2)$  samples (for any constant  $\varepsilon$ ), and that  $\Omega(n)$  samples are necessary. They also establish a formal barrier to achieving  $\alpha - \varepsilon$  with  $o(n^2)$  samples. In comparison, we circumvent their barrier to achieve a competitive ratio of  $\alpha - \varepsilon$  with  $O(n)$  samples, resolving one of their open problems.

## Roadmap

In Section 2, we provide brief preliminaries. Section 3 contains our 2-approximation with a single sample. Section 4 contains our  $(\alpha - \varepsilon)$ -approximation with linearly many samples.

## 2 Preliminaries

There are  $n$  distributions,  $\mathcal{D}_1, \dots, \mathcal{D}_n$ . Online, a gambler sees a random variable  $X_i$  drawn from  $\mathcal{D}_i$  one at a time, and must immediately and irrevocably decide whether to accept (and get reward  $X_i$ ) or reject (and see  $X_{i+1}$ , throwing away  $X_i$  forever). Strategies for a gambler are often termed *stopping rules*, and the competitive ratio of a stopping rule is the worst-case ratio (over all possible  $n$ , and  $\mathcal{D}_1, \dots, \mathcal{D}_n$ ) between its expected reward and  $\mathbb{E}[\max_i \{X_i\}]$ .

Our algorithms will not have knowledge of any  $\mathcal{D}_i$ , but instead will have access to samples. Our algorithms will treat these samples as the only offline input, and decide whether to accept or reject an element based only on the value of that element and the samples.<sup>2</sup> Here, we will count the number of samples from each distribution as our sample complexity.

We will also consider the *i.i.d. setting*, where each  $\mathcal{D}_i = \mathcal{D}$ . Here, we will count the total number of samples from  $\mathcal{D}$  as our sample complexity. In this setting, we will let  $\alpha \approx 0.745$  denote the optimal competitive ratio for an algorithm with knowledge of  $\mathcal{D}$ .

### Continuous versus Discrete Random Variables

All of our algorithm definitions are straight-forward for continuous distributions. For distributions with point masses, the following “reduction” to continuous is needed. Instead of thinking of  $\mathcal{D}$  as a single-variate distribution, we will (overload notation and) think of  $\mathcal{D}$  as a bivariate distribution with the first coordinate drawn from  $\mathcal{D}$ , and the second “tie-breaker” coordinate drawn independently and uniformly from  $[0, 1]$ . Then  $(X_1, t_1) > (X_2, t_2)$  if either  $X_1 > X_2$ , or  $X_1 = X_2$  and  $t_1 > t_2$ . Observe that because the tie-breaker coordinate is continuous, the probability of having  $(X_1, t_1) = (X_2, t_2)$  for any two values during a run of any algorithm is zero. Therefore, if we define  $F_{\mathcal{D}}(X, t) := \Pr_{(Y, u) \leftarrow (\mathcal{D}, U([0, 1]))}[(Y, u) < (X, t)]$ , we have that  $F_{\mathcal{D}}(X, t) < F_{\mathcal{D}}(Y, u) \Leftrightarrow (X, t) < (Y, u)$ . We will not explicitly reference this tie-breaker random variable in the definition of our algorithms, but simply refer to  $X \leftarrow \mathcal{D}$  as the pair  $(X, t)$ .

### Adversaries

Prophet Inequalities are typically studied against an *offline adversary*. That is, the adversary simply picks the distributions  $\mathcal{D}_1, \dots, \mathcal{D}_n$  (and their indices), which is all presented to the gambler offline. Some prophet inequalities hold against the stronger *almighty adversary*, which selects the set of distributions  $\{\mathcal{D}_1, \dots, \mathcal{D}_n\}$  offline, then decides in which order to reveal the random variables  $X_1, \dots, X_n$  based on their realization. Note that previous competitive ratios of  $1/2$  in the non-i.i.d. setting hold against an almighty adversary, and Theorem 2 does as well. Previous competitive ratios of  $\alpha$  in the i.i.d. setting hold against the offline adversary (and are impossible to achieve against the almighty adversary), so Theorem 3 holds against the offline adversary as well.

