# A Bicriterion Concentration Inequality and Prophet Inequalities for $k$-Fold Matroid Unions


**Noga Alon**
Department of Mathematics, Princeton University, NJ, USA
Schools of Mathematics and Computer Science, Tel Aviv University, Tel Aviv, Israel

**Nick Gravin**
Key Laboratory of Interdisciplinary Research of Computation and Economics,
Shanghai University of Finance and Economics, China

**Tristan Pollner**
Department of Management Science and Engineering, Stanford University, CA, USA

**Aviad Rubinstein**
Department of Computer Science, Stanford University, CA, USA

**Hongao Wang**
Department of Computer Science, Purdue University, West Lafayette, IN, USA

**S. Matthew Weinberg**
Department of Computer Science, Princeton University, NJ, USA

**Qianfan Zhang**
Department of Computer Science, Princeton University, NJ, USA



## Abstract

We investigate prophet inequalities with competitive ratios approaching 1, seeking to generalize $k$-uniform matroids. We first show that large girth does *not* suffice: for all $k$, there exists a matroid of girth $\geq k$ and a prophet inequality instance on that matroid whose optimal competitive ratio is $\frac{1}{2}$. Next, we show $k$-fold matroid unions *do* suffice: we provide a prophet inequality with competitive ratio $1 - O(\sqrt{\frac{\log k}{k}})$ for any $k$-fold matroid union. Our prophet inequality follows from an online contention resolution scheme.

The key technical ingredient in our online contention resolution scheme is a novel bicriterion concentration inequality for arbitrary monotone 1-Lipschitz functions over independent items which may be of independent interest. Applied to our particular setting, our bicriterion concentration inequality yields "Chernoff-strength" concentration for a 1-Lipschitz function that is not (approximately) self-bounding.



**2012 ACM Subject Classification** Theory of computation → Online algorithms

**Keywords and phrases** Prophet Inequalities, Online Contention Resolution Schemes, Concentration Inequalities

**Digital Object Identifier** 10.4230/LIPIcs.ITCS.2025.4

**Related Version** *Full Version*: https://arxiv.org/abs/2411.11741

**Funding** *Noga Alon*: Research supported in part by NSF grant DMS-2154082.
*Nick Gravin*: Research is supported by National Key R & D Program of China (2023YFA1009500), NSFC grant 61932002, "the Fundamental Research Funds for the Central Universities in China".
*Aviad Rubinstein*: Research supported in part by NSF CCF-1954927, and a David and Lucile Packard Fellowship.
*S. Matthew Weinberg*: Research supported in part by NSF CCF-1955205. During Professor Weinberg's development of this paper, he participated as an expert witness on behalf of the State of Texas in ongoing litigation against Google (the "Google Litigation").
*Qianfan Zhang*: Research supported in part by NSF CCF-1955205.

**Acknowledgements** The authors are grateful to the anonymous reviewers for helpful feedback on the initial submission of this work.

## 1 Introduction

Prophet inequalities are fundamental problems in optimal stopping theory, whose study dates back to seminal work of Krengel and Sucheston [17], and that have wide applications across Economics and Computer Science (e.g., [8, 11]). A prophet inequality instance contains a ground set $E$ of elements, a family $\mathcal{F} \subseteq 2^E$ of feasible sets, and a collection of distributions $\{\mathcal{D}_e\}_{e\in E}$. For one element at a time, a random variable $v_e$ is drawn from distribution $\mathcal{D}_e$ independently and revealed to a gambler, who immediately and irrevocably decides whether to accept or reject $e$. The gambler must at all times maintain the set of accepted elements $A \in \mathcal{F}$, and gets payoff $\sum_{e\in A} v_e$ at the end of the game. A prophet inequality is $c$-competitive if it guarantees $\mathbf{E}[\sum_{e\in A} v_e] \geq c \cdot \mathbf{E}[\max_{S\in\mathcal{F}} \sum_{e\in S} v_e]$.[1]

Krengel and Sucheston's seminal result establishes a $\frac{1}{2}$-competitive prophet inequality for any instance where $\mathcal{F}$ is a 1-uniform matroid (i.e. at most one element is feasible to accept), and moreover establish that no better guarantee is possible.[2] For $k$-uniform matroids, however, a significantly improved guarantee of $1-O(\frac{1}{\sqrt{k}})$ is possible [1, 14, 10]. This motivates the following question: for a given $\varepsilon > 0$, what conditions on $\mathcal{F}$ suffice for a $(1-\varepsilon)$-competitive prophet inequality?

