Prophet Inequalities via the Expected Competitive Ratio

TOMER EZRA, Center of Mathematical Sciences and Applications, Harvard University, Cambridge, United States
STEFANO LEONARDI, Computer, Control and Management Sciences, Sapienza University of Rome, Rome, Italy
REBECCA REIFFENHAUSER, Computer Science, University of Amsterdam, Amsterdam, Netherlands
MATTEO RUSSO, Dipartimento di Ingegneria Informatica Automatica e Gestionale, Universita degli Studi di Roma La Sapienza, Rome, Italy
ALEXANDROS TSIGONIAS-DIMITRIADIS, Technical University Munich: Technische Universitat Munchen, Munich, Germany

We consider prophet inequalities under general downward-closed constraints. In a prophet inequality problem, a decision-maker sees a series of online elements with values and needs to decide immediately and irrevocably whether or not to select each element upon its arrival, subject to an underlying feasibility constraint. Traditionally, the decision-maker's expected performance has been compared to the expected performance of the *prophet*, i.e., the expected offline optimum. We refer to this measure as the *Ratio of Expectations* (or, in short, RoE). However, a major limitation of the RoE measure is that it only gives a guarantee against what the optimum would be on average, while, in theory, algorithms still might perform poorly compared to the realized ex-post optimal value. Hence, we study alternative performance measures. In particular, we suggest the *Expectation of Ratio* (or, in short, EoR), which is the expectation of the ratio between the value of the

Tomer Ezra is supported by the Harvard University Center of Mathematical Sciences and Applications. Tomer Ezra was also supported by the National Science Foundation under Grant No. DMS-1928930 and by the Alfred P. Sloan Foundation under grant G-2021-16778, while the author was in residence at the Simons Laufer Mathematical Sciences Institute (formerly MSRI) in Berkeley, California, during the Fall 2023 semester. Stefano Leonardi and Matteo Russo are partially supported by the ERC Advanced Grant 788893 AMDROMA "Algorithmic and Mechanism Design Research in Online Markets", by the FAIR (Future Artificial Intelligence Research) project PE0000013, funded by the NextGenerationEU program within the PNRR-PE-AI scheme (M4C2, investment 1.3, line on Artificial Intelligence), and by the PNRR MUR project IR0000013-SoBigData.it. Stefano Leonardi and Matteo Russo are also partially supported by the MUR PRIN grant 2022EKNE5K (Learning in Markets and Society). Alexandros Tsigonias-Dimitriadis further acknowledges the support of the Alexander von Humboldt Foundation with funds from the German Federal Ministry of Education and Research (BMBF), the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Projektnummer 277991500, the COST Action CA16228 "European Network for Game Theory" (GAMENET), and ANID, Chile, grant ACT210005. Most of this work was done while Alexandros Tsigonias-Dimitriadis was at TU Munich and Universidad de Chile. The views expressed in this paper are Alexandros Tsigonias-Dimitriadis.s and do not necessarily reflect those of the European Central Bank or the Eurosystem.
Authors' Contact Information: Tomer Ezra, Center of Mathematical Sciences and Applications, Harvard University, Cambridge, MA, United States; e-mail: tomer@cmsa.fas.harvard.edu; Stefano Leonardi, Computer, Control and Management Sciences, Sapienza University of Rome, Rome, Italy; e-mail: leonardi@dis.uniroma1.it; Rebecca Reiffenhauser, Computer Science, University of Amsterdam, Amsterdam, Noord-Holland, Netherlands; e-mail: r.e.m.reiffenhauser@uva.nl; Matteo Russo, Dipartimento di Ingegneria Informatica Automatica e Gestionale, Universita degli Studi di Roma La Sapienza, Rome, Lazio, Italy; e-mail: russo.matteo0801@gmail.com; Alexandros Tsigonias-Dimitriadis, Technical University Munich: Technische Universitat Munchen, Munich, Germany; e-mail: alexandrostsigdim@gmail.com.

ACM 2167-8375/2025/04-ART7
https://doi.org/10.1145/3717076

algorithm and the value of the prophet. This measure yields desirable guarantees, e.g., a constant EoR implies achieving a constant fraction of the ex-post offline optimum with constant probability. Moreover, in the single-choice setting, we show that the EoR is equivalent (in the worst case) to the probability of selecting the maximum, a well-studied measure in the literature. However, the probability of selecting the maximum does not generalize meaningfully to combinatorial constraints (beyond single-choice), since its direct equivalent is the probability of selecting an optimal overall set. We, thus, introduce the EoR as a cardinal variant of the probability of selecting the maximum, which extends naturally to combinatorial settings. Our main goal is to understand the relation between RoE and EoR in combinatorial settings. Specifically, we establish two reductions: For every feasibility constraint, the RoE and the EoR are at most a constant factor apart on worst-case instances. Additionally, we show that the EoR is a stronger benchmark than the RoE in that, for every instance (feasibility constraint and product distribution), the RoE is at least a constant fraction of the EoR but not vice versa. Both these reductions imply a wealth of EoR results in multiple settings where RoE results are known.

CCS Concepts: • **Theory of computation** → **Online algorithms**; *Algorithmic mechanism design;*

Additional Key Words and Phrases: Online algorithms, stopping time algorithms, posted price mechanisms

ACM Reference Format:
Tomer Ezra, Stefano Leonardi, Rebecca Reiffenhauser, Matteo Russo, and Alexandros Tsigonias-Dimitriadis. 2025. Prophet Inequalities via the Expected Competitive Ratio. *ACM Trans. Econ. Comput.* 13, 2, Article 7 (April 2025), 32 pages. https://doi.org/10.1145/3717076

1 Introduction

Prophet Inequalities are one of optimal stopping theory's most prominent problem classes. In the classic prophet inequality, a decision-maker must select an element e from an online sequence of elements E immediately and irrevocably. The sequence is revealed one-by-one in an online fashion, and the decision-maker wants to maximize the weight of the chosen element, where each element's weight is drawn from some distribution D_e. The decision-maker knows the distributions and is compared to a *prophet*, who knows all the realizations of the weights in advance. A classic result of Krengel and Sucheston [44, 45] and Samuel-Cahn [53] asserts that the decision-maker can extract at least half of the prophet's expected reward and that this result is tight.

A vast body of research has studied the classic prophet inequality and its variants, where the objective function is to maximize the ratio between what the algorithm gets in expectation and the expected weight of the ex-post optimum. We use the shorthand RoE to signify this *Ratio of Expectations*. However, this benchmark has shortcomings for many applications of prophet inequalities. Oftentimes, the decision-maker is not only concerned about the expected value, but she also wants to have some guarantees with respect to the ex-post outcome.

A first suggestion for a guarantee against ex-post outcomes is the ***Probability of Selecting the Maximum* (PSM)**. This was introduced by Gilbert and Mosteller [34] for the case of i.i.d. valued elements, for which the PSM = 0.5801. In the non-i.i.d. case, in worst-case order, Allaart and Islas [4] show a tight bound on PSM of $1/e$.

Another different performance measure that has been studied for ex-post guarantees in stochastic optimization problems is the ***Expectation of Ratio* (EoR)** between what the algorithm achieves and what the ex-post optimum achieves. This was originally suggested as a performance measure in stochastic minimization problems: Scharbrodt et al. [54] study scheduling algorithms average-case performance through EoR. Garg et al. [33] examine the connection between RoE and EoR in minimization problems, including stochastic Steiner Tree and Traveling Salesman Problem.

We introduce the study of EoR in the Prophet Inequality problem and consider its connection to the PSM. Specifically, in Appendix A, we establish that PSM and EoR are essentially identical

Fig. 1. n pairs of boxes: for each pair, $w_{1,i} = 1$, $w_{2,i} \sim D_2 = \{ \begin{smallmatrix} 0, & \text{w.p. } 1/2 \\ 2, & \text{w.p. } 1/2 \end{smallmatrix}$.

measures of performance in single-choice settings for worst-case instances. This result, however, no longer holds for richer variants of prophet inequalities beyond single-choice. In the natural extension of classic prophet inequalities, termed prophet inequalities with combinatorial constraints [50], the decision-maker is allowed to select more than a single element according to a predefined (downward-closed) feasibility constraint. These types of constraints capture the idea that if a given set is feasible, then so are all its subsets: examples include knapsack, matchings, and general matroids as well as their intersection. In the next example, one can observe that, for any online algorithm, the probability of selecting the maximum is exponentially small in the number of elements; put another way, the probability of selecting the exact optimum offline is negligible and, thus, the guarantees of the PSM measure do not extend to combinatorial prophet inequalities. We formalize these considerations in the next example:

Example 1.1. Consider a setting with n pairs of boxes, such that, for each pair i, one box has weight $w_{1,i} = 1$ deterministically, and the second has weight $w_{2,i}$, equal to 0 with probability 1/2 and to 2 with probability 1/2 (see Figure 1). The (downward-closed) feasibility constraint is that at most one box from each pair can be selected (this is a special case of a partition matroid), and the decision-maker gets the sum over the selected set.

The probability of selecting the maximum is the probability that the algorithm chooses the larger realized value for each pair of boxes. An online algorithm cannot select the maximum of each pair with probability greater than 1/2. Since all realizations are independent, we have an upper bound of $\text{PSM} = 1/2^n$ for every online algorithm (see formal proof in Claim B.4). This motivates choosing a different measure of performance in combinatorial settings. In particular, for this example, the algorithm that always selects the first box for each pair guarantees good expected ratio. Indeed, by Jensen's Inequality (see Appendix C), we have

$$\mathbf{E}\left[\frac{\text{ALG}}{\text{OPT}}\right] \geq \frac{n}{\mathbf{E}\left[\sum_{i\in[n]} \max\{w_{1,i}, w_{2,i}\}\right]} = \frac{n}{\frac{3}{2}n} = \frac{2}{3}.$$

Note that this algorithm also guarantees $\text{ALG} \geq \left(\frac{2}{3} - \varepsilon\right) \cdot \text{OPT}$ with high probability.

Combining that EoR and PSM are equivalent in single-choice settings for worst-case instances and that PSM is unachievable beyond single-choice supports the conclusion that EoR is the "right" alternative to the PSM in combinatorial settings. In Appendix E, we discuss other possible extensions of PSM and show their shortcomings in combinatorial settings.

Another possible interpretation of the EoR is that of a cardinal version of the PSM that guarantees to approximate the ex-post optimum with reasonable probability. More precisely, in Appendix B, we provide a series of claims (Claim B.1, Claim B.2, and Claim B.3), which establish that our definition of EoR (in contrast to the RoE) guarantees the best-we-can-hope-for when comparing to the ex-post optimum. More specifically, Claim B.1 shows that a "good" (i.e., relatively high constant) EoR directly implies that we attain a constant fraction of the optimum with constant probability. Moreover, as shown in Claim B.2 and Claim B.3, no (qualitatively) better bi-criteria approximation can be achieved; there exist simple feasibility constraint-distribution pairs for which either no constant approximation to the maximum is possible with high probability or no near-optimal approximation to the maximum can be attained with constant probability. Therefore, settling for a constant of the ex-post optimum with constant probability (which is

Fig. 2. Two boxes: $w_1 = 1$, $w_2 \sim D_2 = \{ \begin{smallmatrix} 0, \text{ w.p. } 1-\varepsilon \\ \frac{1+2\varepsilon}{\varepsilon}, \text{ w.p. } \varepsilon \end{smallmatrix}$.

Fig. 3. Two boxes: $w_1 = 1$, $w_2 \sim D_2 = \{ \begin{smallmatrix} \varepsilon^2, \text{ w.p. } 1-\varepsilon \\ 1/\varepsilon^2, \text{ w.p. } \varepsilon \end{smallmatrix}$.

guaranteed by having a constant EoR) is the best we can hope for. Moreover, instead of aiming directly for such bi-criteria results, our main goal in this article is to suggest a natural alternative performance measure and present its properties and insights it provides.

We further investigate how the notions of the ratio of expectations and the expectation of ratio are connected to each other. As a first step, we illustrate the behavior of algorithms trying to maximize the EoR and RoE and show that one does not imply the other even in a single-choice setting. First, Example 1.2 shows that an algorithm achieving a constant RoE is not necessarily constant competitive in the EoR sense.

Example 1.2. Consider a setting with two boxes. The first box's weight w_1 is deterministically 1, and the second box's weight w_2 is 0 with probability $1-\varepsilon$ and $\frac{1+2\varepsilon}{\varepsilon}$ with probability ε, for $\varepsilon \in (0,1)$ (see Figure 2).

The canonical 1/2-competitive (and tight) RoE algorithm for the single-choice problem is that of setting a single threshold $\tau := \mathbf{E}[\mathsf{OPT}]/2 = \mathbf{E}[\max\{w_1, w_2\}]/2 = (1-\varepsilon+1+2\varepsilon)/2 = 1+\varepsilon/2$ and accepting the first box whose weight exceeds τ. We now analyze the performance of such algorithm, measured according to EoR:

$$\mathbf{E}\left[\frac{\mathsf{ALG}}{\mathsf{OPT}}\right] = (1-\varepsilon)\cdot\frac{0}{1} + \varepsilon\cdot\frac{(1+2\varepsilon)/\varepsilon}{(1+2\varepsilon)/\varepsilon} = \varepsilon,$$

since the algorithm would only accept if the value is at least $1+\varepsilon/2$, which only happens if the second box realization is $(1+2\varepsilon)/\varepsilon$. This algorithm has no EoR guarantee, since ε can be arbitrarily small.

Second, Example 1.3 shows that an algorithm achieving a constant EoR is not necessarily constant competitive in the RoE sense.

Example 1.3. Consider a setting with two boxes, one with a weight $w_1 = 1$ deterministically, and the second with a weight w_2, which is ε^2 with probability $1-\varepsilon$ and $1/\varepsilon^2$ with probability ε, for $\varepsilon \in (0,1)$ (see Figure 3).

The algorithm that always selects the first box achieves $\mathbf{E}[\frac{\mathsf{ALG}}{\mathsf{OPT}}] = (1-\varepsilon)\cdot\frac{1}{1} + \varepsilon\cdot\frac{1}{1/\varepsilon^2} > 1-\varepsilon$. On the contrary, $\mathbf{E}[\mathsf{ALG}] = 1$ and $\mathbf{E}[\mathsf{OPT}] = 1-\varepsilon+\frac{1}{\varepsilon} > \frac{1}{\varepsilon}$. Thus,

$$\frac{\mathbf{E}[\mathsf{ALG}]}{\mathbf{E}[\mathsf{OPT}]} < \frac{1}{1/\varepsilon} = \varepsilon.$$

As ε can be arbitrarily small, this algorithm does not guarantee a constant RoE.

The aforementioned examples demonstrate that algorithms exhibiting a constant guarantee for one performance measure (such as the optimal RoE algorithm, which gives a guarantee of 1/2

for any instance) might fail miserably in some instances for the other performance measure. This motivates us to deeper understand whether, and for which settings, a good algorithm for RoE can be transformed to a good algorithm for EoR and vice versa. The **guiding question** of this article is, therefore,

What is the relation between RoE and EoR?

1.1 Our Contribution

As a motivation for our study, in Appendix A, we show the equivalence under worst-case instances between PSM and EoR in the single-choice setting. We present two proofs for this equivalence; one is an adaptation of the worst-case example for the PSM measure by Esfandiari et al. [28]. The second proof is based on the observation that for each product distribution, we can construct a new product distribution for which the EoR is arbitrarily close to the PSM of the original distribution.

Our main results establish two reductions between EoR and RoE. In particular, we show that for every downward-closed feasibility constraint, RoE and EoR are at most a multiplicative constant factor apart in the worst-case (see Sections 3 and 4). For the following informal statements of the three main results, we first introduce some basic notation. We use $\mathsf{RoE}(\mathcal{F})$ (similarly $\mathsf{EoR}(\mathcal{F})$) to denote the ratio between the performance of the *best* algorithm against the offline optimum on the *worst-case* distribution, given a family of feasibility constraints $\mathcal{F}$. Note that, in principle, we expect the worst-case distributions to be different for the two measures. In the second and third statements, we use the stronger notions of $\mathsf{RoE}(\mathcal{F}, D)$ and $\mathsf{EoR}(\mathcal{F}, D)$. Here, the ratio expresses the guarantee of the best algorithm against the offline optimum on the specific *constraint-distribution pair* $(\mathcal{F}, D)$. This means that the input now consists not only of a family $\mathcal{F}$, but also of a product (i.e., the distributions of the elements' weights are independent) distribution D.

Theorem (Equivalence between RoE and EoR, Corollary 4.2). *For every downward-closed family of feasibility constraints $\mathcal{F}$, it holds that*

$$\frac{RoE(\mathcal{F})}{EoR(\mathcal{F})} \in \Theta(1).$$

In the next result, we show that the EoR is a stronger benchmark than the RoE in the sense that, for every instance composed of a feasibility constraint and a product distribution, the RoE is at least a constant of the EoR.