## 3 The Non-I.I.D. Case: Optimal Ratio with a Single Sample From $\mathcal{D}_i$

The Single Sample Algorithm proceeds as follows. It takes as input  $\tilde{X}_i$  drawn independently from each  $\mathcal{D}_i$ , sets a threshold  $T = \max_i \{\tilde{X}_i\}$ , and accepts the first element exceeding  $T$ . Our goal in this section is to prove Theorem 2 that this algorithm obtains  $\frac{1}{2}$  the reward, in expectation, of the omniscient prophet that always selects the highest value.

<sup>2</sup> In principle, sample-based algorithms might also consider previously viewed elements, but our algorithms don't.

Our analysis will use the principle of deferred decisions: instead of first drawing the samples  $\tilde{X}$ , and then revealing the actual draws  $X$ , we will jointly draw  $2n$  samples  $Y_1, \dots, Y_n, Z_1, \dots, Z_n$ , and then for each  $i$  randomly decide which of  $\{Y_i, Z_i\}$  is equal to  $\tilde{X}_i$  and which is equal to  $X_i$ . Formally, consider the following Deferred-Decisions procedure for drawing  $X, \tilde{X}$ :

1. Draw  $Y_1, \dots, Y_n$  and  $Z_1, \dots, Z_n$  independently each from  $\mathcal{D}_1, \dots, \mathcal{D}_n$ .
2. For ease of notation later, for all  $i$ , relabel so that  $Y_i > Z_i$ .
3. Independently, flip  $n$  fair coins. If coin  $i$  is heads, set  $X_i := Y_i$  and  $\tilde{X}_i := Z_i$ . Otherwise, set  $X_i := Z_i$  and  $\tilde{X}_i := Y_i$ .

► **Observation 4.** *The output of the Deferred-Decisions procedure correctly generates  $\tilde{X}_1, \dots, \tilde{X}_n$  and  $X_1, \dots, X_n$  as independent draws from  $\mathcal{D}_1, \dots, \mathcal{D}_n$ .*

Our analysis will proceed by directly comparing, for any fixed  $Y_1, \dots, Y_n, Z_1, \dots, Z_n$ , the expected reward of the gambler over the randomness in the coin flips of step three to the expected reward of the prophet over the randomness in the coin flips of step three. We note that this analysis is similar to that of the rehearsal algorithm of [5] for  $k$ -uniform matroids (whose competitive ratio is asymptotically optimal for large  $k$ ), and that prior to this it was folklore knowledge that the Single Sample Algorithm achieves a competitive ratio of at least  $1/4$ . The novelty in our analysis is precisely nailing down the tight competitive ratio.

### 3.1 Analysis Setup

For a fixed  $Y_1, \dots, Y_n, Z_1, \dots, Z_n$ , sort the values into descending order, and relabel them as  $W_1, \dots, W_{2n}$ . If  $W_j$  is equal to  $Y_i$  (or  $Z_i$ ), we say that  $W_j$  comes from  $i$ , and denote this with  $\text{index}(W_j) = i$ . Call the *pivotal index*  $j^*$  the minimum  $j$  such that there exists an  $\ell > j$  with  $\text{index}(W_\ell) = \text{index}(W_j)$ . That is, the pivotal index  $j^*$  is such that there are exactly  $j^* - 1$   $Y$  random variables exceeding the largest  $Z$  random variable.

Our analysis will make use of the following concept: for each  $W_1, \dots, W_{j^*-1}$ , let  $C_j$  denote the outcome of the coinflip for  $\text{index}(W_j)$  (which assigns either  $Y_i$  or  $Z_i$  to arrive as a sample and the other to arrive as a real value). Observe, importantly, that the random variables  $C_1, \dots, C_{j^*-1}$  are independent (because they are independent coin flips for different indices). Also importantly, observe that the random variable  $C_{j^*}$  is deterministic conditioned on  $C_1, \dots, C_{j^*-1}$  (because it is exactly the same coin flip as one of the earlier indices).

### 3.2 The Prophet's Expected Reward

► **Proposition 5.** *For fixed  $W_1, \dots, W_{2n}$  and pivotal index  $j^*$ , the prophet's expected reward, over the randomness in the coin flips of step three, is  $\sum_{j=1}^{j^*-1} W_j/2^j + W_{j^*}/2^{j^*-1}$ .*

**Proof.** Observe that the prophet achieves expected reward equal to  $\max_i \{X_i\}$ , so we just want to compute the probability that this is  $W_1, \dots, W_{2n}$ . For each  $j < j^*$ ,  $W_j$  is equal to  $\max_i \{X_i\}$  if and only if  $C_j$  is heads, and  $C_\ell$  is tails for all  $\ell < j$  (recall that all  $W_j$ 's are  $Y$  random variables for  $j < j^*$ ). Because each of the coin flips are independent, this occurs with probability precisely  $1/2^j$ .