### Main Result I: Large Girth does not Suffice

A natural starting point to address this question is to first understand what makes $k$-uniform matroids "special" in the sense that the canonical hard instance cannot be embedded. One conjecture might be because $k$-uniform matroids have large *girth*: there are *no infeasible sets of size* $\leq k$. So, a natural first question to ask is whether $\mathcal{F}$ having large girth suffices in order to conclude that any instance over $\mathcal{F}$ admits a $c$-competitive prophet inequality. Our first main result establishes that large girth does *not* suffice.

▶ **Theorem 1.** *For all $k \geq 1$ and $\varepsilon > 0$, there exists a prophet inequality instance $(E, \mathcal{F}, \{\mathcal{D}\}_{e\in E})$ such that: (a) $(E, \mathcal{F})$ is a graphic matroid with girth $k$, and (b) $(E, \mathcal{F}, \{\mathcal{D}\}_{e\in E})$ does not admit a $(\frac{1}{2} + \varepsilon)$-competitive prophet inequality.*

Our construction leverages dense graphs of high girth (and a particular construction of [18]) in order to effectively embed multiple copies of the canonical hard 1-uniform instance. See Section 3 for further details.

### Main Result II: $k$-fold Matroid Unions Suffice

Theorem 1 motivates richer generalizations of $k$-uniform matroids. We next consider *$k$-fold matroid unions*, observing that $k$-uniform matroids are the union of $k$ 1-uniform matroids. Given a matroid $\mathcal{M} = (E, \mathcal{F})$ over ground set $E$ with feasible sets $\mathcal{F}$, the *$k$-fold union of* $\mathcal{M}$ is a new matroid $\mathcal{M}^k$ with ground set $E$ and feasible sets $\mathcal{F}^k := \{F_1 \cup F_2 \cup \cdots \cup F_k : F_1, F_2, \ldots, F_k \in \mathcal{F}\}$. That is, a set is feasible in $\mathcal{M}^k$ if it can be partitioned into $k$ sets that are each feasible in $\mathcal{M}$.

[1] The expectation is taken with respect to the random variables $\{v_e\}_{e\in E}$, which in turn makes $A$ a random variable.

[2] That is, there exist prophet inequality instances over 1-uniform matroids for which better than a $\frac{1}{2}$-competitive ratio is impossible. The hard instance is quite simple: $v_1 \sim \mathcal{D}_1$ is a point mass at 1, and $v_2 \sim \mathcal{D}_2$ is equal to $\frac{1}{\varepsilon}$ with probability $\varepsilon$ and 0 otherwise. A gambler who sees $v_1$ first cannot achieve expect reward exceeding 1, but a prophet who always takes the maximum can achieve expected reward of $2 - \varepsilon$.

▶ **Theorem 2.** *For every prophet inequality instance* $(E, \mathcal{F}^k, \{\mathcal{D}_e\}_{e\in E})$ *where* $(E, \mathcal{F}^k)$ *is the* $k$*-fold union of a matroid* $(E, \mathcal{F})$*, there exists a* $(1 - O(\sqrt{\frac{\log k}{k}}))$*-competitive prophet inequality.*

Our proof of Theorem 2 follows from a novel Online Contention Resolution Scheme (OCRS). An OCRS is parameterized by a ground set $E$, a feasibility family $\mathcal{F}$, and a vector of probabilities $\boldsymbol{x} \in \text{ConvexHull}(\{\mathbf{1}_F : F \in \mathcal{F}\}) \subseteq [0,1]^E$ (that is, $\boldsymbol{x}$ can be written as a convex combination of indicator vectors of feasible sets). One at a time, elements of $E$ are revealed and *active* with probability $x_e$ independently. If an element is active, it can be accepted or rejected (if inactive, it must be rejected), and the accepted elements must at all times be in $\mathcal{F}$. An OCRS is $c$-selectable if every element $e$ is accepted with probability at least $c \cdot x_e$. In this language, Theorem 2 follows from a novel $(1 - O(\sqrt{\frac{\log k}{k}}))$-selectable OCRS for $k$-fold matroid unions.