Theorem (EoR to RoE Reduction, Theorem 4.1). *For every downward-closed family of feasibility constraints $\mathcal{F}$ and a product distribution D, it holds that*

$$RoE(\mathcal{F}, D) \geq \frac{EoR(\mathcal{F}, D)}{68}.$$

We complement this by showing that the parallel result cannot be achieved in the other direction (i.e., from RoE to EoR).

Theorem (RoE to EoR Impossibility, Corollary 5.1). *For every $\varepsilon > 0$, there exist a feasibility constraint $\mathcal{F}$ and a product distribution D in which $EoR(\mathcal{F}, D) \leq \varepsilon$ and $RoE(\mathcal{F}, D) \in \Omega(1)$.*

1.2 Our Techniques

A key ingredient of our proof is a distinction between cases where the contribution to the value of the prophet comes from a large number of boxes and cases where the contribution mainly comes from a small set of boxes. If we are in the latter case, then one can just run a simple threshold strategy and have a good guarantee. Otherwise, we use our second key ingredient, which is analyzing

the structure of the offline optimum function (value of the prophet) in the event that the threshold algorithm does not have a good enough guarantee. In particular, we show that, under such event, the normalized offline optimum function is *self-bounding* (see Definition 2.3) and, therefore, well-concentrated [11]. Our Proposition 2.1 and Proposition D.1 generalize claims about the self-bounding property of normalized offline optima shown in References [9, 56] (see Appendix D.2 for further discussion).

Feasibility-based reduction: $\mathsf{RoE}(\mathcal{F})$ **vs.** $\mathsf{EoR}(\mathcal{F})$**.** To prove that the $\mathsf{EoR}(\mathcal{F})$ is at least a constant times $\mathsf{RoE}(\mathcal{F})$, we first calculate the threshold for which the maximal value exceeds it with probability of half. We use this to perform a *tail-core* split: Intuitively, if the expected offline optimum is not too large (i.e., close to the threshold), then our algorithm tries to catch a "superstar" (i.e., the first element with a weight above the threshold). To simplify our analysis, we count only the gain in the cases where exactly one such element is realized (the tail event). This happens with constant probability, and since the expected offline optimum is relatively small, we always get a good fraction of it by picking this unique element.

When, instead, the expected offline optimum is large (i.e., far from the threshold), we run a constant competitive RoE algorithm in a black-box fashion. As already pointed out, an algorithm with constant RoE does not necessarily achieve any guarantee for EoR. We overcome this obstacle through our case distinction and the self-bounding properties we show for the optimum. In particular, we upper bound the value of the offline optimum with high enough probability and lower bound the RoE algorithm expected value conditioned on the optimum not being too large.

An immediate corollary of our result and Reference [50] is that, for downward-closed feasibility constraints, the EoR is in $\Omega(\frac{1}{\log^2 n})$. However, we prove a much stronger result: For every specific feasibility constraint, the EoR is a constant away from the RoE (which implies trivially the former assertion). In particular, if for some feasibility constraint the RoE is $\Omega(\frac{1}{\log\log n})$, then our result shows that the EoR is approximately the same (and not that it is just bounded by $\Omega(\frac{1}{\log^2 n})$).

Instance-based reduction: $\mathsf{RoE}(\mathcal{F}, D)$ **vs.** $\mathsf{EoR}(\mathcal{F}, D)$**.** For the other direction, we show a stronger result in that, for every instance (a feasibility constraint, and a product distribution), the RoE is at least a constant fraction of the EoR.

To achieve this result, we show that either the original EoR algorithm achieves up to a constant the same RoE guarantee or that the simple threshold algorithm that achieves half of the expectation of the maximal element guarantees a good RoE.

We remark that both our reductions are constructive: We could interpret them as using an RoE algorithm black-box to design an EoR one and vice versa. Further extensions and implications (XOS objectives, unknown prior, different assumptions on the arrival order, etc.) are discussed in Section 6.

1.3 Related Work

For early work on prophet inequalities, starting from the classic model and some of its most important variants, we refer the reader to the comprehensive survey of Hill and Kertz [38]. The topic of prophet inequalities has recently regained strong interest, primarily among researchers in theoretical computer science, due to its connections to (algorithmic) mechanism design and, in particular, posted price mechanisms [14, 18, 36, 43]. The surveys of Correa et al. [17] and Lucier [46] provide detailed overviews of recent results in prophet inequalities and their connections to mechanism design, respectively.

This recent surge of interest has given rise to a stream of work, extending the classic prophet inequality to more general objective functions beyond single-choice (including submodular [15], XOS [31] and monotone subadditive functions [51]), different assumptions on the arrival order,

and rich combinatorial feasibility constraints. Among the latter, some notable results include k-uniform matroid [2, 36, 40], matching [3, 29, 31, 35], general matroid or knapsack [14, 23, 26, 32, 43], and polymatroid constraints [25]. Among the most general environments considered (in which non-trivial positive results can be achieved) are arbitrary downward-closed feasibility constraints [50, 51]. Combined with our framework, these results immediately give corresponding (lower and upper) bounds on the EoR. Note that Reference [50] also considers non-downward-closed feasibility constraints but shows that it is impossible to achieve an RoE larger than $O(1/n)$.

One of our goals in this article is to go beyond the traditional measure of performance in prophet inequalities, i.e., the ratio of expectations, and understand how natural alternatives perform in a wide range of scenarios. While the EoR measure has not been studied before in the context of prophet inequalities, Garg et al. [33] considered it for Bayesian cost minimization problems, such as the Online Stochastic Steiner tree problem, where they show an upper bound of $O(\log \log n)$ on the gap between RoE and EoR: Whether this gap is constant is an open question up to this day. Furthermore, Hartline and Johnsen [37, Appendix A.1] study a similar notion to the EoR and compare it to the RoE in the context of prior independent mechanism design. Their goal is to measure the performance of an algorithm without knowledge of the input distribution against the best algorithm with full distributional knowledge.

In the prophet inequalities literature, a common underlying assumption is that the decision-maker has full knowledge of the distributions from where the values of the arriving elements are drawn. This is arguably a strong assumption in many practical applications; therefore, a parallel line of work has focused on settings where the distributions are unknown, but a limited number of samples from these distributions is available to the decision-maker. Azar et al. [6] pioneered this idea and showed positive results in several combinatorial settings. In fact, in the classic prophet inequality, just a single sample from each distribution suffices to recover the tight result with full distributional knowledge [52]. Similar insights are obtained when the distributions are identical; this problem was initially studied by Correa et al. [21] and subsequently improved several times [16, 20, 22, 41, 52]. Azar et al. [6], Caramanis et al. [13], Kaplan et al. [41] extended the single-sample framework (which can be viewed as the minimum amount of available information) to several combinatorial problems.

Directly related to prophet inequalities, a limited number of recent papers consider different performance metrics; we can loosely divide them into two main categories. In the first one [5, 12, 24, 30, 47, 49], the goal is to compare against the computationally unbounded optimal online policy. Similar in spirit is the work of Reference [1], where the decision-maker can choose the ordering of the elements, and the main question is whether finding the optimal ordering can be done in polynomial time. In general, given an online Bayesian selection problem, the natural questions are whether it is hard to compute an optimal solution and, if that is the case, how well can we approximate this benchmark with polynomial-time algorithms. The second group of papers [20, 28, 30, 48] studies single-choice problems with the goal of maximizing the probability of picking the element with the highest value. Note that this is the objective of the secretary problem, but in problems with stochastic input (as is the case in prophet inequalities), i.e., where elements are generated from distributions. Alternative measures of performance have also been proposed to capture the behavior of biased (as opposed to rational) agents [42] or to address fairness considerations [19].

2 Preliminaries

2.1 Model and Notation

We consider a setting where there is a ground set of elements E, and each element $e \in E$ is associated with a non-negative weight $w_e \sim D_e$. We assume that the distributions have no point

masses,[1] and we denote by $D = \times_{e \in E} D_e$ the product distribution. The elements are presented with their weights in an online fashion to a decision-maker who needs to decide immediately and irrevocably whether to accept the current element or not. The decision-maker must ensure that the set of selected elements belongs to a predefined family of downward-closed feasibility constraints $\mathcal{F}$ at all times. The goal of the decision-maker is to maximize the weight of the selected set. We make use of the following definitions and notations:

Definition 2.1. Let $\boldsymbol{w} \in \mathbb{R}_{\geq 0}^{|E|}$ be a non-negative weight vector. We define $\mathsf{OPT} : \mathbb{R}_{\geq 0}^{|E|} \to \mathcal{F}$ to be the function mapping a vector of weights to a maximum-weight set in family $\mathcal{F}$. Namely,

$$\mathsf{OPT}(\boldsymbol{w}) = \arg\max_{S \in \mathcal{F}} \sum_{e \in S} w_e.$$

Moreover, we abuse notation of vector $\boldsymbol{w}$ and use $\boldsymbol{w}(S)$ to denote the sum of weights in set S, i.e.,

$$\boldsymbol{w}(S) := \sum_{e \in S} w_e.$$

Definition 2.2. Given a downward-closed family of feasible sets $\mathcal{F}$, we define $f_{\mathcal{F}} : \mathbb{R}_{\geq 0}^{|E|} \to \mathbb{R}$ to be the function that, given a weight vector $\boldsymbol{w}$, returns the maximal weight of a feasible set in $\mathcal{F}$, i.e., $f_{\mathcal{F}}(\boldsymbol{w}) := \boldsymbol{w}(\mathsf{OPT}(\boldsymbol{w}))$. When clear from context, we omit $\mathcal{F}$ and use f instead of $f_{\mathcal{F}}$.

Given an online algorithm ALG, we denote the (possibly random) set chosen by it given an input $\boldsymbol{w}$ by $\mathsf{ALG}(\boldsymbol{w})$. We will denote by $a_{\mathsf{ALG}}(\boldsymbol{w}) := \boldsymbol{w}(\mathsf{ALG}(\boldsymbol{w}))$ the weight of the feasible set chosen by the algorithm for a specific realization $\boldsymbol{w}$, and when clear from context, we omit ALG from the notation and use $a(\boldsymbol{w})$ instead of $a_{\mathsf{ALG}}(\boldsymbol{w})$. Our objective is to design algorithms that maximize the expected ratio between what the online algorithm gets and the weight of the offline optimum. To measure our performance, given a downward-closed family $\mathcal{F}$, a product distribution D, and an algorithm ALG, we define

$$\mathsf{EoR}(\mathcal{F}, D, \mathsf{ALG}) := \mathbf{E}\left[\frac{\boldsymbol{w}(\mathsf{ALG}(\boldsymbol{w}))}{f(\boldsymbol{w})}\right],$$

where the expectation runs over the stochastic generation of the input as well as the (possible) randomness of the algorithm. Similarly, we define

$$\mathsf{EoR}(\mathcal{F}, D) := \sup_{\mathsf{ALG}} \mathsf{EoR}(\mathcal{F}, D, \mathsf{ALG}), \tag{1}$$

and

$$\mathsf{EoR}(\mathcal{F}) := \inf_{D} \mathsf{EoR}(\mathcal{F}, D). \tag{2}$$

We will compare our results to the standard objective of maximizing the ratio of expectations between the algorithm and the offline optimum. Accordingly, we denote

$$\mathsf{RoE}(\mathcal{F}, D, \mathsf{ALG}) := \frac{\mathbf{E}\left[\boldsymbol{w}(\mathsf{ALG}(\boldsymbol{w}))\right]}{\mathbf{E}\left[f(\boldsymbol{w})\right]}.$$

The final benchmark we will compare our results to is the probability of selecting an optimal (offline) set, defined by

$$\mathsf{PSM}(\mathcal{F}, D, \mathsf{ALG}) := \Pr\left[\boldsymbol{w}(\mathsf{ALG}(\boldsymbol{w})) = f(\boldsymbol{w})\right].$$

Analogously to Equations (1), and (2), we define $\mathsf{RoE}(\mathcal{F}, D)$, $\mathsf{RoE}(\mathcal{F})$, $\mathsf{PSM}(\mathcal{F}, D)$, and $\mathsf{PSM}(\mathcal{F})$.

[1]We assume that there are no point masses for simplicity of presentation. All of our theorems can be adjusted to the case where there are point masses.

2.2 Structural Properties

As a first step before stating and proving the main results, we derive several properties of f that may be of interest beyond this article. The main technical tool that we use throughout to guarantee only a constant-factor loss in the reduction is the self-bounding property of the (normalized) offline optimum. Since this property resembles a "smoothness" condition when removing one of the coordinates of the input vector, we can only prove it if we restrict the support of the weights.

Definition 2.3 (Self-bounding Functions). Let $\boldsymbol{x} := (x_1, \ldots, x_n)$ be a vector of independent random variables and $\mathcal{X}$ the corresponding product space. Similarly, let $\boldsymbol{x}^{(i)} := (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ be the same vector deprived of the ith coordinate and $\mathcal{X}^{(i)}$ the corresponding product space. A function $g : \mathcal{X} \to \mathbb{R}$ is said to be self-bounding if there exists a series of functions $\{g_i : \mathcal{X}^{(i)} \to \mathbb{R}\}_{1 \le i \le n}$, such that:

$$\forall i : 0 \le g(\boldsymbol{x}) - g_i(\boldsymbol{x}^{(i)}) \le 1,$$
$$\text{and} \sum_{i \in [n]} \left(g(\boldsymbol{x}) - g_i(\boldsymbol{x}^{(i)}) \right) \le g(\boldsymbol{x}).$$

Proposition 2.1 (Properties of $f_{\mathcal{F}}$). *For every downward-closed family of sets $\mathcal{F}$, the function f $(= f_{\mathcal{F}})$ satisfies the following properties:*

(1) f is 1-Lipschitz.
(2) f is monotone, i.e., if $\boldsymbol{u} \ge \boldsymbol{v}$ point-wise, then $f(\boldsymbol{u}) \ge f(\boldsymbol{v})$.
(3) For every $\tau > 0$, the function f/τ restricted to the domain $[0, \tau]^{|E|}$ is self-bounding.

Proof. Let $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^{|E|}_{\ge 0}$ be nonnegative weight vectors and $\mathsf{OPT}(\boldsymbol{u}), \mathsf{OPT}(\boldsymbol{v})$ be, respectively, their optimal feasible subset.

(1) **Lipschitz.** We assume, without loss of generality, that $f(\boldsymbol{u}) \ge f(\boldsymbol{v})$. Then,

$$\begin{aligned} \mid f(\boldsymbol{u}) - f(\boldsymbol{v}) \mid &= f(\boldsymbol{u}) - f(\boldsymbol{v}) \\ &= \sum_{e \in \mathsf{OPT}(\boldsymbol{u})} u_e - \sum_{e \in \mathsf{OPT}(\boldsymbol{v})} v_e \\ &\le \sum_{e \in \mathsf{OPT}(\boldsymbol{u})} u_e - \sum_{e \in \mathsf{OPT}(\boldsymbol{u})} v_e \\ &\le \sum_{e \in \mathsf{OPT}(\boldsymbol{u})} \mid u_e - v_e \mid \\ &\le \sum_{e \in E} |u_e - v_e| = \|\boldsymbol{u} - \boldsymbol{v}\|_1, \end{aligned}$$

where the first inequality is by Definition 2.1. We conclude that f is 1-Lipschitz.

(2) **Monotonicity.** Let $\boldsymbol{u}, \boldsymbol{v}$ be two vectors such that $\boldsymbol{v} \le \boldsymbol{u}$ component-wise. It holds that

$$f(\boldsymbol{v}) = \boldsymbol{v}(\mathsf{OPT}(\boldsymbol{v})) \le \boldsymbol{u}(\mathsf{OPT}(\boldsymbol{v})) \le \boldsymbol{u}(\mathsf{OPT}(\boldsymbol{u})) = f(\boldsymbol{u}),$$

where the first inequality is by our assumption that $\boldsymbol{u} \ge \boldsymbol{v}$, and the second inequality is by the definition of OPT.

(3) **Self-boundedness.** For every element $e \in E$, we define $\boldsymbol{w}^{(e)} := (w_1, \ldots, w_{e-1}, w_{e+1}, \ldots, w_{|E|})$ to be the weight vector $\boldsymbol{w}$ deprived of eth coordinate and $f_e : \mathbb{R}^{|E|-1}_{\ge 0} \to \mathbb{R}$ to be the function f that takes as input a weight vector $\boldsymbol{w}^{(e)}$ and apply f to the vector $(\boldsymbol{w}^{(e)}, w_e = 0)$. Thus, f/τ in the range of $[0, \tau]^E$ is self-bounding since:

— The first condition of Definition 2.3 that $0 \le \frac{1}{\tau} f(\boldsymbol{w}) - \frac{1}{\tau} f_e(\boldsymbol{w}^{(e)}) \le 1$ is satisfied by monotonicity, the 1-Lipschitzness, and the fact that w_e is restricted to the range of $[0, \tau]$.