For  $j^*$ ,  $W_{j^*}$  is equal to  $\max_i \{X_i\}$  if and only if  $C_\ell$  is tails for all  $\ell < j^*$ , and coin  $C_{j^*}$  is tails. Observe, however, that by definition of the pivotal index  $j^*$ , that when  $C_\ell$  is tails for all  $\ell < j^*$  we have  $C_{j^*}$  as tails as well (because it is the same coin as one of the first  $j^* - 1$ ). Therefore, whenever all of the first  $j^* - 1$  coins are tails,  $\max_i \{X_i\} = W_{j^*}$  (and this happens with probability  $1/2^{j^*-1}$ ).

As the first  $j^* - 1$  coins either contain some heads, or are all tails (and we have counted the prophet's reward in all such cases), we have now fully accounted for the prophet's expected reward over the randomness in the coin flips. ◀

### 3.3 The Single Sample Algorithm's Expected Reward

► **Proposition 6.** *For fixed  $W_1, \dots, W_{2n}$  and pivotal index  $j^*$ , the gambler's expected reward, over the randomness in the coin flips of step three, is at least  $\sum_{j=1}^{j^*-2} W_j/2^{j+1} + W_{j^*-1}/2^{j^*-1}$ .*

**Proof.** Consider the case where  $C_1$  is tails. In this case, the gambler gets no reward because the threshold is higher than all revealed elements. For  $j < j^* - 1$ , consider next the case where  $C_1, \dots, C_j$  are heads, but  $C_{j+1}$  is tails. In this case, the gambler gets reward *at least*  $W_j$  (because the gambler will accept the first non-sample random variable exceeding  $W_{j+1}$ , and these random variables have values  $W_1, \dots, W_j$ ). The probability that this occurs is exactly  $1/2^{j+1}$ .

Consider also the case where  $C_1, \dots, C_{j^*-1}$  are all heads. Then the threshold is set at  $W_{j^*}$ , and the gambler will get at least  $W_{j^*-1}$ . This occurs with probability exactly  $1/2^{j^*-1}$ . ◀

### 3.4 Proof of Theorem 2

**Proof.** We can immediately see that:

$$\begin{aligned} \sum_{j=1}^{j^*-2} W_j/2^{j+1} + W_{j^*-1}/2^{j^*-1} &\geq \sum_{j=1}^{j^*-1} W_j/2^{j+1} + W_{j^*}/2^{j^*} \\ &= \frac{1}{2} \cdot \left( \sum_{j=1}^{j^*-1} W_j/2^j + W_{j^*}/2^{j^*-1} \right). \end{aligned}$$

By Propositions 5 and 6, the right-hand side is exactly half the prophet's expected reward, conditioned on  $W_1, \dots, W_{2n}$  and  $j^*$ , and the left-hand side is exactly the gambler's expected reward (again conditioned on  $W_1, \dots, W_{2n}$  and  $j^*$ ). As the gambler achieves half the prophet's expected reward for all  $W_1, \dots, W_{2n}$  and  $j^*$ , the guarantee holds in expectation as well. ◀

## 4 The I.I.D. Case: Optimal Ratio with Linear Samples From $\mathcal{D}$

We begin with a brief overview of the algorithm from [10] and its main features, followed by a formal specification of our algorithm.

### 4.1 Overview of [10] and Samples-CFHOV

The algorithm of [10] (with one slight modification due to [9]) proceeds as follows. We'll refer to this algorithm as Explicit-CFHOV.

1. As a function only of  $n$ , and independently of  $\mathcal{D}$ , define monotone increasing probabilities  $0 \leq p_1 \leq \dots \leq p_n \leq 1$ .
2. For all  $i$  such that  $p_i \leq \delta = \varepsilon^2/n$ , update  $p_i := 0$  (this is the [9] modification).
3. Accept  $X_i$  if and only if  $F_{\mathcal{D}}(X_i) > 1 - p_i$ . Observe that this is identical to accepting  $X_i$  if and only if  $X_i > \sigma_i = F_{\mathcal{D}}^{-1}(1 - p_i)$ . Also observe that  $X_i$  exceeds  $\sigma_i$  with probability  $p_i$ .