To prove our OCRS, we follow a similar framework as [12], and design a recursive decomposition of $\mathcal{F}$ over which to greedily accept active elements. There are two key challenges to applying their framework, which we overview in greater detail in Subsection 4.2. We give a representative example below.

Applied to the 1-uniform matroid, the [12] algorithm simply proposes "accept any active element independently with probability $b$." Then, linearity of expectation suffices to observe that there are at most $b$ elements in expectation that are both active and accepted,[3] and Markov's inequality suffices to guarantee that with probability at least $1-b$, no elements are accepted at all. This suffices to guarantee that for all $e$: (a) with probability at least $1-b$ it is feasible to accept $e$ when revealed, and (b) independently, we will accept $e$ with probability $b$ conditioned on $e$ being active and feasible. This implies a $b(1-b)$-selectable algorithm, which is optimized at $b = \frac{1}{2}$.

Applied to the $k$-uniform matroid, a natural algorithm would again be "accept any active element independently with probability $b$." Then, linearity of expectation still suffices to observe that there are at most $bk$ elements in expectation that are both active and accepted, but Markov's inequality only guarantees that with probability at least $1-b$, at most $k$ elements are accepted. This would lead to the same $\frac{1}{4}$-selectable OCRS, which is not the desired $1 - O(\sqrt{\frac{\log k}{k}})$. Of course, the obvious fix is to use a significantly stronger concentration inequality than Markov's. E.g., a Chernoff bound suffices to guarantee that with probability at least $1 - \frac{1}{k}$ at most $k-1$ elements are accepted, when $b = 1 - O(\sqrt{\frac{\log k}{k}})$. This leads to the desired $1 - O(\sqrt{\frac{\log k}{k}})$ selectable OCRS for $k$*-uniform* matroids. However, Chernoff bounds are insufficient for the general class of $k$-fold matroid unions – the probability that a particular element is feasible to accept is a highly combinatorial function that depends on the underlying matroid structure. Thus our Theorem 2 has two components: first, a decomposition that reduces the OCRS problem to a concentration inequality and second, a novel concentration inequality, which is our third main result.

## Main Result III: A Bicriterion Concentration Inequality

Putting aside prophet inequalities for a moment, concentration inequalities are a core aspect of applied probability with widespread application across many areas of Computer Science. One representative setting is the following: Let $f : \{0,1\}^E \to \mathbb{R}$ be some function, and let $\boldsymbol{X} = \langle X_e \rangle_{e\in E}$ be a vector of independent Bernoulli random variables, where $X_e \sim \text{Ber}(p_e)$. A canonical question asks: what is the probability that $f(\boldsymbol{X})$ exceeds $\mathbf{E}[f(\boldsymbol{X})] + t$?

[3] There are at most 1 elements in expectation that are active, and each active element is accepted with probability $b$.

On one extreme, McDiarmid's inequality holds whenever $f$ is 1-Lipschitz. On the other, Chernoff bounds are *significantly* stronger, if $f$ is linear (and 1-Lipschitz). In between, "Chernoff-strength" concentration holds whenever $f$ is fractionally-subadditive or (approximately) self-bounding [5, 6, 22, 7, 27], but this provably does not extend even to the case when $f$ is subadditive [27].

Our third main result provides a *bicriterion concentration inequality* for any monotone 1-Lipschitz function. Specifically, if $\boldsymbol{X}$ is a vector of Bernoulli random variables with probability vector $\boldsymbol{p}$, let $\boldsymbol{X}^{(s)}$ denote a vector of Bernoulli random variables with probability vector $e^{-s}\boldsymbol{p}$. That is, each probability $p_i$ has been decreased by a factor of $e^{-s}$. Our new concentration inequality establishes:

▶ **Theorem 3.** *Let $f : \{0,1\}^E \to \mathbb{R}$ be a monotone 1-Lipschitz function. For any $s \in (0,1]$, $t > 0$:*

$$\mathbf{Pr}\left[f(\boldsymbol{X}^{(s)}) \geq \mathbf{E}[f(\boldsymbol{X})] + t\right] \leq e^{-st}.$$