— The second condition of Definition 2.3 holds, since

$$\sum_{e\in E}\left(\frac{1}{\tau}f(\boldsymbol{w})-\frac{1}{\tau}f_e(\boldsymbol{w}^{(e)})\right)=\sum_{e\in \mathrm{OPT}(\boldsymbol{w})}\left(\frac{1}{\tau}f(\boldsymbol{w})-\frac{1}{\tau}f_e(\boldsymbol{w}^{(e)})\right)\le\frac{1}{\tau}\cdot\sum_{e\in \mathrm{OPT}(\boldsymbol{w})}w_e=\frac{1}{\tau}\cdot f(\boldsymbol{w}),$$

where the first equality derives from the fact that if $e\notin \mathrm{OPT}(\boldsymbol{w})$, then $f(\boldsymbol{w})-f_e(\boldsymbol{w}^{(e)})=0$, the inequality is by the 1-Lipschitzness, and the last equality is by the definition of f.

This concludes the proof. □

In Appendix D.2, we generalize the above proposition to arbitrary (extended) XOS functions. The main attribute of self-bounding functions that we will use is the following inequality:

Theorem 2.1 (BLM Inequality [10]). *For a self-bounding function $g:\mathcal{X}\to\mathbb{R}$, it holds that:*

$$\text{For every } z>0,\qquad \Pr\left[g(\boldsymbol{x})\ge \mathbf{E}[g(\boldsymbol{x})]+z\right]\le e^{-\frac{3z^2}{6\mathbf{E}[g(\boldsymbol{x})]+2z}}.$$

$$\text{For every } z<\mathbf{E}[g(\boldsymbol{x})],\quad \Pr\left[g(\boldsymbol{x})\le \mathbf{E}[g(\boldsymbol{x})]-z\right]\le e^{-\frac{z^2}{2\mathbf{E}[g(\boldsymbol{x})]}}.$$

Given a distribution D and an event I, we denote by $D_{|I}$ the conditional distribution of D conditioned on event I.

3 From RoE to EoR: Feasibility-based Reduction

Before presenting our reduction from RoE to EoR, we start with a few definitions and observations. We defer their proofs, as well as other auxiliary claims, to Appendix C. Fixing a parameter $\gamma\in(0,1)$, we define the threshold τ given a set of elements E with corresponding distributions $\{D_e\}_{e\in E}$, to be such that

$$\Pr\left[\tau\ge\max_{e\in E}w_e\right]=\gamma. \tag{3}$$

Such a τ exists and is unique for every γ, since we assume that there are no point masses. For every $e\in E$, we denote by $\overline{D}_e$ the distribution $D_{e|w_e\le\tau}$, as per the τ defined in Equation (3). This is well defined, since $\gamma>0$, and, therefore, the probability that $w_e\le\tau$ is at least $\gamma>0$. Given a realization of $\boldsymbol{w}$, let $\overline{\boldsymbol{w}}\in\mathbb{R}^{|E|}_{\ge 0}$ be the weight vector determined by the following process: For each e, if $w_e\le\tau$, then $\overline{w}_e=w_e$; otherwise, let $\overline{w}_e$ be a fresh (independent) draw from $\overline{D}_e$. Note that the distribution of $\overline{w}$ is a product distribution, where for each $e\in E$, $\overline{w}_e$ is distributed according to $\overline{D}_e$.

We next define the two events that we will use in our analysis.

Definition 3.1. Let us define the following events:

(1) **Core.** $\mathcal{E}_0:=\{\forall e\in E: w_e\le\tau\}$.
(2) **Tail.** $\mathcal{E}_1:=\{\exists! e\in E: w_e>\tau\}$,

where the symbol "$\exists!$" signifies the existence of a unique such element.

The next observation enables us to flexibly change whenever needed from conditioning on $\mathcal{E}_0$ to working directly with the truncated distribution and vice versa.

Observation 3.1. *The distribution $\overline{D}$ is identical to the distribution D conditioned on event $\mathcal{E}_0$.*

We are now ready to present our reduction from RoE to EoR.

Theorem 3.1. *For every downward-closed family of feasibility constraints $\mathcal{F}$, it holds that*

$$EoR(\mathcal{F})\ge\frac{RoE(\mathcal{F})}{12}. \tag{4}$$

ALGORITHM 1: RoE-to-EoR

Data: Ground set E, distributions D_e, feasibility family $\mathcal{F}$, a subroutine $\mathsf{ALG}_{\mathsf{RoE}}$
Parameters: $\gamma \in (0,1)$, $c > 0$
Assumption: $\mathsf{ALG}_{\mathsf{RoE}}$ satisfies $\mathbf{E}\left[a_{\mathsf{ALG}_{\mathsf{RoE}}}(\boldsymbol{w}) \mid \mathcal{E}_0\right] \geq \alpha \cdot \mathbf{E}\left[f(\boldsymbol{w}) \mid \mathcal{E}_0\right]$
Result: Subset $\mathsf{ALG}(\boldsymbol{w}) \subseteq E$ such that $\mathsf{ALG}(\boldsymbol{w}) \in \mathcal{F}$
Calculate τ according to Equation (3);
Calculate $W := \mathbf{E}\left[f(\overline{\boldsymbol{w}})\right]$;
if $W \leq c \cdot \tau$ **then**
  Return the first element $e^* \in E$ such that $w_{e^*} \geq \tau$;
  If no such element exists, return $\emptyset$;
else
  As long as $w_e \leq \tau$ run $\mathsf{ALG}_{\mathsf{RoE}}(\boldsymbol{w})$;
  If $w_e > \tau$, reject all remaining elements;
end

Note that in the reduction of Algorithm 1, part of the input is a subroutine $\mathsf{ALG}_{\mathsf{RoE}}$ that has an RoE at least as large as α. The condition on the event $\mathcal{E}_0$ just means that all we need to know is that this α-guarantee holds when all weights are below the chosen threshold τ. Starting from $\mathsf{ALG}_{\mathsf{RoE}}$, we design an algorithm that uses $\mathsf{ALG}_{\mathsf{RoE}}$ as a black box in one of the two cases and achieves an EoR, which is at most a multiplicative constant factor away from the RoE.

To prove Theorem 3.1, by the definition of $\mathsf{RoE}(\mathcal{F})$, we will assume the existence of an algorithm $\mathsf{ALG}_{\mathsf{RoE}}$ that satisfies:

$$\begin{aligned}\mathbf{E}_{\boldsymbol{w}\sim D}\left[a_{\mathsf{ALG}_{\mathsf{RoE}}}(\boldsymbol{w}) \mid \mathcal{E}_0\right] &= \mathbf{E}_{\overline{\boldsymbol{w}}\sim \overline{D}}\left[a_{\mathsf{ALG}_{\mathsf{RoE}}}(\boldsymbol{w})\right] \\ &\geq \alpha \cdot \mathbf{E}_{\overline{\boldsymbol{w}}\sim \overline{D}}\left[f(\overline{\boldsymbol{w}})\right] = \alpha \cdot \mathbf{E}_{\boldsymbol{w}\sim D}\left[f(\boldsymbol{w}) \mid \mathcal{E}_0\right],\end{aligned} \tag{5}$$

where the equalities hold by Observation 3.1. Our analysis distinguishes between two cases according to whether $W := \mathbf{E}\left[f(\overline{\boldsymbol{w}})\right] \leq c \cdot \tau$. Lemma 3.1 analyzes the case where $W \leq c \cdot \tau$ and Lemma 3.2 the case where $W > c \cdot \tau$. In the remainder, for ease of notation, we use $a(\boldsymbol{w})$ instead of $a_{\mathsf{ALG}_{\mathsf{RoE}}}(\boldsymbol{w})$.

Lemma 3.1 ("Catch the Superstar"). *For all constants $\delta \geq 2, k \geq 3$ and $c \geq \frac{8}{3}\log\frac{k}{\alpha}$, if $W \leq c \cdot \tau$, then Algorithm 1 satisfies*

$$\mathbf{E}\left[\frac{a(\boldsymbol{w})}{f(\boldsymbol{w})}\right] \geq \frac{\gamma \log(1/\gamma)}{c+1}. \tag{6}$$

Proof. In this scenario, we know that Algorithm 1 will select the first element $e \in E$ such that $w_e \geq \tau$, which we denote by e^*. It may happen that no such element exists, in which case the algorithm gets a contribution of 0. However, by Claim C.5 (deferred to Appendix C), $\Pr\left[\mathcal{E}_1\right] \geq \gamma \log(1/\gamma)$. Then, conditioned on this event, Algorithm 1 surely (with probability 1) selects e^*, and we have

$$\begin{aligned}\frac{a(\boldsymbol{w})}{f(\boldsymbol{w})} &\geq \frac{w_{e^*}}{w_{e^*} + f(\overline{\boldsymbol{w}})} \\ &\geq \frac{\tau}{\tau + f(\overline{\boldsymbol{w}})},\end{aligned} \tag{7}$$

where the first inequality follows from Claim C.4 (deferred to Appendix C) and the second by observing that, since $w_e > \tau$, the ratio is minimized when $w_e = \tau$. Hence, we get that

$$\begin{aligned}\mathbf{E}\left[\frac{a(\boldsymbol{w})}{f(\boldsymbol{w})}\right] &\geq \mathbf{E}\left[\frac{a(\boldsymbol{w})}{f(\boldsymbol{w})}\,\middle|\,\mathcal{E}_1\right]\cdot \Pr\left[\mathcal{E}_1\right]\\ &\geq \mathbf{E}\left[\frac{\tau}{\tau+f(\overline{\boldsymbol{w}})}\right]\cdot \gamma\log\left(1/\gamma\right)\\ &\geq \frac{\tau}{\tau+W}\cdot\gamma\log\left(1/\gamma\right) \geq \frac{\gamma\log\left(1/\gamma\right)}{c+1},\end{aligned}$$

where the first inequality is by the law of total expectation, the second follows from Equation (7) and Claim C.5, the third is by Jensen's inequality (see Claim C.2), and the last inequality is due to our assumption that $W \leq c\cdot\tau$. □

In the following, Lemma 3.2, Claim 3.1, and Claim 3.2 are mainly dedicated to expressing the expected ratio through various manipulations in a convenient form, such that the concentration property of the offline optimum (see Theorem 2.1) can be repeatedly applied.

Lemma 3.2 ("Run the Combinatorial Algorithm"). *For all constants $\delta \geq 2, k \geq 3$, and $c \geq \frac{8}{3}\log\frac{k}{\alpha}$, then Algorithm 1 satisfies*

$$\mathbf{E}\left[\frac{a(\boldsymbol{w})}{f(\boldsymbol{w})}\right] \geq \frac{\gamma}{\delta}\frac{k-\delta}{k}\alpha. \tag{8}$$

To prove Lemma 3.2, we will make use of the two following claims:

Claim 3.1. *For all constants $\delta \geq 2, k \geq 3$ and $c \geq \frac{8}{3}\log\frac{k}{\alpha}$, if $W > c\cdot\tau$, then we have*

$$\Pr\left[f(\boldsymbol{w}) > \delta W \mid \mathcal{E}_0\right] \leq \frac{\alpha}{k}.$$

Proof. By Theorem 2.1, choosing $z = \frac{1}{\tau}(\delta-1)W$, we have

$$\Pr\left[f(\boldsymbol{w}) > \delta W \mid \mathcal{E}_0\right] = \Pr\left[\frac{f(\boldsymbol{w})}{\tau} > \frac{\delta W}{\tau}\,\middle|\,\mathcal{E}_0\right] \leq e^{-\frac{3(\delta-1)^2}{4+2\delta}\frac{W}{\tau}} \leq e^{-\frac{3(\delta-1)^2}{4+2\delta}c} \leq \frac{\alpha}{k},$$

since $W > c\cdot\tau$ by assumption, and since $c \geq \frac{8}{3}\log\frac{k}{\alpha} \geq \frac{4+2\delta}{3(\delta-1)^2}\log\frac{k}{\alpha}$. □

Claim 3.2. *For all constants $\delta \geq 2, k \geq 3$, and $c \geq \frac{8}{3}\log\frac{k}{\alpha}$, if $W > c\cdot\tau$, then we have*

$$\boldsymbol{E}\left[a(\boldsymbol{w})\,\middle|\, f(\boldsymbol{w}) \leq \delta W, \mathcal{E}_0\right] \geq \left(1-\frac{\delta-1}{k}\right)\alpha W.$$

Proof. We know that

$$\begin{aligned}\alpha\cdot W \leq \mathbf{E}\left[a(\boldsymbol{w})\,\middle|\,\mathcal{E}_0\right] &= \mathbf{E}\left[a(\boldsymbol{w})\,\middle|\, f(\boldsymbol{w}) \leq \delta W, \mathcal{E}_0\right]\cdot\Pr\left[f(\boldsymbol{w}) \leq \delta W \mid \mathcal{E}_0\right]\\ &\quad+\mathbf{E}\left[a(\boldsymbol{w})\,\middle|\, f(\boldsymbol{w}) > \delta W, \mathcal{E}_0\right]\cdot\Pr\left[f(\boldsymbol{w}) > \delta W \mid \mathcal{E}_0\right]\\ &\leq \mathbf{E}\left[a(\boldsymbol{w})\,\middle|\, f(\boldsymbol{w}) \leq \delta W, \mathcal{E}_0\right]\\ &\quad+\mathbf{E}\left[f(\boldsymbol{w})\,\middle|\, f(\boldsymbol{w}) > \delta W, \mathcal{E}_0\right]\cdot\Pr\left[f(\boldsymbol{w}) > \delta W \mid \mathcal{E}_0\right],\end{aligned} \tag{9}$$

where the first inequality is by Equation (5), the equality is by the Law of Total Expectation, and the second inequality is since a probability is bounded by 1, and $a(\boldsymbol{w}) \leq f(\boldsymbol{w})$.

Next, we show that

$$\begin{aligned}
&\Pr[f(\boldsymbol{w}) > \delta W \mid \mathcal{E}_0] \cdot \mathbf{E}[f(\boldsymbol{w}) \mid f(\boldsymbol{w}) > \delta W, \mathcal{E}_0] \\
&= \Pr[f(\boldsymbol{w}) > \delta W \mid \mathcal{E}_0] \cdot \frac{1}{\Pr[f(\boldsymbol{w}) > \delta W \mid \mathcal{E}_0]} \cdot \int_{\delta W}^{\infty} \Pr[f(\boldsymbol{w}) > z \mid \mathcal{E}_0] dz \\
&= \int_{\delta W}^{\infty} \Pr[f(\boldsymbol{w}) > z \mid \mathcal{E}_0] dz = \int_{(\delta-1)W}^{\infty} \Pr[f(\boldsymbol{w}) > W + z \mid \mathcal{E}_0] dz \\
&= \int_{(\delta-1)W}^{\infty} \Pr\left[\frac{f(\boldsymbol{w})}{\tau} > \frac{1}{\tau}(W+z) \,\middle|\, \mathcal{E}_0\right] dz \overset{(1)}{\leq} \int_{(\delta-1)W}^{\infty} e^{-\frac{3(z/\tau)^2}{6W/\tau + 2z/\tau}} dz \\
&\overset{(2)}{\leq} \int_{(\delta-1)W}^{\infty} e^{-\frac{3(\delta-1)}{4+2\delta}\frac{z}{\tau}} dz = \frac{4+2\delta}{3(\delta-1)} \tau e^{-\frac{3(\delta-1)^2}{4+2\delta}\frac{W}{\tau}} \overset{(3)}{<} \frac{4+2\delta}{3c(\delta-1)} W e^{-\frac{3(\delta-1)^2}{4+2\delta} c} \\
&\overset{(4)}{\leq} (\delta-1)\frac{\alpha}{k} W. \qquad (10)
\end{aligned}$$

Here, (1) follows from applying Theorem 2.1, (2) from noticing that $W \leq z/(\delta-1)$ within the integral limits, (3) from the fact that $W > c \cdot \tau$, (4) since $c \geq \frac{8}{3}\log\frac{k}{\alpha} \geq \frac{4+2\delta}{3(\delta-1)^2}\log\frac{k}{\alpha}$.