► **Theorem 7** ([10, 9]). *In the i.i.d. setting, Explicit-CFHOV has competitive ratio  $\alpha - \varepsilon$ .<sup>3</sup>*

<sup>3</sup> Without step two, the algorithm achieves a competitive ratio of  $\alpha$ .

That is, Explicit-CFHOV sets, for each  $i \in [n]$ , a probability  $p_i$  independent of  $\mathcal{D}$ , and sets a threshold  $\sigma_i$  for accepting  $X_i$  which is exceeded with probability exactly  $p_i$ .

If instead of explicit access to  $\mathcal{D}$ , we're given  $m$  i.i.d. samples from  $\mathcal{D}$ , the challenge is simply that we can no longer compute  $F_{\mathcal{D}}(X_i)$  exactly and run Explicit-CFHOV. The algorithm of [9] observes that  $m = O(n^2)$  samples suffices to estimate the quantiles sufficiently well. Our algorithm observes that in fact  $m = O(n)$  samples suffice (which is asymptotically tight, by a lower bound in [9]). Our algorithm proceeds as follows, which we call Samples-CFHOV.

1. **As a function only of  $n$ , and independently of  $\mathcal{D}$** , define monotone increasing probabilities  $0 \leq p_1 \leq \dots \leq p_n \leq 1$ , exactly as in Explicit-CFHOV.
2. Round down each  $p_i$  to the nearest integer power of  $(1 + \varepsilon)$ ; we denote the rounded value by  $\lfloor p_i \rfloor \in \{(1 + \varepsilon)^{-1}, (1 + \varepsilon)^{-2}, \dots\}$ .
3. Set  $\tilde{p}_i := \lfloor p_i \rfloor / (1 + \varepsilon)$  (that is, we have rounded down each  $p_i$ , then further divided by  $(1 + \varepsilon)$ ).
4. From our  $m$  samples, let  $\tau_i$  denote the value of the  $(\tilde{p}_i \cdot m)$ -th highest sample.
5. Accept  $X_i$  if and only if  $X_i > \tau_i$ .

That is, Samples-CFHOV provides an estimate  $\tau_i$  of  $\sigma_i$  via the  $m$  samples. Intuitively, we are trying to overestimate  $\sigma_i$  so that it is unlikely that Samples-CFHOV will ever choose to accept an element that Explicit-CFHOV would not. We'll prove Theorem 3 as a corollary of Theorem 8:

► **Theorem 8.** *For any distribution  $\mathcal{D}$ , with  $m = O(n/\varepsilon^6)$  samples, the expected value achieved by Samples-CFHOV is at least a  $(1 - O(\varepsilon))$ -fraction of that of Explicit-CFHOV.*

We briefly remark that our proof of Theorem 8 actually holds for any choice of  $p_i$ 's (all  $\notin (0, \delta)$ ). That is, if Explicit-CFHOV achieves a competitive ratio of  $\gamma(\vec{p})$  with a particular choice of  $\vec{p}$ , Samples-CFHOV achieves a competitive ratio of  $\gamma(\vec{p}) - O(\varepsilon)$  (as long as each  $p_i \notin (0, \delta)$ ).

## 4.2 Brief Comparison to [9]

The algorithm employed by [9] using  $O(n^2)$  samples is conceptually similar in that they also wish to set thresholds  $\tau_i$  such that  $F_{\mathcal{D}}(\tau_i) \approx 1 - p_i$ . The main difference is that we target a multiplicative  $(1 - \varepsilon)$ -approximation to each, whereas they target an additive  $1/n$ -approximation for each threshold. That is, they aim to ensure that for each  $p_i$ , the threshold  $\tau_i$  has  $|F_{\mathcal{D}}(\tau_i) - p_i| \leq 1/n$ . They prove, using the Dvoretzky-Kiefer-Wolfowitz Inequality, that  $O(n^2)$  samples suffice for this, then further argue that these small additive errors in the CDF don't cost much.

The same paper also establishes a barrier to moving beyond  $\Omega(n^2)$  samples. Specifically, they establish that  $\Omega(n^2)$  samples are necessarily just to guarantee for a *single*  $i$  with  $p_i \approx 1/2$  that  $|F_{\mathcal{D}}(\tau_i) - p_i| \leq 1/n$ . Our approach circumvents this bound by seeking a significantly weaker guarantee for such  $i$  (only that  $|F_{\mathcal{D}}(\tau_i) - p_i| \leq O(\varepsilon p_i)$  – see Equation (1)). So the two key differences in our approach is (a) we show that  $O(n)$  samples suffice to learn the thresholds up to a multiplicative  $(1 + \varepsilon)$  error in the CDF and (b) establishing that this (significantly weaker) estimation suffices for a good approximation.