A helpful comparison point is McDiarmid's inequality, which instead proves the following: $\mathbf{Pr}\left[f(\boldsymbol{X}) \geq \mathbf{E}[f(\boldsymbol{X})] + t\right] \leq e^{-2t^2/|E|}$. The distinctions are: (a) our concentration inequality is bicriterion – we analyze $f(\boldsymbol{X}^{(s)})$ instead of $f(\boldsymbol{X})$, and (b) our concentration has an exponent of $-st$ instead of $-2t^2/|E|$. In particular, McDiarmid's inequality depends on the dimension $|E|$ and cannot possibly kick in for $t \ll \sqrt{|E|}$, whereas our concentration inequality can kick in for any $t > 1/s$. A representative example to have in mind might be $s = \sqrt{\log(1/\varepsilon)/\,\mathbf{E}[f(\boldsymbol{X})]}$ and $t = \sqrt{\log(1/\varepsilon)\,\mathbf{E}[f(\boldsymbol{X})]}$. This results in a tail probability of $\varepsilon$ for exceeding $\mathbf{E}[f(\boldsymbol{X})]$ by $\sqrt{\log(1/\varepsilon)}$ multiples of $\sqrt{\mathbf{E}[f(\boldsymbol{X})]}$, which is "Chernoff-strength". *But*, this concentration holds only for $f(\boldsymbol{X}^{(s)})$, rather than $f(\boldsymbol{X})$. This suffices for our application.

To prove Theorem 3, we utilize the entropy method for self-bounding functions [5, 6, 22, 7] in an unconventional way. We give a more detailed technical overview in Subsection 5.1.

## 1.1 Related Work

There are three strands of related work: prophet inequalities, concentration inequalities, and attempts to generalize $k$-uniform guarantees.

### Prophet Inequalities

Prophet inequalities have a long history in Mathematics, Computer Science, and Operations Research. Representative results include Krengel and Sucheston's initial $\frac{1}{2}$-approximation [17], Samuel-Cahn's elegant thresholding strategy [25], Chawla et al.'s connection to Bayesian mechanism design [8], Kleinberg and Weinberg's extension to matroids [16], and Dutting et al.'s connection to Price of Anarchy [11].

Of particular relevance to our work are prophet inequalities for $k$-uniform matroids. The first $1 - O(\sqrt{\frac{\log k}{k}})$ approximation was developed by [13], and the first asymptotically tight $1 - O(\frac{1}{\sqrt{k}})$ approximation was developed by [1]. Subsequent works achieve the same $1 - O(\frac{1}{\sqrt{k}})$ approximation with sample access [4], the optimal OCRS [14], or a simpler OCRS [10]. The most technically related paper to our work is [12], whose OCRS framework we leverage. It remains an open question whether the prophet inequality for $k$-fold matroid unions can be improved to $1 - O(\frac{1}{\sqrt{k}})$.

### Concentration Inequalities

The most related concentration inequalities fit the same framework but consider different $f$. McDiarmid's inequality [21] holds for all 1-Lipschitz $f$, Schechtman's inequality holds for $f$ that are subadditive [26], Bucheron et al. derive an inequality for $f$ that are self-bounding functions [5], and Vondrák derives an inequality for $f$ that are fractionally subadditive [27]. These inequalities are commonly used across Theoretical Computer Science, and especially within combinatorial prophet inequalities and Bayesian mechanism design [24, 23].

### Generalizing $k$-uniform matroids

Recent work of [9] considers (offline) contention resolution and correlation gap inequalities. Here too, guarantees for $k$-uniform matroids are significantly stronger than what is achievable for arbitrary matroids. Their work similarly extends guarantees achievable for $k$-uniform matroids to $k$-fold matroid unions. In comparison to our work: (a) the general motivation is the same – both works seek to extend stronger guarantees for $k$-uniform matroids to more general settings, (b) the problems studied and technical aspects are orthogonal,[4] (c) our work also proposes a bicriterion concentration inequality.

Another generalization of $k$-uniform matroids are *packing constraints*, where each element has a $d$-dimensional size in $[0,1]^d$ and one can accept a subset of elements if their size vectors sum to at most $k$ in every coordinate. Packing constraints have been studied in various online settings, including secretary model [15], prophet model [2], and mixed model [3].

## 2 Preliminaries

### Prophet Inequalities

In the *prophet inequality* problem, we are given a ground set of elements $E$, a downward-closed family of feasible sets $\mathcal{F} \subseteq 2^E$, and a distribution $\mathcal{D}_e$ associated with each element $e \in E$. Elements arrive in an adversarial order.[5] As each element $e$ arrives, its value $v_e$, independently drawn from $\mathcal{D}_e$, is revealed. At this point, an irrevocable decision must be made whether to include $e$ in its output $A$, while keeping $A \in \mathcal{F}$.