Combining Equations (9) and (10), we get

$$\mathbf{E}\left[a(\boldsymbol{w}) \,\middle|\, f(\boldsymbol{w}) \leq \delta W, \mathcal{E}_0\right] \geq \left(1 - \frac{\delta-1}{k}\right)\alpha W,$$

which concludes the proof of the claim. □

Proof of Lemma 3.2. We conservatively assume that if there is at least one element with a weight exceeding τ, then the contribution of the algorithm is 0. By Claim C.5, $\Pr[\mathcal{E}_0] = \gamma$. We then have that:

$$\begin{aligned}
\mathbf{E}\left[\frac{a(\boldsymbol{w})}{f(\boldsymbol{w})}\right] &\geq \mathbf{E}\left[\frac{a(\boldsymbol{w})}{f(\boldsymbol{w})} \,\middle|\, \mathcal{E}_0\right] \cdot \Pr[\mathcal{E}_0] \\
&\geq \mathbf{E}\left[\frac{a(\boldsymbol{w})}{f(\boldsymbol{w})} \,\middle|\, f(\boldsymbol{w}) \leq \delta W, \mathcal{E}_0\right] \cdot \Pr[f(\boldsymbol{w}) \leq \delta W \mid \mathcal{E}_0] \cdot \Pr[\mathcal{E}_0] \\
&\geq \frac{\mathbf{E}[a(\boldsymbol{w}) \mid f(\boldsymbol{w}) \leq \delta W, \mathcal{E}_0]}{\delta W} \cdot (1 - \Pr[f(\boldsymbol{w}) > \delta W \mid \mathcal{E}_0]) \cdot \gamma \\
&= \frac{\gamma}{\delta W} \cdot (1 - \Pr[f(\boldsymbol{w}) > \delta W \mid \mathcal{E}_0]) \cdot \mathbf{E}[a(\boldsymbol{w}) \mid f(\boldsymbol{w}) \leq \delta W, \mathcal{E}_0]. \qquad (11)
\end{aligned}$$

We bound $\Pr[f(\boldsymbol{w}) > \delta W \mid \mathcal{E}_0]$ from above in Claim 3.1 and $\mathbf{E}[a(\boldsymbol{w}) \mid f(\boldsymbol{w}) \leq \delta W, \mathcal{E}_0]$ from below in Claim 3.2.

Combining Equation (11) with Claim 3.1 and Claim 3.2, we get that

$$\mathbf{E}\left[\frac{a(\boldsymbol{w})}{f(\boldsymbol{w})}\right] \geq \frac{\gamma}{\delta}\left(1 - \frac{\alpha}{k}\right)\left(1 - \frac{\delta-1}{k}\right)\alpha \geq \frac{\gamma}{\delta}\frac{k-\delta}{k}\alpha,$$

which concludes the proof. □

To complete the proof of Theorem 3.1, we need to carefully choose the parameters of the previous claims and lemmas (including the threshold of Equation 3, used for the case distinction and beyond) so the multiplicative loss from the two cases balances to a (relatively good) constant.

Proof of Theorem 3.1. If $\alpha = \mathsf{RoE}(\mathcal{F})$, then there exists an algorithm $\mathsf{ALG}_{\mathsf{RoE}}$ that satisfies Equation (5). Algorithm 1 guarantees the minimum between the expected competitive ratios of

ALGORITHM 2: EoR-to-RoE

Data: Ground set E, distributions D_e, feasibility family $\mathcal{F}$, a subroutine $\mathsf{ALG}_{\mathsf{EoR}}$
Assumption: $\mathsf{ALG}_{\mathsf{EoR}}(\boldsymbol{w})$ satisfies that $\mathbf{E}_{\boldsymbol{w}\sim D}\left[a_{\mathsf{ALG}_{\mathsf{EoR}}}(\boldsymbol{w})/f(\boldsymbol{w})\right] = \alpha$
Result: Subset $\mathsf{ALG}(\boldsymbol{w}) \subseteq E$ such that $\mathsf{ALG}(\boldsymbol{w}) \in \mathcal{F}$
Let $A := \mathbf{E}_{w\sim D}[\max_{e\in E} w_e]$;
if $A \geq \alpha \cdot \boldsymbol{E}_{\boldsymbol{w}\sim D}[f(\boldsymbol{w})]/34$ **then**
  Return the first element $e^* \in E$ such that $w_{e^*} \geq \frac{A}{2}$;
  If no such element exists, return $\emptyset$;
else
  Run $\mathsf{ALG}_{\mathsf{EoR}}(\boldsymbol{w})$;
end

Lemma 3.1 and Lemma 3.2. Since the only parameter we do not control is α, as it depends on the feasibility structure $\mathcal{F}$, we choose $\gamma = 1/2, \delta = 2, k = 3, c = \frac{8}{3}\log\frac{3}{\alpha}$, and we get

$$\mathsf{EoR} \geq \min\left\{\frac{\frac{1}{2}\log 2}{\frac{8}{3}\log\frac{3}{\alpha}+1}, \frac{\alpha}{12}\right\} = \frac{\alpha}{12}.$$

This concludes the proof of the theorem. □

4 From EoR to RoE: Instance-based Reduction

In this section, we show an instance-based reduction from EoR to RoE. Unlike Theorem 3.1, our next result shows that the RoE is always at least a constant fraction of the EoR for every pair of (downward-closed) feasibility constraint $\mathcal{F}$ and product distribution D. We will assume the existence of an algorithm $\mathsf{ALG}_{\mathsf{EoR}}(\boldsymbol{w})$ that satisfies $\mathbf{E}_{\boldsymbol{w}\sim D}[a_{\mathsf{ALG}_{\mathsf{EoR}}}(\boldsymbol{w})/f(\boldsymbol{w})] \geq \alpha$. For ease of notation, we denote the value of Algorithm 2 on $\boldsymbol{w}$ by $a(\boldsymbol{w})$.

Theorem 4.1. *For every downward-closed family of feasibility constraints $\mathcal{F}$ and every product distribution D, it holds that*

$$RoE(\mathcal{F}, D) \geq \frac{EoR(\mathcal{F}, D)}{68}. \tag{12}$$

Before proceeding with the proof of the above theorem, we show the following claim, which will prove useful in the remainder. To this end, let $w'_e = w_e \cdot \mathbf{1}_{w_e \leq 2A}$, for every $e \in E$, let D'_e be the distribution of w'_e, and let D' be their product distribution.

Claim 4.1. *It holds that $\boldsymbol{E}_{\boldsymbol{w}\sim D}[\sum_{e\in E} w_e - w'_e] \leq 2\alpha \cdot \frac{\boldsymbol{E}_{w\sim D}[f(w)]}{34}$.*

Proof. Let event $B_e = \{w_{e'} \leq 2A, \forall e' \neq e\}$. We have that

$$\mathbf{E}\left[\sum_{e\in E} w_e - w'_e\right] = \mathbf{E}\left[\sum_{e\in E} w_e \cdot \mathbf{1}_{w_e > 2A}\right] \leq 2 \cdot \mathbf{E}\left[\sum_{e\in E} w_e \cdot \mathbf{1}_{w_e>2A} \cdot \mathbf{1}_{B_e}\right] \leq 2A, \tag{13}$$

where the first inequality follows from noting that (1) B_e is independent of the realization of w_e, and (2) $\Pr[B_e] \geq \frac{1}{2}$, since $\Pr[B_e] \geq \Pr[\max_{e\in E} w_e \leq 2A]$ and by applying Markov's Inequality on random variable $\max_{e\in E} w_e$. The second inequality follows from the fact that we are accounting for selecting the maximum only when there is a unique element whose weight exceeds $2A$ (which is at most the expected maximum). The claim then holds by combining the last inequality and the assumption on A from being in this case. □

Proof of Theorem 4.1. Let $\alpha = \mathsf{EoR}(\mathcal{F}, D)$, and let $A = \mathbf{E}_{w\sim D}[\max_{e\in E} w_e]$ (as per Algorithm 2's pseudocode). Our algorithm has two cases, depending on the value of A.

In the first case, $A \geq \alpha \cdot \frac{\mathbf{E}[f(\boldsymbol{w})]}{34}$: Algorithm 2 sets a threshold of $\frac{A}{2}$ and selects the first element that exceeds it. The algorithm achieves

$$\mathbf{E}\left[a(\boldsymbol{w})\right] \geq \frac{A}{2} \geq \alpha \cdot \frac{\mathbf{E}[f(\boldsymbol{w})]}{68},$$

where the first inequality follows by the Prophet Inequality [53] and the last from the assumption on A in this case.

Otherwise, we have that $A < \alpha \cdot \frac{\mathbf{E}[f(\boldsymbol{w})]}{34}$, in which case Algorithm 2 simply runs the $\mathsf{ALG}_{\mathsf{EoR}}$ subroutine as a blackbox. By Proposition 2.1, the function $\frac{f(\boldsymbol{w})}{2A}$ is self-bounding in the domain $[0, 2A]^{|E|}$. Let $Q = \frac{\mathbf{E}_{\boldsymbol{w}'\sim D'}[f(\boldsymbol{w}')]}{2A}$. It holds that

$$\begin{aligned} Q = \frac{\mathbf{E}_{\boldsymbol{w}'\sim D'}[f(\boldsymbol{w}')]}{2A} &> \frac{34 \cdot \mathbf{E}_{\boldsymbol{w}'\sim D'}[f(\boldsymbol{w}')]}{2\alpha \cdot \mathbf{E}_{\boldsymbol{w}\sim D}[f(\boldsymbol{w})]} \\ &\geq \frac{17 \cdot (\mathbf{E}_{\boldsymbol{w}\sim D}[f(\boldsymbol{w})] - 2A)}{\alpha \cdot \mathbf{E}_{\boldsymbol{w}\sim D}[f(\boldsymbol{w})]} > \frac{17}{\alpha}\left(1 - \frac{2\alpha}{34}\right) \geq \frac{16}{\alpha}, \end{aligned} \tag{14}$$

where the first and third inequalities follow simply by using $A < \alpha \cdot \frac{\mathbf{E}[f(\boldsymbol{w})]}{34}$. The second inequality follows by Claim 4.1, by Lipschitzness of f, Inequality (13), and the value of A. The last inequality is since $\alpha \leq 1$. We, therefore, have that

$$\Pr_{\boldsymbol{w}'\sim D'}\left[\frac{f(\boldsymbol{w}')}{2A} \leq \frac{Q}{2}\right] \leq e^{-Q/8} \overset{(14)}{\leq} e^{-2/\alpha} \leq \frac{\alpha}{2}, \tag{15}$$

where the first inequality follows from Theorem 2.1, and the last inequality is by the identity $e^{-x} \leq 1/x$ for all $x > 0$. Let $V = \mathbf{E}_{\boldsymbol{w}\sim D}[f(\boldsymbol{w})]$, then it holds that

$$\frac{V}{4} \leq \frac{\mathbf{E}_{\boldsymbol{w}\sim D}[f(\boldsymbol{w})] - 2A}{2} \leq \frac{\mathbf{E}_{\boldsymbol{w}'\sim D'}[f(\boldsymbol{w}')]}{2} = Q \cdot A, \tag{16}$$

where the first inequality is by using $A < \alpha \cdot \frac{\mathbf{E}[f(\boldsymbol{w})]}{34}$, and the second inequality is again by Claim 4.1, by Lipschitzness of f, Inequality (13), and the value of A. Thus,

$$\Pr_{\boldsymbol{w}\sim D}\left[f(\boldsymbol{w}) \leq \frac{V}{4}\right] \leq \Pr_{\boldsymbol{w}'\sim D'}[f(\boldsymbol{w}') \leq Q \cdot A] \overset{(15)}{\leq} \frac{\alpha}{2}, \tag{17}$$

where the first inequality holds, since f is monotone, since $\boldsymbol{w}' \leq \boldsymbol{w}$ component-wise, and since, by Equation (16), $\frac{V}{4} \leq Q \cdot A$. We next bound $\mathbf{E}_{\boldsymbol{w}\sim D}[\frac{a(\boldsymbol{w})}{f(\boldsymbol{w})} \mid f(\boldsymbol{w}) \geq \frac{V}{4}]$. By definition of α, it holds that

$$\begin{aligned} \alpha &\leq \mathbf{E}_{\boldsymbol{w}\sim D}\left[\frac{a(\boldsymbol{w})}{f(\boldsymbol{w})}\right] \\ &= \mathbf{E}_{\boldsymbol{w}\sim D}\left[\frac{a(\boldsymbol{w})}{f(\boldsymbol{w})} \,\middle|\, f(\boldsymbol{w}) < \frac{V}{4}\right] \cdot \Pr\left[f(\boldsymbol{w}) < \frac{V}{4}\right] \\ &+ \mathbf{E}_{\boldsymbol{w}\sim D}\left[\frac{a(\boldsymbol{w})}{f(\boldsymbol{w})} \,\middle|\, f(\boldsymbol{w}) \geq \frac{V}{4}\right] \cdot \Pr\left[f(\boldsymbol{w}) \geq \frac{V}{4}\right] \\ &\leq \frac{\alpha}{2} + \mathbf{E}_{\boldsymbol{w}\sim D}\left[\frac{a(\boldsymbol{w})}{f(\boldsymbol{w})} \,\middle|\, f(\boldsymbol{w}) \geq \frac{V}{4}\right]. \end{aligned}$$

By rearranging, we get that

$$\mathbf{E}_{\boldsymbol{w}\sim D}\left[\frac{a(\boldsymbol{w})}{f(\boldsymbol{w})} \,\middle|\, f(\boldsymbol{w}) \geq \frac{V}{4}\right] \geq \frac{\alpha}{2}. \tag{18}$$

This implies that

$$\mathsf{RoE}(\mathcal{F}, D) \geq \frac{\mathbf{E}_{w\sim D}[a(\boldsymbol{w})]}{V} \geq \mathbf{E}_{\boldsymbol{w}\sim D}\left[a(\boldsymbol{w}) \,\middle|\, f(\boldsymbol{w}) \geq \frac{V}{4}\right] \cdot \Pr\left[f(\boldsymbol{w}) \geq \frac{V}{4}\right] \cdot \frac{1}{V}$$
$$\geq \frac{\alpha \cdot V}{8} \cdot \left(1 - \frac{\alpha}{2}\right) \cdot \frac{1}{V} \geq \frac{\alpha}{16},$$

where the third inequality holds by Equation (17) and Equation (18), and the last inequality is since $\alpha \leq 1$. This concludes the proof. □

An immediate corollary of Theorem 4.1 is:

Corollary 4.1. *For every downward-closed family of feasibility constraints $\mathcal{F}$ it holds that*

$$RoE(\mathcal{F}) \geq \frac{EoR(\mathcal{F})}{68}.$$

The above corollary and Theorem 3.1 imply

Corollary 4.2. *For every downward-closed family of feasibility constraints $\mathcal{F}$, it holds that*

$$\frac{RoE(\mathcal{F})}{EoR(\mathcal{F})} \in \Theta(1).$$

5 From RoE to EoR: an Impossibility of Instance-based Reduction

In this section, we show that an instance-based reduction from RoE to EoR is unachievable. This is in contrast to the reduction of Theorem 4.1 from EoR to RoE. In particular, we show that there are a feasibility constraint $\mathcal{F}$ and a product distribution D for which $\mathsf{RoE}(\mathcal{F}, D)$ is constant but $\mathsf{EoR}(\mathcal{F}, D)$ is sub-constant. We show this using the following stronger claim:

Proposition 5.1. *For every feasibility constraint $\mathcal{F}$ there exists a product distribution D such that $EoR(\mathcal{F}, D) \in O(RoE(\mathcal{F}))$ while $RoE(\mathcal{F}, D) \in \Omega(1)$.*

Proof. Let $\mathsf{RoE}(\mathcal{F}) = x$. By definition of $\mathsf{RoE}(\mathcal{F})$ as an infimum over all product distributions D, there exists a product distribution D such that $\mathsf{RoE}(\mathcal{F}, D) \leq \frac{69x}{68}$. We now consider the product distribution D' constructed from leaving all elements' distributions unaltered but modifying the distribution D_{e^*} of an arbitrary element e^* into D'_{e^*} as follows: Consider the random variable w_{e^*} representing the weight of element e^* and let

$$w'_{e^*} := w_{e^*} + \frac{\mathbf{E}\left[\sum_{e\in E} w_e\right]}{x} \cdot X,$$

where $X \sim \text{Bernoulli}(x)$. It is easy to see that $\mathsf{RoE}(\mathcal{F}, D') \geq \frac{1}{2}$, since the algorithm that always selects element e^* obtains $\mathbf{E}[w'_{e^*}] \geq \mathbf{E}[\sum_{e\in E} w_e]$, while the prophet can achieve at most $2\cdot\mathbf{E}[\sum_{e\in E} w_e]$. To conclude, we have that

$$\mathsf{EoR}(\mathcal{F}, D') \leq \Pr[X = 1] \cdot 1 + \Pr[X = 0] \cdot \mathsf{EoR}(\mathcal{F}, D)$$
$$\leq x + \mathsf{EoR}(\mathcal{F}, D) \leq x + 68 \cdot \mathsf{RoE}(\mathcal{F}, D) \leq 70x,$$

where the first inequality above derives from the fact that, if the Bernoulli random variable X is 1, then we upper bound the algorithm's performance by that of the optimum; otherwise, we upper bound the performance by $\mathsf{EoR}(\mathcal{F}, D)$, the third inequality follows from Theorem 4.1, and the last by recalling that $\mathsf{RoE}(\mathcal{F}, D) \leq \frac{69x}{68}$. □

We know by the example presented in Appendix B of Reference [50] that is based on an example from Reference [7] for a different setting that there exists a feasibility constraint with n elements

such that $\mathsf{RoE}(\mathcal{F}) \in O(\frac{\log\log(n)}{\log(n)})$. Combining Proposition 5.1 with this example for large enough n implies that:

Corollary 5.1. *For every $\epsilon > 0$, there exist a feasibility constraint $\mathcal{F}$ and a product distribution D in which $EoR(\mathcal{F}, D) \leq \epsilon$ and $RoE(\mathcal{F}, D) \in \Omega(1)$.*

6 Discussion

In this article, we studied the performance of combinatorial prophet inequalities via the expected ratio (EoR). We focus on its connections to the standard measure of performance in the literature, i.e., the ratio of expectations (RoE). We establish that, for every downward-closed feasibility constraint, the gap between $\mathsf{RoE}(\mathcal{F})$ and $\mathsf{EoR}(\mathcal{F})$ is at most a constant. Moreover, we show that the EoR is an even stronger benchmark in the sense that $\mathsf{RoE}(\mathcal{F}, D)$ is at least a constant of $\mathsf{EoR}(\mathcal{F}, D)$ but not vice versa.