### 4.3 Proof of Theorem 8

Our proof breaks down into two simple claims. The first establishes that with high probability, our sample-based thresholds are “good.” The second establishes that “good” thresholds yield a good approximation. Below, recall that  $\delta = \varepsilon^2/n$ .

► **Lemma 9.** *With  $m = 12 \ln(1/\varepsilon)/(\varepsilon^3\delta) = O(n/\varepsilon^6)$  samples, with probability at least  $1 - \varepsilon$ , we have that for every  $i$  (simultaneously),*

$$\frac{p_i}{(1 + \varepsilon)^3} \leq \Pr_{x \sim \mathcal{D}}[x > \tau_i] \leq p_i. \quad (1)$$

Note that Equation (1) does not reference the values of the actual elements  $X_1, \dots, X_n$  at all – it is just a claim about the thresholds  $\tau$  being set. That is, the probability  $1 - \varepsilon$  is taken only over the randomness in *drawing the samples* (and in particular independent of the values of the actual elements). We will call a set of thresholds “good” if they satisfy Equation (1).

**Proof.** Recall that  $\tau_i$  is set by first rounding down  $p_i$  to  $\lfloor p_i \rfloor$ , then further dividing by  $(1 + \varepsilon)$  to  $\tilde{p}_i$ , then set equal to the  $(\tilde{p}_i \cdot m)$ -th highest of  $m$  samples. To proceed, let  $L_i$  be such that

$$\Pr_{x \sim \mathcal{D}}[x > L_i] = \lfloor p_i \rfloor.$$

Similarly, let  $H_i$  be such that

$$\Pr_{x \sim \mathcal{D}}[x > H_i] = (1 + \varepsilon)^{-2} \lfloor p_i \rfloor.$$

Then (1) certainly holds whenever  $L_i < \tau_i < H_i$ . The remainder of the proof establishes that we are likely to have  $L_i < \tau_i < H_i$  for all  $i$ .

Indeed, observe that we expect to see  $\lfloor p_i \rfloor m$  samples greater than  $L_i$ . We say that  $\lfloor p_i \rfloor$  is *bad* if the number of samples greater than  $L_i$  is *not* in the range  $\left[(1 + \varepsilon)^{-1}(\lfloor p_i \rfloor m), (1 + \varepsilon)(\lfloor p_i \rfloor m)\right]$ . Note that whenever neither  $\lfloor p_i \rfloor$  nor  $(1 + \varepsilon)^{-2} \lfloor p_i \rfloor$  is bad, then we indeed have  $L_i < \tau_i < H_i$ .

Because the number of samples greater than  $p$  is an average of  $m$  independent  $\{0, 1\}$  random variables with expectation  $p$ , the multiplicative Chernoff bound implies that the probability that a particular  $p$  is bad is upper bounded by:

$$\Pr[p \text{ is bad}] < e^{-\varepsilon^2 pm/3}.$$

If all  $p \in \{(1 + \varepsilon)^{-1}, \dots, \delta\}$  are not bad, then our desired claim holds. Taking union bound over this  $(1 + \varepsilon)$ -multiplicative net, we have that the probability that some  $p \in \{(1 + \varepsilon)^{-1}, (1 + \varepsilon)^{-2}, \dots, \delta\}$  is bad is bounded by:

$$\sum_{i=0}^{O(\ln(1/\delta)/\varepsilon)} e^{-\varepsilon^2(1-\varepsilon)^{-i}\delta m/3} \leq \sum_{i=0}^{\infty} e^{-\varepsilon^2(1-\varepsilon)^{-i}\delta m/3} \leq \sum_{i=0}^{\infty} e^{-\varepsilon^3 i \delta m/6} \leq e^{-\varepsilon^3 \delta m/12}$$

Above, the first term is simply a union bound over all  $p$  in this net. The second inequality follows as  $(1 - \varepsilon)^{-i} \geq \varepsilon i/2$  for all  $\varepsilon \in (0, 1)$  and  $i \geq 0$ . The final inequality holds (at least) when  $m \geq 6/(\varepsilon^3\delta)$ . Therefore, setting  $m = 12 \ln(1/\varepsilon)/(\varepsilon^3\delta)$  satisfies the desired claim. ◀

Next, we wish to show that whenever the thresholds are “good”, the algorithm performs well in expectation. Below, let  $t_1$  denote the stopping time of Explicit-CFHOV (i.e. the random variable denoting the element it chooses to accept), and let  $t_2$  denote the stopping time of Samples-CFHOV.