For $c \in [0,1]$, we say an online algorithm implies a $c$-*competitive* prophet inequality for $\mathcal{F}$, if for any distributions $\{\mathcal{D}_e\}_{e\in E}$,

$$\mathbf{E}\left[\sum_{e\in A} v_e\right] \geq c \cdot \mathbf{E}\left[\max_{S\in\mathcal{F}} \sum_{e\in S} v_e\right]$$

where the expectation is taken with respect to random variables $\{v_e\}_{e\in E}$ and the internal randomness of the algorithm.

### Online Contention Resolution Schemes

Given a ground set of elements $E$ and a downward-closed family of feasible sets $\mathcal{F} \subseteq 2^E$, we define the polytope of $\mathcal{F}$ as the convex hull of all characteristic vectors of feasible sets, i.e., $\mathcal{P}_{\mathcal{F}} = \text{ConvexHull}(\{\mathbf{1}_F : F \in \mathcal{F}\}) \subseteq [0,1]^E$.

[4] While in principle, contention resolution and online contention resolution may appear similar, the relevant techniques are fundamentally different with little overlap. Similarly, while correlation gap inequalities are sometimes a useful tool in prophet inequalities, in this case there is no overlap.

[5] There are various adversarial models. The weakest is the *fixed-order adversary*, which sets the arrival order offline, based solely on the distributions. The strongest is the *almighty adversary*, which sets the arrival order online, with full knowledge of all realizations of randomness and the algorithm's past decisions. Our negative result in Section 3 applies to fixed-order adversary, while our positive result in Section 4 holds against almighty adversary.

An *online contention resolution scheme* (OCRS) takes a vector $\boldsymbol{x} \in \mathcal{P}_{\mathcal{F}}$ as input. Let $R(\boldsymbol{x}) \subseteq E$ be a random set where each element $e \in E$ is in $R(\boldsymbol{x})$ independently with probability $x_e$. The OCRS sees membership in $R(\boldsymbol{x})$ of elements in $E$, arriving in an adversarial order; when each $e \in E$ arrives, if $e \in R(\boldsymbol{x})$ (i.e., $e$ is "active"), the scheme must decide irrevocably whether to include $e$ in its output $A$, while keeping $A \in \mathcal{F}$.

For $c \in [0,1]$, an OCRS is called *$c$-selectable* for $\mathcal{F}$, if for any $\boldsymbol{x} \in \mathcal{P}_{\mathcal{F}}$,

$$\mathbf{Pr}[e \in A \mid e \in R(\boldsymbol{x})] \geq c \quad \forall e \in E$$

where $A \in \mathcal{F}$ is the output of the OCRS, and the probability is measured with respect to $R(\boldsymbol{x})$ and internal randomness of the OCRS. As shown in [12], a $c$-selectable OCRS directly implies a $c$-competitive prophet inequality.

▶ **Lemma 4** ([12]). *For a ground set $E$ and a family of feasible sets $\mathcal{F} \subseteq 2^E$, a $c$-selectable OCRS for $\mathcal{F}$ implies a $c$-competitive prophet inequality for $\mathcal{F}$.*

## Matroids

A *matroid* $\mathcal{M} = (E, \mathcal{I})$ is defined by a ground set of elements $E$ and a non-empty downward-closed family of independent sets $\mathcal{I} \subseteq 2^E$ with the *exchange property*, i.e., for every $A, B \in \mathcal{I}$ where $|A| > |B|$, there exists an element $e \in A \setminus B$ such that $B \cup \{e\} \in \mathcal{I}$. Given a matroid $\mathcal{M} = (E, \mathcal{I})$, the following notations are used throughout the paper:

- The *rank* of a set $S \subseteq E$ is the size of the largest independent set contained in $S$: $\text{rank}(S) = \max\{|I| : I \subseteq S, I \in \mathcal{I}\}$.
- The *span* of a set $S \subseteq E$ is the set of elements that is not independent from $S$: $\text{span}(S) = \{e \in E : \text{rank}(S) = \text{rank}(S \cup \{e\})\}$.
- The *restriction* of $\mathcal{M}$ to a set $S \subseteq E$ is a matroid $\mathcal{M}|_S = (S, \mathcal{I}|_S) = (S, \{I \in \mathcal{I} : I \subseteq S\})$.
- The *girth* of $\mathcal{M}$ is the size of the smallest dependent set: $\text{girth}(\mathcal{M}) = \min\{|S| : S \subseteq E, S \notin \mathcal{I}\}$.