We want to remark that Algorithm 1 and Algorithm 2 are constructive ways to prove Theorem 3.1 and Theorem 4.1. For example, the purpose of Algorithm 1 is to show that for every family of feasibility constraints $\mathcal{F}$, $\mathsf{EoR}(\mathcal{F}) \geq \mathsf{RoE}(\mathcal{F})/12$. For $\alpha = \mathsf{RoE}(\mathcal{F})$, by definition of $\mathsf{RoE}(\mathcal{F})$, there exists an algorithm that satisfies the assumption made in the algorithm, and for the proof of the theorem, this can be used to show an existence of an algorithm with $\mathsf{EoR}(\mathcal{F}) \geq \alpha/12$. Therefore, the assumption simply restates the starting point of the reduction of Theorem 3.1.

In the remainder of this section, we state some remarks and discuss extensions that follow from our proofs and techniques.

Arrival order. We first note that our results hold for any arrival order of the elements, such as random order [27], free order [57], or batch arrival order [29], since neither do our algorithms use any knowledge of the order, nor do our results rely on any restricting assumptions on this.

Single-sample. We can also consider scenarios in which the decision-maker does not have full knowledge of the distributions of the elements' weights. In fact, our reductions can be adjusted (with slightly worse constants) to scenarios in which the decision-maker has only a single sample from each distribution (see Appendix D.1 for a formal discussion).

Extension to XOS functions. Our results extend beyond additive functions over downward-closed feasibility constraints, namely, to extended-XOS functions (see Appendix D.2 for a formal discussion). We generalize the optimum function (Definition 2.1) to be $\mathsf{OPT}(\boldsymbol{w}) := \arg\max_{S\in\mathcal{F}} \max_{i\in[\ell]} \langle b_i, \boldsymbol{w}_S\rangle$, where $\boldsymbol{w}_S \in \mathbb{R}^{|E|}$ is the vector of elements weights in S (and 0 for elements not in S), while each $b_i \in \mathbb{R}^{|E|}$ is a vector of nonnegative coefficients. Similarly to the additive case, we have that $f(\boldsymbol{w}) = \max_{\substack{i\in[\ell]\\ S\in\mathcal{F}}} \langle b_i, \boldsymbol{w}_S\rangle$. We can show that, by setting $a_{ij}^S := b_{ij} \cdot \mathbf{1}_{j\in S}$, $f(\boldsymbol{w})$ can be expressed as $f(\boldsymbol{w}) = \max_{\substack{i\in[\ell]\\ S\in\mathcal{F}}} \left\langle a_i^S, \boldsymbol{w}_S\right\rangle$. Moreover, the functions resulting from projecting all such f's onto $\{0,1\}^{|E|}$ not only are XOS but describe all XOS functions. We can now run Algorithm 1 and perform the case distinction with the modified threshold τ being such that $\Pr\left[\exists j \in [|E|] : w_j > \frac{\tau}{\max_{i\in[\ell]} a_{ij}}\right] = \gamma$. With this at hand, the "catch the superstar" subroutine becomes selecting the first element j with $w_j > \frac{\tau}{\max_{i\in[\ell]} a_{ij}}$, while the "run the combinatorial algorithm" remains unaltered (up to slight modifications, described in Appendix D.2). For Algorithm 2, we redefine $A := \mathbf{E}[\max_{i,j} a_{ij} \cdot w_j]$ and repeat a similar analysis to the one above. All in all, we lose an additional $\max_{i,j} a_{ij}$ factor in the expected ratio of Theorem 3.1 and Theorem 4.1, and we get that:

$$\mathsf{EoR}(\mathcal{F}) \geq \frac{\mathsf{RoE}(\mathcal{F})}{12 \cdot \max_{i,j} a_{ij}}, \quad \text{and} \quad \mathsf{RoE}(\mathcal{F}, D) \geq \frac{\mathsf{EoR}(\mathcal{F}, D)}{68 \cdot \max_{i,j} a_{ij}}.$$

Other measures of performance. In this work, we study the expected ratio as our measure of performance. Another natural performance measure is the expected inverse ratio. We now show that such a measure may be unbounded even for the single-choice case. Let us consider again Example 1.3 and denote by D the product distribution of the first and second box distributions. Fix a (randomized) algorithm that selects the first element with probability p, and let $a_p(\boldsymbol{w})$ be its performance on input $\boldsymbol{w}$. Then, we have

$$\begin{aligned}
\mathsf{EoIR}(\mathcal{F}) &\geq \mathsf{EoIR}(\mathcal{F}, D) \\
&\geq \min_p \mathbf{E}_D \left[\frac{f(\boldsymbol{w})}{a_p(\boldsymbol{w})} \right] \\
&\geq \min_p (1-\varepsilon) \cdot \left(p \cdot \frac{1}{1} + (1-p) \cdot \frac{1}{\varepsilon^2} \right) + \varepsilon \cdot \left(p \cdot \frac{1/\varepsilon^2}{1} + (1-p) \cdot \frac{1/\varepsilon^2}{1/\varepsilon^2} \right) \\
&= \min_p \frac{1-\varepsilon}{\varepsilon^2} + \varepsilon - \left(\frac{1-\varepsilon}{\varepsilon^2} - \frac{1}{\varepsilon} - 1 + 2\varepsilon \right) p \geq \frac{1}{\varepsilon} + 1 - \varepsilon > \frac{1}{\varepsilon},
\end{aligned}$$

where both $\mathsf{EoIR}(\mathcal{F}), \mathsf{EoIR}(\mathcal{F}, D)$ are defined analogously to the other quantities in Equation 2. Hereby, the second inequality holds, since the algorithm selects the second box (given it is realized) with probability at most $1 - p$, while the third inequality follows from setting $p = 1$ to minimize the expression. As ε can be made arbitrarily small, we have an unbounded EoIR. Note that, in Example 1.3, $\frac{1}{\mathsf{RoE}}$ is unbounded only for that specific algorithm, but we know that there exists an algorithm with $\frac{1}{\mathsf{RoE}} \leq 2$, while for EoIR any algorithm gives an unbounded ratio.

In spite of the above simple impossibility result in maximization problems, the same measure of performance could be of use when the decision-maker seeks to minimize a function subject to, e.g., covering constraints with stochastic inputs. As mentioned in Section 1.3, Garg et al. [33] study the relation between RoE and EoR for minimization problems, such as Online Steiner Tree and Traveling Salesman Problem with stochastic inputs. However, the EoIR measure in this setting remains unexplored. It would be interesting to understand whether the reductions provided in Sections 3 and 4 are generalizable to the minimization setting (both for the EoR and the EoIR).

Gap between EoR and RoE.. When RoE and EoR differ by a constant factor, we have shown that it is possible to devise a policy that approximates both metrics. This allows for a "best of both worlds" interpretation of our results, where even if the primary focus is RoE, the policy will also perform well on the EoR metric. Despite the similarity between the benchmarks, it is not at all obvious whether the maximal gap between RoE and EoR for prophet settings and every downward-closed feasibility constraint would be constant (it is an open question whether this gap is constant in other Bayesian settings [33]). In our feasibility-based reduction from RoE to EoR, we lose a constant of $\frac{1}{12}$. Losing a constant is unavoidable already from the single-choice setting, where there is a (tight) gap of $\frac{2}{e}$ (since $\mathsf{EoR} = \frac{1}{e}$ and $\mathsf{RoE} = \frac{1}{2}$). It would be interesting to study whether this is the worst gap possible in all combinatorial settings. In the other direction (i.e., from EoR to RoE), the gap in the reduction is also not tight; it is even possible that the $\mathsf{RoE}(\mathcal{F})$ is at least the $\mathsf{EoR}(\mathcal{F})$ for every feasibility constraint $\mathcal{F}$. However, we know that $\mathsf{RoE}(\mathcal{F}, D)$ can be smaller than $\mathsf{EoR}(\mathcal{F}, D)$ by a factor of 2 by Example 1.2 and can be unboundedly larger than $\mathsf{EoR}(\mathcal{F}, D)$. Additionally, finding the exact value of $\mathsf{EoR}(\mathcal{F})$ for specific downward-closed constraints (e.g., matching, matroid, knapsack) is an interesting open question.

Acknowledgments

Most of this work was done while the author was at TU Munich and Universidad de Chile. The views expressed in this article are the author's and do not necessarily reflect those of the European Central Bank or the Eurosystem.

References

[1] Shipra Agrawal, Jay Sethuraman, and Xingyu Zhang. 2020. On optimal ordering in the optimal stopping problem. In *Proceedings of the 21st ACM Conference on Economics and Computation (EC'20)*. ACM.

[2] Saeed Alaei. 2011. Bayesian combinatorial auctions: Expanding single buyer mechanisms to many buyers. In *Proceedings of the IEEE 52nd Annual Symposium on Foundations of Computer Science*. 512–521.

[3] Saeed Alaei, MohammadTaghi Hajiaghayi, and Vahid Liaghat. 2012. Online prophet-inequality matching with applications to ad allocation. In *Proceedings of the 13th ACM Conference on Electronic Commerce (EC'12)*. ACM Press.

[4] Pieter C. Allaart and José A. Islas. 2016. A sharp lower bound for choosing the maximum of an independent sequence. *J. Appl. Probab.* 53, 4 (2016), 1041–1051.

[5] Nima Anari, Rad Niazadeh, Amin Saberi, and Ali Shameli. 2019. Nearly optimal pricing algorithms for production constrained and laminar Bayesian selection. In *Proceedings of the ACM Conference on Economics and Computation (EC'19)*. ACM, 91–92.

[6] Pablo Daniel Azar, Robert Kleinberg, and S. Matthew Weinberg. 2019. Prior independent mechanisms via prophet inequalities with limited information. *Games Econ. Behav.* 118 (2019), 511–532.

[7] Moshe Babaioff, Nicole Immorlica, and Robert Kleinberg. 2007. Matroids, secretary problems, and online mechanisms. In *Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'07)*. SIAM, 434–443.

[8] Maryam Bahrani, Hedyeh Beyhaghi, Sahil Singla, and S. Matthew Weinberg. 2021. Formal barriers to simple algorithms for the matroid secretary problem. In *Proceedings of the 17th International Conference on Web and Internet Economics (WINE'21)*. Springer, 280–298.

[9] Avrim Blum, Ioannis Caragiannis, Nika Haghtalab, Ariel D. Procaccia, Eviatar B. Procaccia, and Rohit Vaish. 2017. Opting into optimal matchings. In *Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'17)*, Philip N. Klein (Ed.). SIAM, 2351–2363.

[10] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. 2000. A sharp concentration inequality with application. *Rand. Struct. Algor.* 16, 3 (May 2000), 277–292.

[11] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. 2009. On concentration of self-bounding functions. *Electronic Journal of Probability [electronic only]* 14 (2009), 1884–1899. Retrieved from http://eudml.org/doc/225950

[12] Mark Braverman, Mahsa Derakhshan, and Antonio Molina Lovett. 2022. Max-weight online stochastic matching: Improved approximations against the online benchmark. In *Proceedings of the 23rd ACM Conference on Economics and Computation*.

[13] Constantine Caramanis, Paul Dütting, Matthew Faw, Federico Fusco, Philip Lazos, Stefano Leonardi, Orestis Papadigenopoulos, Emmanouil Pountourakis, and Rebecca Reiffenhäuser. 2022. Single-sample prophet inequalities via greedy-ordered selection. In *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'22)*. Society for Industrial and Applied Mathematics, 1298–1325.

[14] Shuchi Chawla, Jason D. Hartline, David L. Malec, and Balasubramanian Sivan. 2010. Multi-parameter mechanism design and sequential posted pricing. In *Proceedings of the 42nd ACM Symposium on Theory of Computing (STOC'10)*, Leonard J. Schulman (Ed.). ACM, 311–320.

[15] Chandra Chekuri and Vasilis Livanos. 2021. On submodular prophet inequalities and correlation gap. In *Proceedings of the 14th International Symposium on Algorithmic Game Theory (SAGT'21) (Lecture Notes in Computer Science,* Vol. 12885*)*. Springer, 410.

[16] José Correa, Andrés Cristi, Boris Epstein, and José A. Soto. 2024. Sample-driven optimal stopping: From the secretary problem to the I.I.D. prophet inequality. *Math. Oper. Res.* 49, 1 (2024), 441–475.

[17] Jose Correa, Patricio Foncea, Ruben Hoeksma, Tim Oosterwijk, and Tjark Vredeveld. 2019. Recent developments in prophet inequalities. *SIGecom Exch.* 17, 1 (May 2019), 61–70.

[18] José Correa, Patricio Foncea, Dana Pizarro, and Victor Verdugo. 2019. From pricing to prophets, and back! *Oper. Res. Lett.* 47, 1 (2019), 25–29.

[19] José R. Correa, Andrés Cristi, Paul Duetting, and Ashkan Norouzi-Fard. 2021. Fairness and bias in online selection. In *Proceedings of the 38th International Conference on Machine Learning (ICML'21) (Proceedings of Machine Learning Research,* Vol. 139*)*. PMLR, 2112–2121.

[20] José R. Correa, Andrés Cristi, Boris Epstein, and José A. Soto. 2020. The two-sided game of Googol and sample-based prophet inequalities. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA'20)*, Shuchi Chawla (Ed.). SIAM, 2066–2081.

[21] José R. Correa, Paul Dütting, Felix A. Fischer, and Kevin Schewior. 2019. Prophet inequalities for I.I.D. random variables from an unknown distribution. In *Proceedings of the ACM Conference on Economics and Computation (EC'19)*. ACM, 3–17.

[22] José R. Correa, Paul Dütting, Felix A. Fischer, Kevin Schewior, and Bruno Ziliotto. 2021. Unknown I.I.D. prophets: Better bounds, streaming algorithms, and a new impossibility (extended abstract). In *Proceedings of the 12th Innovations in Theoretical Computer Science Conference (ITCS'21) (LIPIcs,* Vol. 185*)*. 86:1–86:1.

[23] Paul Dütting, Michal Feldman, Thomas Kesselheim, and Brendan Lucier. 2020. Prophet inequalities made easy: Stochastic optimization by pricing nonstochastic inputs. *SIAM J. Comput.* 49, 3 (2020), 540–582.

[24] Paul Dütting, Evangelia Gergatsouli, Rojin Rezvan, Yifeng Teng, and Alexandros Tsigonias-Dimitriadis. 2023. Prophet secretary against the online optimal. In *Proceedings of the ACM Conference on Economics and Computation (EC'23)*. ACM, 561–581.

[25] Paul Dütting and Robert Kleinberg. 2015. Polymatroid prophet inequalities. In *Proceedings of the 23rd Annual European Symposium on Algorithms (ESA'15)*. 437–449.

[26] Soheil Ehsani, MohammadTaghi Hajiaghayi, Thomas Kesselheim, and Sahil Singla. 2018. Prophet secretary for combinatorial auctions and matroids. In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'18)*. SIAM, 700–714.

[27] Hossein Esfandiari, MohammadTaghi Hajiaghayi, Vahid Liaghat, and Morteza Monemizadeh. 2017. Prophet secretary. *SIAM J. Discret. Math.* 31, 3 (2017), 1685–1701.

[28] Hossein Esfandiari, MohammadTaghi Hajiaghayi, Brendan Lucier, and Michael Mitzenmacher. 2020. Prophets, secretaries, and maximizing the probability of choosing the best. In *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics (AISTATS'20) (Proceedings of Machine Learning Research,* Vol. 108*)*. PMLR, 3717–3727.

[29] Tomer Ezra, Michal Feldman, Nick Gravin, and Zhihao Gavin Tang. 2020. Online stochastic max-weight matching: Prophet inequality for vertex and edge arrival Models. In *Proceedings of the 21st ACM Conference on Economics and Computation (EC'20)*. ACM, 769–787.