▷ **Claim 10.** Conditioned on (1) holding for every  $i$ ,  $t_2 \geq t_1$ . That is, Samples-CFHOV selects an element *later* than Explicit-CFHOV.

*Proof.* For every  $i$ , we have that by (1), the threshold used by Samples-CFHOV is at least the threshold used by Explicit-CFHOV. Therefore, the first time they deviate (if any) is when Explicit-CFHOV accepts an element but Samples-CFHOV does not. ◁

► **Lemma 11.** *Conditioned on (1) holding for every  $i$ , the following holds for every  $v$ :*

$$\Pr[X_{t_2} > v] \geq \frac{1}{(1+\varepsilon)^3} \Pr[X_{t_1} > v]. \quad (2)$$

**Proof.** We prove that (2) holds uniformly for every  $i \in [n]$ , i.e.

$$\Pr[(X_{t_2} > v) \wedge (t_2 = i)] \geq \frac{1}{(1+\varepsilon)^3} \Pr[(X_{t_1} > v) \wedge (t_1 = i)]. \quad (3)$$

The event  $(X_{t_b} > v) \wedge (t_b = i)$  (for either  $b \in \{1, 2\}$ ) occurs if and only if the corresponding algorithm *doesn't* stop before  $i$ , and  $X_i$  is larger than both  $v$  and the threshold set (by the corresponding algorithm). Of course, whether or not an algorithm stops before  $i$  is completely independent of  $X_i$ . We claim that the following holds on the probability that the two algorithms accept  $X_i$  (conditioned on making it to  $X_i$ ). Intuitively, Claim 12 establishes that, even though the threshold  $\tau_i$  is stricter than  $\sigma_i$ , we are still roughly as likely to accept an  $X_i$  exceeding  $v$  using  $\tau_i$  versus  $\sigma_i$ , for all  $v$ .

▷ **Claim 12.** Conditioned on (1) holding for every  $i$ , then for every  $v$  and  $i$  such that  $p_i \geq \delta$ :

$$(1+\varepsilon)^3 \Pr[(X_i > v) \wedge (X_i > \tau_i)] \geq \Pr[(X_i > v) \wedge (X_i > \sigma_i)].$$

*Proof.* We consider the following three cases: perhaps  $v > \tau_i$ , or perhaps  $v \in (\sigma_i, \tau_i)$ , or perhaps  $v < \sigma_i$ . We claim that the following three inequalities hold:

$$v \geq \tau_i \Rightarrow$$

$$\Pr[(X_i > v) \wedge (X_i > \tau_i)] = \Pr[X_i > v] = \Pr[(X_i > v) \wedge (X_i > \sigma_i)].$$

$$v \in (\sigma_i, \tau_i) \Rightarrow$$

$$\Pr[(X_i > v) \wedge (X_i > \sigma_i)] \leq \Pr[X_i > \sigma_i] \leq \frac{\Pr[X_i > \tau_i]}{(1+\varepsilon)^3} = \frac{\Pr[(X_i > v) \wedge (X_i > \tau_i)]}{(1+\varepsilon)^3}.$$

$$v < \sigma_i \Rightarrow$$

$$\Pr[(X_i > v) \wedge (X_i > \sigma_i)] = \Pr[X_i > \sigma_i] \leq \frac{\Pr[X_i > \tau_i]}{(1+\varepsilon)^3} = \frac{\Pr[(X_i > \tau_i) \wedge (X_i > v)]}{(1+\varepsilon)^3}.$$

Indeed, the first implication follows as  $v$  exceeds both  $\sigma_i$  and  $\tau_i$ . The second implication follows as  $v > \sigma_i$ , and then by condition (1). The third implication follows from condition (1). ◁