Following are some special matroids that we will use later.

▶ **Example 5** (Uniform matroid). A *$k$-uniform matroid* $\mathcal{M} = (E, \mathcal{I})$ is a matroid in which the independent sets are exactly the sets that contains at most $k$ elements for an integer $k \geq 1$, i.e, $\mathcal{I} = \{I \subseteq E : |I| \leq k\}$.

▶ **Example 6** (Graphical matroid). A *graphical matroid* $\mathcal{M} = (E, \mathcal{I})$ is a matroid in which the independent sets are the forests in a given undirected graph $G = (V, E)$, i.e., $\mathcal{I} = \{I \subseteq E : I \text{ is acyclic in } G\}$.

We formally define $k$-fold matroid union as follows.

▶ **Definition 7** ($k$-fold matroid union). *Given a matroid $\mathcal{M} = (E, \mathcal{I})$ and an integer $k \geq 1$, the* $k$-fold union *of $\mathcal{M}$ is defined as $\mathcal{M}^k = \underbrace{\mathcal{M} \vee \mathcal{M} \vee \cdots \vee \mathcal{M}}_{k \text{ times}} = (E, \mathcal{I}^k)$ where*

$$\mathcal{I}^k = \{I_1 \cup I_2 \cup \cdots \cup I_k : I_1, I_2, \ldots, I_k \in \mathcal{I}\}.$$

*In other words, a set $I$ is independent in $\mathcal{M}^k$ if and only if $I$ can be partitioned into at most $k$ independent sets in $\mathcal{M}$. Note that $\mathcal{M}^k$ remains a matroid by the closure property of matroid union.*

# 3 Large Girth is Not Sufficient

In this section, we prove that a large girth is not sufficient for matroids (specifically, graphical matroids) to have a prophet inequality with a competitive ratio better than $\frac{1}{2}$.

▶ **Theorem 1.** *For all $k \geq 1$ and $\varepsilon > 0$, there exists a prophet inequality instance $(E, \mathcal{F}, \{\mathcal{D}\}_{e\in E})$ such that: (a) $(E, \mathcal{F})$ is a graphic matroid with girth $k$, and (b) $(E, \mathcal{F}, \{\mathcal{D}\}_{e\in E})$ does not admit a $(\frac{1}{2} + \varepsilon)$-competitive prophet inequality.*

To construct a hard instance, we start with a dense graph of large girth. We then transform the graph by splitting each edge $(v_1, v_2)$ into two edges $(v_1, u)$ and $(v_2, u)$, where $u$ is a newly introduced vertex. We obtain the final hard instance of the prophet inequality problem by embedding the hard instance of the single-item case into each of these edge pairs $(v_1, u), (v_2, u)$.

The hardness of this instance arises from the following observation: without accepting both edges in a pair, the instance essentially reduces to $|E|$ independent hard instances of the single-item case. On the other hand, one can accept at most $|V| - 1$ extra pairs of edges (in addition to $|E|$ single-item problems) at the same time without forming a cycle, which could not contribute a lot to the final solution because the graph is dense.

**Proof of Theorem 1.** We employ a construction of [18] which provides dense graphs of large girth. In particular, we will use that for any fixed $k$ there exists some arbitrarily large $n$ such that there is a graph $G_n$ on $n$ vertices with at least $n \log n$ edges and girth at least $k$.[6] Specifically, consider the graph $G_n$ with vertices $V(G_n) = \{v_1, v_2, \ldots, v_n\}$ and edges $E(G_n) = \{e_1, e_2, \ldots, e_m\}$, where $m \geq n \log n$ and each edge $e_i = (a(i), b(i)) \in E(G_n)$ connects vertices $a(i)$ and $b(i)$ in $V(G_n)$. We construct a new graph $H_n$ with $n + m$ vertices as follows:

- Begin with a set of $n+m$ vertices, labeled $V(H_n) := \{w_1, w_2, \ldots, w_n\} \sqcup \{u_1, u_2, \ldots, u_m\}$.
- For each edge $e_i$ in $G_n$ connecting $v_{a(i)}$ and $v_{b(i)}$, add in $H_n$ an edge between $u_i$ and $w_{a(i)}$ (call it $f_i$) as well as an edge between $u_i$ and $w_{b(i)}$ (call it $f'_i$).