[30] Tomer Ezra, Michal Feldman, Nick Gravin, and Zhihao Gavin Tang. 2023. "Who is next in line?" On the significance of knowing the arrival order in Bayesian online settings. In *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'23)*. 3759–3776.

[31] Michal Feldman, Nick Gravin, and Brendan Lucier. 2015. Combinatorial auctions via posted prices. In *Proceedings of the 26th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'15*), Piotr Indyk (Ed.). SIAM, 123–135.

[32] Moran Feldman, Ola Svensson, and Rico Zenklusen. 2021. Online contention resolution schemes with applications to Bayesian selection problems. *SIAM J. Comput.* 50, 2 (2021), 255–300.

[33] Naveen Garg, Anupam Gupta, Stefano Leonardi, and Piotr Sankowski. 2008. Stochastic analyses for online combinatorial optimization problems. In *Proceedings of the 19th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'08)*. 942–951.

[34] John P. Gilbert and Frederick Mosteller. 1966. Recognizing the maximum of a sequence. *J. Amer. Statist. Assoc.* 61, 313 (1966), 35–73.

[35] Nikolai Gravin and Hongao Wang. 2019. Prophet inequality for bipartite matching: Merits of being simple and non adaptive. In *Proceedings of the ACM Conference on Economics and Computation (EC'19)*. ACM, 93–109.

[36] Mohammad Taghi Hajiaghayi, Robert Kleinberg, and Tuomas Sandholm. 2007. Automated online mechanism design and prophet inequalities. In *Proceedings of the 22nd National Conference on Artificial Intelligence (AAAI'07)*. AAAI Press, 58–65.

[37] Jason D. Hartline and Aleck C. Johnsen. 2021. Lower bounds for prior independent algorithms. *CoRR* abs/2107.04754 (2021).

[38] Theodore P. Hill and Robert P. Kertz. 1992. A survey of prophet inequalities in optimal stopping theory. *Contemporary Mathematics* 125 (1992), 191–207.

[39] Martin Hoefer and Bojana Kodric. 2017. Combinatorial secretary problems with ordinal information. In *Proceedings of the 44th International Colloquium on Automata, Languages, and Programming (ICALP'17) (LIPIcs,* Vol. 80*)*. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 133:1–133:14.

[40] Jiashuo Jiang, Will Ma, and Jiawei Zhang. 2022. Tight guarantees for multi-unit prophet inequalities and online stochastic knapsack. In *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'22)*. Society for Industrial and Applied Mathematics, 1221–1246.

[41] Haim Kaplan, David Naori, and Danny Raz. 2020. Competitive analysis with a sample and the secretary problem. In *Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'20)*. SIAM, 2082–2095.

[42] Jon Kleinberg, Robert Kleinberg, and Sigal Oren. 2021. Optimal stopping with behaviorally biased agents: The role of loss aversion and changing reference points. In *Proceedings of the 22nd ACM Conference on Economics and Computation*. ACM.

[43] Robert Kleinberg and S. Matthew Weinberg. 2019. Matroid prophet inequalities and applications to multi-dimensional mechanism design. *Games Econ. Behav.* 113 (2019), 97–115.
[44] Ulrich Krengel and Louis Sucheston. 1977. Semiamarts and finite values. *Bull. Amer. Math. Soc.* 83, 4 (1977), 745–747.
[45] Ulrich Krengel and Louis Sucheston. 1978. On semiamarts, amarts, and processes with finite value. *Advan. Probab. Relat. Topics* 4 (1978), 197—266.
[46] Brendan Lucier. 2017. An economic view of prophet inequalities. *ACM SIGecom Exch.* 16, 1 (2017), 24–47.
[47] Rad Niazadeh, Amin Saberi, and Ali Shameli. 2018. Prophet inequalities vs. approximating optimum online. In *Web and Internet Economics*. Springer International Publishing, 356–374.
[48] Pranav Nuti. 2022. The secretary problem with distributions. In *Integer Programming and Combinatorial Optimization*. Springer International Publishing, 429–439.
[49] Christos Papadimitriou, Tristan Pollner, Amin Saberi, and David Wajc. 2021. Online stochastic max-weight bipartite matching: Beyond prophet inequalities. In *Proceedings of the 22nd ACM Conference on Economics and Computation*. ACM.
[50] Aviad Rubinstein. 2016. Beyond matroids: Secretary problem and prophet inequality with general constraints. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing (STOC'16)*. ACM, 324–332.
[51] Aviad Rubinstein and Sahil Singla. 2017. Combinatorial prophet inequalities. In *Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'17)*. SIAM, 1671–1687.
[52] Aviad Rubinstein, Jack Z. Wang, and S. Matthew Weinberg. 2020. Optimal single-choice prophet inequalities from samples. In *Proceedings of the 11th Innovations in Theoretical Computer Science Conference (ITCS'20) (LIPIcs,* Vol. 151*)*. 60:1–60:10.
[53] Ester Samuel-Cahn. 1984. Comparison of threshold stop rules and maximum for independent nonnegative random variables. *Ann. Probab.* 12, 4 (1984), 1213–1216.
[54] Mark Scharbrodt, Thomas Schickinger, and Angelika Steger. 2006. A new average case analysis for completion time scheduling. *J. ACM* 53, 1 (2006), 121–146.
[55] José A. Soto, Abner Turkieltaub, and Victor Verdugo. 2021. Strong algorithms for the ordinal matroid secretary problem. *Math. Oper. Res.* 46, 2 (2021), 642–673.
[56] Jan Vondrák. 2010. A note on concentration of submodular functions. *CoRR* abs/1005.2791 (2010). arXiv:1005.2791
[57] Qiqi Yan. 2011. Mechanism design via correlation gap. In *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'11)*. 710–719.
[58] Andrew Chi-Chin Yao. 1977. Probabilistic computations: Toward a unified measure of complexity. In *18th Annual Symposium on Foundations of Computer Science*. 222–227.

Appendices

A The Single-choice Case: PSM and EoR Are Equivalent

In this appendix, we will show that for the single-choice prophet inequality, the EoR is equivalent to the PSM. The (tight) bound on the PSM of $1/e$ was already shown in Esfandiari et al. [28], and Proposition A.1 shows that by similar arguments the EoR is also $1/e$. On the one hand, it is immediate to show that the EoR is at least the PSM. On the other hand, we show how to adapt the tight example of the PSM to work also with respect to the EoR. Before proceeding, let us note that in single-choice EoR $= \mathbf{E}[\frac{w_{\text{ALG}}}{\max_{e\in E} w_e}]$, where w_{ALG} is the value selected by the algorithm.

Proposition A.1 (EoR-PSM equivalence). *In the single-choice prophet inequality ($\mathcal{F} = \{S \subseteq E \mid |S| \leq 1\}$), it holds that*

$$EoR(\mathcal{F}) = PSM(\mathcal{F}).$$

Proof. The direction of $\text{EoR}(\mathcal{F}) \geq \text{PSM}(\mathcal{F})$ is trivial, since, if we count only the cases when an algorithm selects the maximum, then $w_{\text{ALG}}/\max_{e\in E} w_e = 1$, which happens with probability of at least $\text{PSM}(\mathcal{F})$. The other direction follows by considering the following instance: for $i \in [|E|]$, let

$$w_1 \sim D_1 = 1,\ w_i \sim D_i = \begin{cases} 0, \text{ w.p. } 1 - \frac{1}{|E|} \\ M^{i-1}, \text{ w.p. } \frac{1}{|E|} \end{cases}, \forall i > 1.$$

By Yao's Minimax Principle [58], we assume without loss of generality that the best algorithm (with respect to EoR) is a deterministic algorithm. Moreover, observe that any optimal algorithm

can be assumed to never select zeroes and be history-independent. The latter is true, because either the algorithm sees 0 (which should never be selected) or a value that is the maximum so far, and thus, all the preceding values are irrelevant for both the algorithm and the optimum. Moreover, for this instance, any deterministic, history-independent algorithm that does not select zeroes can be described as selecting a value if and only if it belongs to a fixed subset of values $S \subseteq \{1, M, \ldots, M^{|E|-1}\}$. We now distinguish between two cases: If $1 \in S$, then $\mathsf{EoR}(\mathsf{ALG}, D, \mathcal{F}) \leq 1 \cdot (1 - \frac{1}{|E|})^{|E|-1} + \frac{1}{M} \approx \frac{1}{e}$, where the approximation is since we can take arbitrarily large M and $|E|$. If $1 \notin S$, then for every $0 < i < j$, such that $M^i \in S$, and $M^j \notin S$, the performance of the algorithm improves by replacing S by $S \setminus \{M^i\} \cup \{M^j\}$. Repeating this argument, we get that S can be described as $\{M^i \mid i \geq k\}$, for some constant $k > 0$, which means it can be described by a single-threshold deterministic algorithm.

Note that if the algorithm picks a non-maximal element, then the $\mathsf{EoR} \leq 1/M$, which tends to 0 for M tending to infinity. This implies that any constant competitive algorithm (in terms of expected ratio) has to select the maximum exactly. Let us now denote by ρ_i the probability that the maximum is selected, had the algorithm chosen threshold $\sigma = M^{i-1}$. We have for all $i > 1$,

$$\rho_i := \Pr\left[\frac{w_{\mathsf{ALG}}}{\max_{e \in E} w_e} = 1 \,\middle|\, \sigma = M^{i-1}\right] = \sum_{j \geq i}\left(\frac{1}{|E|} \cdot \prod_{\substack{k > i \\ k \neq j}}\left(1 - \frac{1}{|E|}\right)\right)$$
$$= \prod_{j \geq i}\left(1 - \frac{1}{|E|}\right) \cdot \sum_{k \geq i} \frac{1/|E|}{1 - 1/|E|},$$

where the first equality is since if the threshold of the algorithm is $\sigma = M^{i-1}$, then ALG selects the maximum if exactly one among $(w_i, \ldots, w_{|E|})$ is not 0. We now simplify the above expression and obtain

$$\rho_1 := \left(1 - \frac{1}{|E|}\right)^{|E|-1} \quad \text{and} \quad \rho_i = \frac{|E| - i + 1}{|E|} \cdot \left(1 - \frac{1}{|E|}\right)^{|E|-i}.$$

We note that ρ_i is a decreasing sequence and is maximized for $i = 1$. It holds that for every $M, |E|$, $\mathsf{EoR}(\mathcal{F}) \leq \frac{1}{M} + \rho_1$, thus, when $M, |E|$ are approaching infinity, we get that

$$\mathsf{EoR}(\mathcal{F}) \leq \lim_{M, |E| \to \infty}\left(\frac{1}{M} + \rho_1\right) = \frac{1}{e},$$

as desired. □

Below, we show an even stronger result for single-choice prophet. For a constant $M > 0$ and a product distribution D of dimension n, we denote by M^D the following product distribution: A vector $\boldsymbol{w}$ is drawn from D, and the realized vector is then $M^{\boldsymbol{w}} := (M^{w_1}, \ldots, M^{w_n})$. The next result states that, for the single-choice feasibility constraint ($\mathcal{F} = \{S \subseteq E \mid |S| \leq 1\}$), and every product distribution D, $\mathsf{PSM}(\mathcal{F}, D)$ equals to the limit of $\mathsf{EoR}(\mathcal{F}, M^D)$, for M that goes to infinity.

Proposition A.2. *In the single-choice prophet inequality ($\mathcal{F} = \{S \subseteq E \mid |S| \leq 1\}$), it holds that*

$$\lim_{M \to \infty} \mathit{EoR}(\mathcal{F}, M^D) = \mathsf{PSM}(\mathcal{F}, D).$$

Proof. Without loss of generality, let us assume that each distribution D_e is defined over the support $\mathbb{R}_{\geq \varepsilon}$. Furthermore, we discretize the support into ε-sized bins and consider the distribution D'_e, resulting from the original distribution when each element in a given bin is associated to the bin's left endpoint. For each D'_e, let us define distribution M^D_e, whose support is

$\{M^w : w \in \text{support}(D'_e)\}$, to be the distribution obtained from drawing $w_e \sim D_e$, and raising M to the power of w_e. We observe that $M^D = \times_{e\in E} M_e^D$, and has support $\bigcup_{e\in E} \{M^w : w \in \text{support}(D'_e)\}$.

Suppose that ALG selects $\max_{e\in E} w_e = M^t$ with probability $\mathsf{PSM}(\mathcal{F}, D)$. Note that for any $s < t$, the ratio $M^s/M^t \le M^{-\varepsilon}$. We take $M \gg 1/\varepsilon$ and obtain

$$\lim_{M\to\infty} \mathsf{EoR}(\mathcal{F}, M^D) \le \lim_{M\to\infty} \mathsf{PSM}(\mathcal{F}, M^D) + (1 - \mathsf{PSM}(\mathcal{F}, M^D)) \cdot M^{-\varepsilon} = \mathsf{PSM}(\mathcal{F}, D).$$

This concludes the proof. □

An alternative proof of Proposition A.1 follows immediately from Proposition A.2. This provides an alternative proof to the fact that EoR and PSM are equivalent in the single-choice setting. Indeed, we have

$$\mathsf{EoR}(\mathcal{F}) = \inf_D \lim_{M\to\infty} \mathsf{EoR}(\mathcal{F}, M^D) = \inf_D \mathsf{PSM}(\mathcal{F}, D) = \mathsf{PSM}(\mathcal{F}).$$

B Implications of "Good" EoR

In this appendix, we formalize the discussion presented in the introduction, regarding the implications of a good EoR.

Claim B.1. *For a (downward-closed) feasibility constraint $\mathcal{F}$ and a product distribution D, if algorithm* ALG *satisfies $\mathsf{EoR}(\mathcal{F}, D, ALG) \ge \alpha$, then*

$$\Pr\left[a(\boldsymbol{w}) \ge \frac{\alpha}{2} \cdot f(\boldsymbol{w})\right] \ge \frac{\alpha}{2}.$$

Proof. $\Pr\left[a(\boldsymbol{w}) \ge \frac{\alpha}{2} \cdot f(\boldsymbol{w})\right] \ge \mathsf{EoR}(\mathcal{F}) - \Pr\left[a(\boldsymbol{w}) < \frac{\alpha}{2} f(\boldsymbol{w})\right] \cdot \frac{\alpha}{2} \ge \frac{\alpha}{2}$. □

Remark B.1. We remark that our RoE to EoR black-box reductions have stronger guarantees than what Claim B.1 states. Indeed, Section 3 shows that with constant probability (as opposed to a probability depending on α, and thus the feasibility constraint), an $O(\alpha)$ approximation of the ex-post optimum can be achieved.

Claim B.2. *Even in single-choice settings, no algorithm can select the maximum (or even any constant approximation to the maximum) with probability larger than $\frac{1}{e}$ (see Appendix A for details).*

Claim B.3. *For every constants $c, \varepsilon > 0$, there exist a (downward-closed) feasibility constraint $\mathcal{F}$ and a product distribution D for which $\mathsf{EoR}(\mathcal{F})$ is constant, and such that the probability that any algorithm achieves better than $\left(\frac{2}{3} + \varepsilon\right)$-approximation of the offline optimum is arbitrarily small. That is,*

$$\Pr\left[a(\boldsymbol{w}) \ge \left(\frac{2}{3} + \varepsilon\right) \cdot f(\boldsymbol{w})\right] \le 5e^{-\frac{\varepsilon^2 n}{12}}.$$

Proof. The proof of this claim derives directly from Example 1.1. Let us first show that the expected ratio is constant (no matter what D is). This is indeed the case, because an algorithm that selects either all the first elements or all the second elements of the pairs with probability $\frac{1}{2}$ each will achieve an $\mathsf{EoR}(\mathcal{F}) \ge \frac{1}{2}$.