Claim 12 is the heart of the proof, and we can now wrap up. Observe that  $\Pr[(X_{t_2} > v) \wedge (t_2 = i)] = \Pr[t_2 \geq i] \cdot \Pr[(X_i > v) \wedge (X_i > \tau_i)]$ . Similarly,  $\Pr[(X_{t_1} > v) \wedge (t_1 = i)] = \Pr[t_1 \geq i] \cdot \Pr[(X_i > v) \wedge (X_i > \sigma_i)]$ . By the work above,  $(1+\varepsilon)^3 \Pr[(X_i > v) \wedge (X_i > \tau_i)] \geq \Pr[(X_i > v) \wedge (X_i > \sigma_i)]$ . By Claim 10,  $\Pr[t_2 \geq i] \geq \Pr[t_1 \geq i]$ . Therefore, we've proven the desired claim for every  $i \in [n]$ . As  $\Pr[X_{t_b} > v] = \sum_i \Pr[(X_{t_b} > v) \wedge (t_b = i)]$ , this completes the proof of Lemma 11. ◀



**Proof of Theorem 8.** The proof of Theorem 8 now follows as a direct corollary of Lemmas 9 and 11. Lemma 11 asserts that whenever the thresholds are “good”, Samples-CFHOV achieves at least a  $1/(1+\varepsilon)^3$  fraction of the expected reward of Explicit-CFHOV (this is because the expected reward of Samples-CFHOV is simply  $\int_0^\infty \Pr[X_{t_2} > v]dv \geq \int_0^\infty \Pr[X_{t_1} > v]dv/(1+\varepsilon)^3$ , and the expected reward of Explicit-CFHOV is precisely  $\int_0^\infty \Pr[X_{t_2} > v]dv$ ). Lemma 9 asserts that the thresholds are good with probability at least  $1-\varepsilon$ . So together, Samples-CFHOV achieves at least a  $\frac{1-\varepsilon}{(1+\varepsilon)^3}$  of the expected reward of Explicit-CFHOV. ◀

## References

- 1 Melika Abolhassani, Soheil Ehsani, Hossein Esfandiari, MohammadTaghi Hajiaghayi, Robert D. Kleinberg, and Brendan Lucier. Beating  $1-1/e$  for ordered prophets. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 61–71, 2017. doi:10.1145/3055399.3055479.
- 2 Marek Adamczyk and Michal Włodarczyk. Random Order Contention Resolution Schemes. In *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018*, pages 790–801, 2018. doi:10.1109/FOCS.2018.00080.
- 3 Saeed Alaei. Bayesian Combinatorial Auctions: Expanding Single Buyer Mechanisms to Many Buyers. In *the 52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 512–521, 2011. doi:10.1109/FOCS.2011.90.
- 4 Nima Anari, Rad Niazadeh, Amin Saberi, and Ali Shameli. Nearly Optimal Pricing Algorithms for Production Constrained and Laminar Bayesian Selection. In *Proceedings of the 2019 ACM Conference on Economics and Computation, EC 2019, Phoenix, AZ, USA, June 24-28, 2019.*, pages 91–92, 2019. doi:10.1145/3328526.3329652.
- 5 Pablo Daniel Azar, Robert Kleinberg, and S. Matthew Weinberg. Prophet Inequalities with Limited Information. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 1358–1377, 2014. doi:10.1137/1.9781611973402.100.
- 6 Yossi Azar, Ashish Chiplunkar, and Haim Kaplan. Prophet Secretary: Surpassing the  $1-1/e$  Barrier. In *Proceedings of the 2018 ACM Conference on Economics and Computation, Ithaca, NY, USA, June 18-22, 2018*, pages 303–318, 2018. doi:10.1145/3219166.3219182.
- 7 Shuchi Chawla, Jason D. Hartline, David L. Malec, and Balasubramanian Sivan. Multi-Parameter Mechanism Design and Sequential Posted Pricing. In *the 42nd ACM Symposium on Theory of Computing (STOC)*, 2010.
- 8 Shuchi Chawla, J. Benjamin Miller, and Yifeng Teng. Pricing for Online Resource Allocation: Intervals and Paths. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2019, San Diego, California, USA, January 6-9, 2019*, pages 1962–1981, 2019. doi:10.1137/1.9781611975482.119.
- 9 José R. Correa, Paul Dütting, Felix A. Fischer, and Kevin Schewior. Prophet Inequalities for I.I.D. Random Variables from an Unknown Distribution. In *Proceedings of the 2019 ACM Conference on Economics and Computation, EC 2019, Phoenix, AZ, USA, June 24-28, 2019.*, pages 3–17, 2019. doi:10.1145/3328526.3329627.
- 10 José R. Correa, Patricio Foncea, Ruben Hoeksma, Tim Oosterwijk, and Tjark Vredeveld. Posted Price Mechanisms for a Random Stream of Customers. In Constantinos Daskalakis, Moshe Babaioff, and Hervé Moulin, editors, *Proceedings of the 2017 ACM Conference on Economics and Computation, EC '17, Cambridge, MA, USA, June 26-30, 2017*, pages 169–186. ACM, 2017. doi:10.1145/3033274.3085137.
- 11 José R. Correa, Raimundo Saona, and Bruno Ziliotto. Prophet Secretary Through Blind Strategies. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2019, San Diego, California, USA, January 6-9, 2019*, pages 1946–1961, 2019. doi:10.1137/1.9781611975482.118.