Hence $H_n$ has a total of $2m$ edges. For $1 \leq i \leq m$, let the associated random variable $X_{f_i}$ of $f_i$ be a constant 1, and let the associated random variable $X_{f'_i}$ of $f'_i$ follow a distribution which takes a value of $\frac{1}{\varepsilon}$ with probability $\varepsilon$, and a value of 0 with probability $1 - \varepsilon$. We consider an instance of the prophet inequality problem where the online algorithm is presented edges in the order $(f_1, f'_1, f_2, f'_2, \ldots, f_m, f'_m)$.

We first lower bound $\mathrm{OPT}(H_n)$, the expected value the optimal offline algorithm gets on this instance. Note that an offline algorithm could simply look at each pair $\{f_i, f'_i\}$ and take whichever edge has higher realized weight; this cannot create a cycle because every edge selected will be incident to a vertex of degree 1. We hence have the bound

$$\mathrm{OPT}(H_n) \geq \sum_{i=1}^{m} \left( \varepsilon \cdot \frac{1}{\varepsilon} + (1 - \varepsilon) \cdot 1 \right) = m(2 - \varepsilon).$$

Fix an online algorithm $\mathcal{A}$, and we now give an upper bound on its expected performance $\mathcal{A}(H_n)$ on the instance. The lower bound relies on the following observation.

▷ Claim 8. There are at most $n - 1$ values of $i$ in $\{1, 2, \ldots, m\}$ such that $\mathcal{A}$ accepts both $f_i$ and $f'_i$.

[6] In fact, [18] prove a significantly stronger result, but the weaker version stated above suffices for our purposes.

Proof. Suppose there are at least $n$ such values of $i$; call them $i_1$, $i_2$, ..., $i_n$. As the original graph $G$ has $n$ vertices, and a forest on $n$ vertices has at most $n-1$ edges, we clearly see that there is a cycle among $\{e_{i_1}, e_{i_2}, \dots, e_{i_n}\}$. That however would imply there is a cycle in $H$; namely, follow the cycle that existed in $G$, but replace each edge $e_i$ with the edge $f_i$ followed by the edge $f_i'$. ◁

For each $1 \le i \le m$, we now consider cases for what $\mathcal{A}$ gets in expectation from $\{f_i, f_i'\}$ right after $f_i$ arrives:

- If $\mathcal{A}$ rejects $f_i$, then it clearly gets in expectation at most 1 from $\{f_i, f_i'\}$ because $\mathbb{E}[X_{f_i'}] = 1$.
- If $\mathcal{A}$ accepts $f_i$ and rejects $f_i'$, then it clearly gets weight at most 1 from $\{f_i, f_i'\}$.
- If $\mathcal{A}$ accepts $f_i$ and accepts $f_i'$, then it clearly gets weight at most $1 + \frac{1}{\varepsilon}$ from $\{f_i, f_i'\}$.

Let $C_1$ denote the set of all $i \in [m]$ such that $\mathcal{A}$ rejects $f_i$, let $C_2$ denote the set of all $i \in [m]$ such that $\mathcal{A}$ accepts $f_i$ and rejects $f_i'$, and let $C_3$ denote the set of all $i \in [m]$ such that $\mathcal{A}$ accepts $f_i$ and $f_i'$. Note $C_1$, $C_2$, and $C_3$ are random (disjoint) sets that may depend on the values realized by $\{X_{f_i}, X_{f_i'}\}_{i=1}^m$ and any randomness in $\mathcal{A}$. By the above cases, we can see that in expectation, $\mathcal{A}$ gets score at most

$$\sum_{i \in C_1} 1 + \sum_{i \in C_2} 1 + \sum_{i \in C_3} \left(1 + \frac{1}{\varepsilon}\right) = |C_1| + |C_2| + |C_3| \cdot \left(1 + \frac{1}{\varepsilon}\right).$$