We now show the second part of the claim: By Chernoff bound, we have that

$$\Pr\left[f(\boldsymbol{w}) < \frac{3-\varepsilon}{2} \cdot n\right] \le e^{-\frac{\varepsilon^2 n}{12}},$$

since all boxes realizations are independent Bernoulli random variables and $\mathbf{E}[f(\boldsymbol{w})] = \frac{3}{2}n$. Let us now consider any algorithm and define the following random variables: A is the number of times the second box in each pair realized in a 2, while B_i indicates whether or not the algorithm has

selected the larger of the two boxes in the ith pair, and we denote $B := \sum_{i\in[n]} B_i$. It is easy to verify that A is binomial and also B is, since B_i's are independent Bernoulli random variables, even though A and B are dependent. We also observe that, by law of total probability,

$$\begin{aligned}\Pr\left[a(\boldsymbol{w}) > \frac{2+\varepsilon}{2}\cdot n\right] &= \underbrace{\Pr\left[a(\boldsymbol{w}) > \frac{2+\varepsilon}{2}\cdot n \,\middle|\, f(\boldsymbol{w}) < \frac{3-\varepsilon}{2}\cdot n\right]}_{\le 1}\cdot \Pr\left[f(\boldsymbol{w}) < \frac{3-\varepsilon}{2}\cdot n\right] \\ &\quad + \Pr\left[a(\boldsymbol{w}) > \frac{2+\varepsilon}{2}\cdot n \,\middle|\, f(\boldsymbol{w}) \ge \frac{3-\varepsilon}{2}\cdot n\right]\cdot \underbrace{\Pr\left[f(\boldsymbol{w}) \ge \frac{3-\varepsilon}{2}\cdot n\right]}_{\le 1} \\ &\le e^{-\frac{\varepsilon^2 n}{12}} + \Pr\left[A+B > \frac{2+\varepsilon}{2}\cdot n \,\middle|\, A \ge \frac{1-\varepsilon}{2}\cdot n\right].\end{aligned}$$

Hereby, the last inequality derives from the fact that $f(\boldsymbol{w}) = n + A$, while $a(\boldsymbol{w}) = n + A - (n - B) = A + B$. To bound the last term, we can write

$$\begin{aligned}\Pr\left[A+B > \frac{2+\varepsilon}{2}\cdot n \,\middle|\, A \ge \frac{1-\varepsilon}{2}\cdot n\right] &\le \Pr\left[A+B > \frac{2+\varepsilon}{2}\cdot n\right] + \Pr\left[A < \frac{1-\varepsilon}{2}\cdot n\right] \\ &\le \Pr\left[B > \frac{2+\varepsilon}{4}\cdot n\right] + \Pr\left[A > \frac{2+\varepsilon}{4}\cdot n\right] + \Pr\left[A < \frac{1-\varepsilon}{2}\cdot n\right] \\ &\le 2e^{-\frac{\varepsilon^2 n}{16+4\varepsilon}} + e^{-\frac{\varepsilon^2 n}{12}} \\ &\le 3e^{-\frac{\varepsilon^2 n}{12}},\end{aligned}$$

where the second step follows from the union bound. All in all, we can bound the probability that any algorithm obtains a large overall value by $\Pr\left[a(\boldsymbol{w}) > \frac{2+\varepsilon}{2}\cdot n\right] \le 4e^{-\frac{\varepsilon^2 n}{12}}$. Thus,

$$\begin{aligned}\Pr\left[a(\boldsymbol{w}) < \left(\frac{2}{3}+\varepsilon\right)\cdot f(\boldsymbol{w})\right] &\ge 1 - \Pr\left[f(\boldsymbol{w}) < \frac{3-\varepsilon}{2}\cdot n \vee a(\boldsymbol{w}) > \frac{2+\varepsilon}{2}\cdot n\right] \\ &\ge 1 - e^{-\frac{\varepsilon^2 n}{12}} - 4e^{-\frac{\varepsilon^2 n}{12}} \\ &\ge 1 - 5e^{-\frac{\varepsilon^2 n}{12}},\end{aligned}$$

where the second step follows again from the union bound. □

Claim B.4. *There exists a family of feasibility constraints $\mathcal{F}$ such that both RoE($\mathcal{F}$) and EoR($\mathcal{F}$) are constants, but PSM($\mathcal{F}$) $\le 2^{-\frac{|E|}{2}}$, i.e., the probability of selecting the maximum is exponentially small in the cardinality of the ground set.*

Proof. Let us take Example 1.1. For the first part, as seen in Claim B.3, the algorithm that selects either all the first elements or all the second elements of the pairs guarantees an RoE and an EoR of at least 1/2. Moreover, by Jensen's Inequality, we have

$$\underbrace{\mathbf{E}\left[\frac{a(\boldsymbol{w})}{f(\boldsymbol{w})}\right]}_{\le \mathsf{EoR}(\mathcal{F},D)} \ge \overbrace{\frac{\mathbf{E}\left[a(\boldsymbol{w})\right]}{\mathbf{E}\left[f(\boldsymbol{w})\right]}}^{\le \mathsf{RoE}(\mathcal{F},D)} = \frac{n}{\mathbf{E}\left[\sum_{i\in[n]}\max\{w_{1,i}, w_{2,i}\}\right]} = \frac{n}{\frac{3}{2}n} = \frac{2}{3}.$$

For the second part of the claim, let us observe that the probability of selecting the maximum is the probability that, for each pair of boxes i, any algorithm ALG chooses the larger realized value

of the two boxes, $w_{1,i}, w_{2,i}$. Denoting by $w_{\mathsf{ALG},i}$ the weight of the box chosen by the algorithm for pair i, we have

$$\mathsf{PSM}(\mathcal{F}) = \Pr\left[\forall i \in [n], w_{\mathsf{ALG},i} = \max\{w_{1,i}, w_{2,i}\}\right] \leq 2^{-n},$$

where the last equality follows from the following two facts: First, no online algorithm can choose the maximum of each pair with probability greater than $\frac{1}{2}$ by construction. Moreover, all realizations are independent, and thus we get the claimed upper bound since $n = \frac{|E|}{2}$. □

C Auxiliary Notions from Section 3

In Section 3, we have used the following definitions and claims, as well as the properties of self-bounding functions (see Definition 2.3).

Claim C.1 (Jensen's Inequality). *Let u be a random variable and g be any function, then*

$$g(\boldsymbol{E}[u]) \leq \boldsymbol{E}[g(u)], \text{ if } g \text{ is convex.}$$
$$g(\boldsymbol{E}[u]) \geq \boldsymbol{E}[g(u)], \text{ if } g \text{ is concave.}$$

Claim C.2. *Let $a > 0$ be a constant, and $u, v > 0$ two independent random variables, then we have*

$$\frac{a}{a + \boldsymbol{E}[u]} \leq \boldsymbol{E}\left[\frac{a}{a+u}\right],$$
$$\frac{\boldsymbol{E}[u]}{a + \boldsymbol{E}[v]} \leq \boldsymbol{E}\left[\frac{u}{a+v}\right].$$

Proof. The first inequality is a direct consequence of Jensen's Inequality (Claim C.1) with function $g(X) = a/(a+X)$, which is convex. The second inequality above holds, because by independence of x, y, we have

$$\mathbf{E}\left[\frac{X}{a+Y}\right] = \mathbf{E}[X] \cdot \mathbf{E}\left[\frac{1}{a+Y}\right] \geq \frac{\mathbf{E}[X]}{a + \mathbf{E}[Y]},$$

where the last step follows again from Jensen's Inequality. □

Claim C.3. *For any collection of $p_1, \ldots, p_n \in (0,1)$, let $p := \left(\prod_{i\in[n]} p_i\right)^{1/n}$. We have that*

$$n \cdot \frac{1-p}{p} \leq \sum_{i\in[n]} \frac{1-p_i}{p_i}.$$

Proof. Let us consider the base case for $p_1, p_2 \in (0,1)$:

$$\frac{2(1-\sqrt{p_1 p_2})}{\sqrt{p_1 p_2}} \leq \frac{1-p_1}{p_1} + \frac{1-p_2}{p_2}.$$

We now obtain that

$$\begin{aligned}
\frac{n(1-p)}{p} &= n \cdot \frac{1 - e^{\log\left(\prod_{i\in[n]} p_i\right)^{1/n}}}{e^{\log\left(\prod_{i\in[n]} p_i\right)^{1/n}}} = n \cdot \frac{1 - e^{\frac{1}{n}\sum_{i\in[n]} \log p_i}}{e^{\frac{1}{n}\sum_{i\in[n]} \log p_i}} \\
&\leq n \cdot \frac{1 - 1/n \sum_{i\in[n]} p_i}{1/n \sum_{i\in[n]} p_i} \\
&\leq \sum_{i\in[n]} \frac{1-p_i}{p_i},
\end{aligned}$$

where the first inequality is a consequence of the AM-GM inequality. □

Claim C.4. *For every realization of $\boldsymbol{w}$, it holds that $f(\boldsymbol{w}) \leq f(\overline{\boldsymbol{w}}) + \sum_{e\in E} w_e \cdot \mathbf{1}[w_e > \tau]$.*

Proof. Let $S = \{e \mid w_e \le \tau\}$. We have that

$$\begin{aligned} f(\boldsymbol{w}) &= \sum_{e \in \mathrm{OPT}(\boldsymbol{w}) \cap S} w_e + \sum_{e \in \mathrm{OPT}(\boldsymbol{w}) \setminus S} w_e \\ &= \sum_{e \in \mathrm{OPT}(\boldsymbol{w}) \cap S} \overline{w}_e + \sum_{e \in \mathrm{OPT}(\boldsymbol{w}) \setminus S} w_e \\ &\le f(\overline{\boldsymbol{w}}) + \sum_{e \in E \setminus S} w_e = f(\overline{\boldsymbol{w}}) + \sum_{e \in E} w_e \cdot \mathbf{1}\left[w_e > \tau\right], \end{aligned}$$

where the second equality is since $\overline{w}_e = w_e$ for every element $e \in S$, and the inequality is since $\mathrm{OPT}(\boldsymbol{w}) \setminus S$ is a feasible set, and $\mathrm{OPT}(\boldsymbol{w}) \setminus S \subseteq E \setminus S$. □

Claim C.5 (Adapted from Theorem 1 in Reference [28]). *For any set of distributions $\{D_e\}_{e \in E}$,*

$$\Pr\left[\mathcal{E}_0\right] = \gamma,$$
$$\Pr\left[\mathcal{E}_1\right] \ge \gamma \log\left(\frac{1}{\gamma}\right).$$

Proof. By the way we have defined τ, we have $\Pr\left[\tau > \max_{e \in E} w_e\right] = \gamma$. This is equivalent to saying

$$\Pr\left[\forall\, e \in E :\ w_e \le \tau\right] = \gamma.$$

We immediately notice that the probability of the complement of the above is the same as the probability that at least one of w_e is with a value of at least τ, i.e.,

$$\Pr\left[\tau < \max_{e \in E} w_e\right] = \Pr\left[\exists\, e \in E :\ w_e > \tau\right].$$

We also know that the RHS can be written as follows by union bound of disjoint events: Letting $p_e := \Pr\left[w_e < \tau\right]$, it holds that

$$\begin{aligned} \Pr\left[\exists\, e \in E :\ w_e > \tau\right] &\ge \Pr\left[\exists!\, e \in E :\ w_e > \tau\right] \\ &= \sum_{e \in E} \left((1 - p_e) \cdot \prod_{e' \ne e} p_{e'} \right) \\ &= \prod_{e \in E} p_e \cdot \sum_{e \in E} \frac{1 - p_e}{p_e}. \end{aligned}$$

Hereby, the first equality follows from the fact that events of the form "$\exists!\, e \in E :\ w_e > T$" are disjoint, and the union bound holds with equality. Let us note that the threshold is larger than the maximum over all realizations if and only if all w_e's realize in a value less than the threshold τ, which happens by definition with probability γ, which means that $\prod_{e \in E} p_e = \gamma$. By Claim C.3 (below), we let $p := \left(\prod_{e \in E} p_e\right)^{1/|E|}$ and get that

$$\begin{aligned} \sum_{e \in E} \frac{1 - p_e}{p_e} &\ge |E| \cdot \frac{(1 - p)}{p} \\ &\ge |E| \cdot \frac{1 - \gamma^{1/|E|}}{\gamma^{1/|E|}} \ge \log\left(\frac{1}{\gamma}\right). \end{aligned}$$

The second inequality follows from noticing again that $p = \gamma^{1/|E|}$, while the third by observing that the LHS of the same inequality is non-increasing in $|E|$ and the smallest value is attained for

$|E| \to \infty$. All in all, we get that

$$\Pr\left[\exists!\, e \in E:\ w_e > \tau\right] \geq \gamma \cdot \log\left(\frac{1}{\gamma}\right),$$

which concludes the proof of the claim. □

D Extensions of Our Results

D.1 Extension to Single-sample

In this appendix, we discuss an extension of our results to the case of single-sample settings. In our results and definitions, we assumed that the algorithm has *full knowledge* of the distributions from where the elements' weights are drawn. This is used to distinguish between running two different subroutines: The first selects a single element, while the second runs the combinatorial algorithm (see also Algorithms 1 and 2). However, as mentioned in Section 1.3, there has been a significant effort towards designing prophet inequalities with provable guarantees when the decision-maker has only access to some samples from each distribution. We explain now how our result can be extended to single sample settings. To do so, we first define equivalent definitions to those of Section 2.

To measure our performance, given a downward-closed family $\mathcal{F}$, an algorithm ALG, and a product distribution D, we define

$$\text{S-EoR}(\mathcal{F}, \text{ALG}, D) := \mathbf{E}\left[\frac{\boldsymbol{w}(\text{ALG}(\boldsymbol{w}))}{f(\boldsymbol{w})}\right],$$

where the expectation runs over the stochastic generation of the input as well as the (possible) randomness of the algorithm. Similarly, we define

$$\text{S-EoR}(\mathcal{F}, \text{ALG}) := \inf_{D} \text{S-EoR}(\mathcal{F}, \text{ALG}, D), \tag{19}$$

and

$$\text{S-EoR}(\mathcal{F}) := \sup_{\text{ALG}} \text{S-EoR}(\mathcal{F}, \text{ALG}). \tag{20}$$

We note that, since we consider sample-based algorithms, the algorithm is oblivious to the distributions but may depend on the sampled values and the online values observed so far.

We will compare our results to the standard objective of maximizing the ratio of expectations between the algorithm and the offline optimum. Accordingly, we denote

$$\text{S-RoE}(\mathcal{F}, \text{ALG}, D) := \frac{\mathbf{E}\left[\boldsymbol{w}(\text{ALG}(\boldsymbol{w}))\right]}{\mathbf{E}\left[f(\boldsymbol{w})\right]}.$$

Analogously to Equations (19), and (20), we define S-RoE($\mathcal{F}$, ALG) and S-RoE($\mathcal{F}$).

Corollary D.1 (Single-sample simulation). *For every downward-closed family of feasibility constraints $\mathcal{F}$, it holds that*

$$\text{S-EoR}(\mathcal{F}) \geq \frac{\text{S-RoE}(\mathcal{F})}{144}. \tag{21}$$

Proof. Let us first observe that we cannot compute the expected optimum of an instance and perform the case distinction, but we can still flip a fair coin and run either of the two following subroutines: With probability 1/2 (the coin lands heads), we select the first element that exceeds $\hat{\tau} := \max_{e \in E} s_e$, and with probability 1/2 (the coin lands tails), we run the S-ALG$_{\text{RoE}}$ subroutine, an algorithm that achieves S-RoE $= \alpha$ only using samples and which exists by the assumption of the claim. Hereby, we denote by $S = \{s_e \mid e \in E\}$ the sample set and by $O = \{o_e \mid e \in E\}$ the online set. Moreover, let $R = S \cup O$ and $r_1 > r_2 > \cdots > r_{2|E|}$ denote the weights in R sorted in decreasing

order. Although the values W, τ (see analysis of Theorem 3.1) are not known to the algorithm, the analysis can still be partitioned as in the analysis of Theorem 3.1.

If $W \leq c \cdot \tau$ and the coin lands heads, then our algorithm selects the first online value $o_e \geq \hat{\tau}$. Let us only count cases where the following conditions occur jointly: (1) exactly one element in S and exactly one element in O exceed τ, (2) the element with the largest weight is in the online set. It is not difficult to see that this happens with probability at least $\frac{1}{2} \cdot \gamma^2 \log^2(1/\gamma)$. In this case, the approximation, following the analysis of Lemma 3.1, is $1/(c+1)$. From this first case, we obtain

$$\text{S-EoR}(\mathcal{F}) \geq \frac{\gamma^2 \log^2(1/\gamma)}{4(c+1)}.$$

If $W > c \cdot \tau$ and the coin lands tails, then we run algorithm $\text{S-ALG}_{\text{RoE}}$ that, given a single sample from each distribution in the input, satisfies $\mathbf{E}\left[\text{S-ALG}_{\text{RoE}}(\boldsymbol{w}) \mid \mathcal{E}_0\right] \geq \alpha \cdot \mathbf{E}\left[f(\boldsymbol{w}) \mid \mathcal{E}_0\right]$. As before, we only count cases where no sample and no online element exceeds threshold τ, which happens with probability γ^2. Following the analysis of Lemma 3.2, from this case, we have that

$$\text{S-EoR}(\mathcal{F}) \geq \frac{\gamma^2}{2\delta} \frac{k-\delta}{k} \alpha,$$

which means that we lose a multiplicative factor of $\gamma/2$ compared to the full information case. All in all, the described algorithm guarantees

$$\text{S-EoR}(\mathcal{F}) \geq \min\left\{\frac{\log^2 2}{16 \cdot (\frac{8}{3}\log\frac{3}{\alpha} + 1)}, \frac{\alpha}{72}\right\} \geq \frac{\alpha}{144},$$

where we have chosen the parameters as in Theorem 3.1. This concludes the proof. □

Note that, for the other direction, running each of the subroutines used in Algorithm 2 with probability 1/2, we get $\text{S-RoE}(\mathcal{F}, D) \geq \text{S-EoR}(\mathcal{F}, D)/136$.