- 12 Paul Duetting, Michal Feldman, Thomas Kesselheim, and Brendan Lucier. Prophet Inequalities Made Easy: Stochastic Optimization by Pricing Non-Stochastic Inputs. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 540–551, 2017. doi:10.1109/FOCS.2017.56.
- 13 Paul Dütting and Thomas Kesselheim. Posted Pricing and Prophet Inequalities with Inaccurate Priors. In *Proceedings of the 2019 ACM Conference on Economics and Computation, EC 2019, Phoenix, AZ, USA, June 24-28, 2019.*, pages 111–129, 2019. doi:10.1145/3328526.3329576.
- 14 Soheil Ehsani, MohammadTaghi Hajiaghayi, Thomas Kesselheim, and Sahil Singla. Prophet Secretary for Combinatorial Auctions and Matroids. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, pages 700–714, 2018. doi:10.1137/1.9781611975031.46.
- 15 Hossein Esfandiari, MohammadTaghi Hajiaghayi, Vahid Liaghat, and Morteza Monemizadeh. Prophet Secretary. In *Algorithms - ESA 2015 - 23rd Annual European Symposium, Patras, Greece, September 14-16, 2015, Proceedings*, pages 496–508, 2015. doi:10.1007/978-3-662-48350-3\_42.
- 16 Moran Feldman, Ola Svensson, and Rico Zenklusen. Online Contention Resolution Schemes. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016*, pages 1014–1033, 2016. doi:10.1137/1.9781611974331.ch72.
- 17 Oliver Göbel, Martin Hoefer, Thomas Kesselheim, Thomas Schleiden, and Berthold Vöcking. Online Independent Set Beyond the Worst-Case: Secretaries, Prophets, and Periods. In *Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part II*, pages 508–519, 2014. doi:10.1007/978-3-662-43951-7\_43.
- 18 Nikolai Gravin and Hongao Wang. Prophet Inequality for Bipartite Matching: Merits of Being Simple and Non Adaptive. In *Proceedings of the 2019 ACM Conference on Economics and Computation, EC 2019, Phoenix, AZ, USA, June 24-28, 2019.*, pages 93–109, 2019. doi:10.1145/3328526.3329604.
- 19 Theodore P. Hill and Robert P. Kertz. Comparisons of stop rule and supremum expectations of i.i.d. random variables. *The Annals of Probability*, 10:336–345, 1982.
- 20 Robert P. Kertz. Stop rule and supremum expectations of i.i.d. random variables: a complete comparison by conjugate duality. *Journal of Multivariate Analysis*, 19:88–112, 1986.
- 21 Robert Kleinberg and S. Matthew Weinberg. Matroid prophet inequalities. In *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 123–136, 2012. doi:10.1145/2213977.2213991.
- 22 Ulrich Krengel and Louis Sucheston. On Semiamarts, amarts, and processes with finite value. *Advances in Probability and Related Topics*, 4:197–266, 1978.
- 23 Aviad Rubinstein. Beyond matroids: secretary problem and prophet inequality with general constraints. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 324–332, 2016. doi:10.1145/2897518.2897540.
- 24 Aviad Rubinstein and Sahil Singla. Combinatorial Prophet Inequalities. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, Hotel Porta Fira, January 16-19*, pages 1671–1687, 2017. doi:10.1137/1.9781611974782.110.
- 25 Ester Samuel-Cahn. Comparison of threshold stop rules and maximum for independent nonnegative random variables. *Annals of Probability*, 12(4):1213–1216, 1984.