D.2 Extension to XOS Functions

In this appendix, we describe in detail the extension of results contained in Sections 3 and 4 to a setting where the decision-maker's objective function is XOS, rather than just additive.

Definition D.1. A function $g : 2^{|E|} \to \mathbb{R}_{\geq 0}$ is XOS if there exist ℓ vectors $b_1, \ldots, b_\ell \in \mathbb{R}_{\geq 0}^{|E|}$, such that $g(S) = \max_{i \in [\ell]} \sum_{j \in S} (b_i)_j$.

We extend Definition D.1 to vectors of weights in the following way:

Definition D.2. A function $g : \mathbb{R}_{\geq 0}^{|E|} \to \mathbb{R}_{\geq 0}$ is extended-XOS if there exist ℓ vectors $b_1, \ldots, b_\ell \in \mathbb{R}_{\geq 0}^{|E|}$, such that $g(\boldsymbol{w}) = \max_{i \in [\ell]} \langle b_i, \boldsymbol{w} \rangle$.

In Claim D.1, we show that the projection of an extended-XOS function to $\{0,1\}^{|E|}$ is an XOS function and that the projections of all extended-XOS functions to $\{0,1\}^{|E|}$ are all XOS functions.

We next extend Definition 2.1 to extended-XOS functions:

Definition D.3. Let $\boldsymbol{w} \in \mathbb{R}_{\geq 0}^{|E|}$ be a nonnegative weight vector. We define $\text{OPT} : \mathbb{R}_{\geq 0}^{|E|} \to \mathcal{F}$ to be the function mapping a vector of weights to a maximum-weight set in family $\mathcal{F}$. Namely,

$$\text{OPT}(\boldsymbol{w}) = \arg\max_{S \in \mathcal{F}} \max_{i \in [\ell]} \langle b_i, \boldsymbol{w}_S \rangle,$$

where $\boldsymbol{w}_S \in \mathbb{R}^{|E|}$ is the vector of elements weights in S (and 0 for elements not in S), while $b_i \in \mathbb{R}^{|E|}$ is a vector of nonnegative coefficients. Moreover, we let $f_{\mathcal{F}}(\boldsymbol{w}) = \max_{S \in \mathcal{F}} \max_{i \in [\ell]} \langle b_i, \boldsymbol{w}_S \rangle$.

For simplicity, when clear from context, we denote $f_{\mathcal{F}}$ by f.

Claim D.1. *$f(\boldsymbol{w})$ can be expressed as $f(\boldsymbol{w}) = \max_{\substack{i\in[\ell]\\ S\in\mathcal{F}}} \left\langle a_i^S, \boldsymbol{w}_S \right\rangle$ and is extended-XOS. Moreover, the function resulting from projecting f onto $\{0,1\}^{|E|}$ (i.e., $f : \{0,1\}^{|E|} \to \mathbb{R}$ is now a set function) is XOS and describes all XOS functions.*

Proof. For the first part of the claim, consider a set T. Then,

$$\begin{aligned} f(T) &= \max_{S\in\mathcal{F}} \max_{i\in[\ell]} \sum_{j\in S\cap T} b_{ij} \\ &= \max_{S\in\mathcal{F}} \max_{i\in[\ell]} \sum_{j\in T} a_{ij}^S, \end{aligned} \tag{22}$$

where the second equality follows by setting $a_{ij}^S := b_{ij} \cdot \mathbf{1}_{j\in S}$. Hence, $f(\boldsymbol{w}) = \max_{\substack{i\in[\ell]\\ S\in\mathcal{F}}} \left\langle a_i^S, \boldsymbol{w}_S \right\rangle$ and it is extended-XOS with coefficient vectors a_i^S, as per Definition D.3.

The second part of the claim follows immediately by Equation (22), since this is maximum over additive functions.

The third part of the claim follows, since for every XOS function f, which is defined with additive functions $a_1, \ldots, a_\ell$, the projection of the extended-XOS function defined by the same additive functions is f. □

We have the following properties for f:

Proposition D.1 (Properties of XOS $f_{\mathcal{F}}$). *For every downward-closed family of sets $\mathcal{F}$, the function f $(= f_{\mathcal{F}})$ satisfies the following properties:*

(1) f is $(\max_{i,j} a_{ij})$-Lipschitz.
(2) f is monotone, i.e., if $\boldsymbol{u} \geq \boldsymbol{v}$ point-wise, then $f(\boldsymbol{u}) \geq f(\boldsymbol{v})$.
(3) For every $\tau > 0$, the function $\frac{f}{\tau \cdot \max_{i,j} a_{ij}}$ restricted to the domain $[0,\tau]^{|E|}$ is self-bounding.

Proof. Let $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}_{\geq 0}^{|E|}$ be nonnegative weight vectors and $\mathsf{OPT}(\boldsymbol{u}), \mathsf{OPT}(\boldsymbol{v})$ be, respectively, their optimal feasible subset.

(1) **Lipschitz.** We assume, without loss of generality, that $f(\boldsymbol{u}) \geq f(\boldsymbol{v})$. Then, letting $i(\boldsymbol{u}), i(\boldsymbol{v})$ be, respectively, the indices of the coefficients maximizing $f(\boldsymbol{v}), f(\boldsymbol{u})$, we get

$$\begin{aligned} \mid f(\boldsymbol{u}) - f(\boldsymbol{v}) \mid &= f(\boldsymbol{u}) - f(\boldsymbol{v}) \\ &= \max_{i\in[\ell]} \sum_{j\in\mathsf{OPT}(\boldsymbol{u})} a_{ij}u_j - \max_{i\in[\ell]} \sum_{j\in\mathsf{OPT}(\boldsymbol{v})} a_{ij}v_j \\ &= \sum_{j\in\mathsf{OPT}(\boldsymbol{u})} a_{i(\boldsymbol{u})j}u_j - \sum_{j\in\mathsf{OPT}(\boldsymbol{v})} a_{i(\boldsymbol{v})j}v_j \\ &\leq \sum_{j\in\mathsf{OPT}(\boldsymbol{u})} a_{i(\boldsymbol{u})j}u_j - a_{i(\boldsymbol{v})j}v_j \\ &\leq \max_{i,j} a_{ij} \cdot \sum_{j\in\mathsf{OPT}(\boldsymbol{u})} \mid u_j - v_j \mid \\ &\leq \max_{i,j} a_{ij} \cdot \sum_{j\in E} \left|u_j - v_j\right| = \max_{i,j} a_{ij} \cdot \|\boldsymbol{u} - \boldsymbol{v}\|_1, \end{aligned}$$

(2) **Monotonicity.** Let $\boldsymbol{u}, \boldsymbol{v}$ be two vectors such that $\boldsymbol{v} \leq \boldsymbol{u}$ component-wise. Then, we know that $\langle a_i, \boldsymbol{v}\rangle \leq \langle a_i, \boldsymbol{u}\rangle$ for all $i \in [\ell]$, and thus letting $i(\boldsymbol{u}), i(\boldsymbol{v})$ be, respectively, the indices of the coefficients maximizing $f(\boldsymbol{v}), f(\boldsymbol{u})$, we get

$$f(\boldsymbol{v}) = \left\langle a_{i(\boldsymbol{v})}, \boldsymbol{v}\right\rangle \leq \left\langle a_{i(\boldsymbol{v})}, \boldsymbol{u}\right\rangle \leq \left\langle a_{i(\boldsymbol{u})}, \boldsymbol{u}\right\rangle = f(\boldsymbol{u}).$$

(3) **Self-boundedness.** For every element $e \in E$, we define f_e to be the function f where the weight of element e is 0; i.e., $f_e(\boldsymbol{w}^{(e)}) := f(w_1, \ldots, w_{e-1}, 0, w_{e+1}, \ldots, w_{|E|})$. Thus, $\frac{f}{\tau \cdot \max_{i,j} a_{ij}}$ in the range of $[0, \tau]^{|E|}$ is self-bounding, since:
— The first condition of Definition 2.3 that $0 \leq \frac{1}{\tau \cdot \max_{i,j} a_{ij}} f(\boldsymbol{w}) - \frac{1}{\tau \cdot \max_{i,j} a_{ij}} f_e(\boldsymbol{w}^{(e)}) \leq 1$ is satisfied by monotonicity, the $\max_{i,j} a_{ij}$-Lipschitzness, and the fact that w_e is restricted to the range of $[0, \tau]$.
— We multiply the second condition of Definition 2.3 by factor $\tau \cdot \max_{i,j} a_{ij}$ and get

$$\begin{aligned}
\sum_{e \in E} \left(f(\boldsymbol{w}) - f_e(\boldsymbol{w}^{(e)}) \right) &= \sum_{e \in \mathsf{OPT}(\boldsymbol{w})} \underbrace{f(\boldsymbol{w})}_{=\langle a_{i^*}, \boldsymbol{w} \rangle} - \underbrace{f_e(\boldsymbol{w}^{(e)})}_{\geq \langle a_{i^*}, \boldsymbol{w}^{(e)} \rangle} \\
&\leq \sum_{e \in \mathsf{OPT}(\boldsymbol{w})} \left\langle a_{i^*}, \boldsymbol{w} - \boldsymbol{w}^{(e)} \right\rangle \\
&= \sum_{e \in \mathsf{OPT}(\boldsymbol{w})} a_{i^* e} \cdot w_e = f(\boldsymbol{w}),
\end{aligned}$$

where the first equality derives from the fact that if $e \notin \mathsf{OPT}(\boldsymbol{w})$, then $f(\boldsymbol{w}) - f_e(\boldsymbol{w}^{(e)}) = 0$, and the first inequality derives by setting a_{i^*} to be the maximizing coefficient for f.

This concludes the proof. □

Remark D.1. Before proceeding, we note that Proposition D.1 generalizes similar claims in previous literature. Namely, Vondrák [56] proves self-boundedness for set functions that are XOS or submodular. The key difference between our claim and those in Reference [56] is that our function f is not a combinatorial function, in that f receives a vector of arbitrary weights instead of a set. Therefore, Lemma 2.2 in Reference [56] follows by projecting Proposition D.1 onto the hypercube. Similarly, Proposition D.1 also generalizes Lemma 2.4 in Blum et al. [9], where they prove the self-bounding property for the maximum matching cardinality function. Proposition D.1 is again a generalization in the following senses: First, the constraint is not necessarily a matching but can be a general packing one; second, f can be an arbitrary extended-XOS function and not just the cardinality function.

Below, we illustrate generalizations of feasibility-based and instance-based reductions under extended XOS functions.

Feasibility-based reduction with extended XOS functions. We redefine threshold τ as follows for reasons that will be explained later:

$$\tau : \Pr\left[\exists j \in E : w_j > \frac{\tau}{\max_{i \in [\ell]} a_{ij}} \right] = \gamma.$$

Given the above, let us run Algorithm 1, performing the case distinction with the redefined threshold τ. In the "catch the superstar" case (Lemma 3.1), the algorithm selects the first element j for which $w_j > \frac{\tau}{\max_{i \in [\ell]} a_{ij}}$. Hence, with probability $\gamma \cdot \log(1/\gamma)$, we catch a unique element with value $w_j \cdot \max_{i,j} a_{ij} > \tau$. Since, in this case $W \leq c \cdot \tau$, we get a $\frac{\gamma \cdot \log(1/\gamma)}{c+1}$ expected ratio.

For the "run the combinatorial algorithm" case (Lemma 3.2), we have that, since $\frac{f}{\tau \cdot \max_{i,j} a_{ij}}$ is self-bounding in $[0, \tau]$ (Proposition D.1), we observe that to maintain the same concentration, we need $c \geq \frac{4+2\delta}{3(\delta-1)^2} \log \frac{k}{\alpha} \cdot \max_{i,j} a_{ij}$. Thus, for the extended-XOS function f, we will achieve a similar

performance to Theorem 3.1 but losing an additional factor of $\max_{i,j} a_{ij}$: namely,

$$\mathsf{EoR}(\mathcal{F}) \geq \frac{\mathsf{RoE}(\mathcal{F})}{12 \cdot \max_{i,j} a_{ij}}.$$

Instance-based reduction with extended XOS functions. An almost identical analysis to the above follows in the case of Algorithm 2. Namely, let us redefine A to be

$$A := \mathbf{E}\left[\max_{i,j} a_{ij} \cdot w_j\right].$$

With this at hand, Theorem 4.1 now corresponds to the following statement, again with $f_{\mathcal{F}}$ being an extended-XOS function:

$$\mathsf{RoE}(\mathcal{F}, D) \geq \frac{\mathsf{EoR}(\mathcal{F}, D)}{68 \cdot \max_{i,j} a_{ij}}.$$

E Alternatives to the EoR

In this appendix, we delve into the discussion regarding alternative measures to the EoR, expanding upon what we introduced in Section 1. In particular, the following extension of PSM, named for simplicity PSM_p, is studied in References [8, 39, 55]. The algorithm's goal is to select each element in the ex-post optimal set with probability at least p for largest possible value of p. Formally, we have that

$$\mathsf{PSM}_p(\mathcal{F}, D, \mathsf{ALG}) := \min_{e \in E} \Pr[e \in \mathsf{ALG}(\boldsymbol{w}) \mid e \in \mathsf{OPT}(\boldsymbol{w})],$$

and

$$\mathsf{PSM}_p(\mathcal{F}) := \inf_{D} \sup_{\mathsf{ALG}} \mathsf{PSM}_p(\mathcal{F}, D, \mathsf{ALG}).$$

Note that for single-choice settings, the above measure reduces to PSM. In Bahrani et al. [8, Definition 3], this notion is referred to as *probability-competitive* algorithms. It is a major open question whether, for general matroids in secretary settings, there exists a constant probability-competitive algorithm.

We now show the shortcoming of this measure for downward-closed prophet settings via the following example. In particular, we present a feasibility constraint and an instance in which it is possible to achieve a good approximation, but no algorithm can guarantee a performance of better than $O(1/n)$ according to PSM_p.

Example E.1. Consider a setting with n pairs of boxes, such that, for each pair i, one box has weight $w_{1,i} = 1$ deterministically, and the second has weight $w_{2,i} \sim \mathcal{U}[1, 2]$. All the deterministic boxes arrive first and the rest arrive later. The (downward-closed) feasibility constraint is that only elements from at most one pair can be selected.

According to the measure of PSM_p, the optimal set will always be a pair of two boxes, and any algorithm will have to guess what the pair exactly is to select each element of the optimal set, which happens with probability at most $1/n$. After the arrival of the first n deterministic boxes, let us consider an element the algorithm has selected with probability at most $1/n$: This element will belong to the optimal set with probability $1/n$. Hence, given that $e \in \mathsf{OPT}(\boldsymbol{w})$ with probability $1/n$, the algorithm will have selected e with probability at most $1/n$, i.e., $\mathsf{PSM}_p(\mathcal{F}) \leq 1/n$. Since the trivial algorithm that selects a random pair always achieves a $\mathsf{PSM}_p(\mathcal{F}) = 1/n$ and, as shown, no other algorithm can have better performance, this measure becomes uninformative of the algorithm's quality in combinatorial settings.

It is easy to see that both the RoE and EoR are both at least a constant for this specific $\mathcal{F}$: Let us consider the single-choice algorithm that tries to catch the box of maximum weight out of the $2n$

Fig. 4. n pairs of boxes: for each pair, $w_{1,i} = 1$, $w_{2,i} \sim D_2 = \mathcal{U}[1,2]$.

boxes and then selects the pair corresponding to this box. This algorithm is essentially treating the instance in Figure 4 as if $2n$ boxes arrived singularly. As discussed in Appendix A, an algorithm aiming to catch the box with maximum weight achieves an EoR of $1/e$. The algorithm then selects the pair where this box belongs. This means that, with probability $1/e$, any other box can have a weight at most that of the maximum, i.e., the total weight in any other pair can be at most twice as large as the weight the selected pair. Thus, $\mathsf{EoR}(\mathcal{F}) \geq 1/2e$. Furthermore, we have that, for all $\boldsymbol{w}$, the optimal pair weight $f(\boldsymbol{w}) \leq 2\tilde{f}(\boldsymbol{w})$, where $\tilde{f}(\boldsymbol{w})$ is the maximum weight out of the $2n$ boxes. Finally, this same algorithm satisfies

$$\mathbf{E}[a(\boldsymbol{w})] \geq \mathbf{E}[a(\boldsymbol{w}) \mid a(\boldsymbol{w}) = \tilde{f}(\boldsymbol{w})] \cdot \Pr\left[a(\boldsymbol{w}) = \tilde{f}(\boldsymbol{w})\right] \geq 1/e \cdot \mathbf{E}[\tilde{f}(\boldsymbol{w})] \geq 1/2e \cdot \mathbf{E}[f(\boldsymbol{w})].$$

Hence, $\mathsf{RoE}(\mathcal{F}) \geq 1/2e$.

Received 22 January 2024; revised 16 September 2024; accepted 26 January 2025