Although $|C_1|$, $|C_2|$, and $|C_3|$ are random variables, $|C_1| + |C_2| \le m$ always, and by Claim 8 we have $|C_3| \le n-1$ always. Hence, in expectation (averaging over all possible realizations of $C_1$, $C_2$, and $C_3$), we can bound the performance of $\mathcal{A}$ on $H_n$ by $\mathcal{A}(H_n) \le m + n\left(1 + \frac{1}{\varepsilon}\right)$. As $n$ grows, we can compute

$$\liminf_{n \to \infty} \frac{\mathcal{A}(H_n)}{\mathrm{OPT}(H_n)} \le \lim_{n \to \infty} \frac{m + n\left(1 + \frac{1}{\varepsilon}\right)}{m(2-\varepsilon)} = \frac{1}{2-\varepsilon}.$$

Taking $\varepsilon \to 0$ demonstrates the claimed result. ◀

## 4 $k$-Fold Unions are Sufficient

Our main goal in the section is to construct a good OCRS for $k$-fold matroid unions (Theorem 9). Combining with the reduction from prophet inequalities to OCRSs by [12] (Lemma 4), this immediately implies the existence of good prophet inequality for all $k$-fold matroid unions (Theorem 2).

▶ **Theorem 9.** *There exists a $(1 - O(\sqrt{\frac{\log k}{k}}))$-selectable OCRS for any $k$-fold matroid union $\mathcal{M}^k$.*

Our OCRS for $k$-fold matroid unions builds on the chain decomposition approach used in the matroid OCRS by [12], outlined in Subsection 4.1. We overview our approach and highlight main difficulties in Subsection 4.2. The construction is then formally given and analyzed in Subsection 4.3, where the bicriterion concentration inequality in Section 5 is used to bound its selectability.

## 4.1 Recap: OCRS for general matroids

We briefly describe the idea of the $\frac{1}{4}$-selectable matroid OCRS by [12]. Specifically, they show that for any parameter $b \in (0,1)$, there exists a $(1-b)$-selectable OCRS for any matroid $\mathcal{M} = (E, \mathcal{I})$ and $\boldsymbol{x} \in b \cdot \mathcal{P}_{\mathcal{M}}$. Note that one can "scale down" a vector $\boldsymbol{x}$ from $\mathcal{P}_{\mathcal{M}}$ to $b \cdot \mathcal{P}_{\mathcal{M}}$ by only considering each element independently with probability $b$. Formally:

▶ **Fact 10.** *For $b, c \in (0,1)$ and any matroid $\mathcal{M}$, a $c$-selectable OCRS for all $\boldsymbol{x} \in b \cdot \mathcal{P}_{\mathcal{M}}$ implies a $bc$-selectable OCRS for all $\boldsymbol{x} \in \mathcal{P}_{\mathcal{M}}$.*

Therefore, it follows that a $b(1-b)$-selectable ORCS exists for any matroid $\mathcal{M}$ and $\boldsymbol{x} \in \mathcal{P}_{\mathcal{M}}$. By letting $b = \frac{1}{2}$, they obtain a $\frac{1}{4}$-selectable matroid OCRS.

### The greedy algorithm

Let us start with the simple greedy algorithm that always accepts the active element whenever possible. When $\mathcal{M}$ is a 1-uniform matroid, the greedy algorithm is actually $(1-b)$-selectable for $\boldsymbol{x} \in b \cdot \mathcal{P}_{\mathcal{M}}$ (i.e., $\sum_{e \in E} x_e \le b$ since $\mathcal{M}$ is 1-uniform), since the selectability of an element $e \in E$ can be easily lower bounded as

$$\begin{aligned}\mathbf{Pr}[e \text{ is accepted} \mid e \text{ is active}] &\ge \mathbf{Pr}[\text{no other element is active} \mid e \text{ is active}] \\ &\ge \mathbf{Pr}[\text{no element is active}].\end{aligned}$$

The first inequality holds because when there is no active elements besides $e$, the greedy algorithm can always accept $e$ even if it arrives at the end. The second inequality holds due to the independence between elements. Moreover, by Markov's inequality,

$$\begin{aligned}\mathbf{Pr}[\text{no element is active}] &= 1 - \mathbf{Pr}[|R(\boldsymbol{x})| \ge 1] \\ &\ge 1 - \mathbf{E}[|R(\boldsymbol{x})|] = 1 - \sum_{e \in E} x_e \ge 1 - b.\end{aligned}$$

(Recall that $R(\boldsymbol{x})$ is the set of active elements.)

The first half of argument applies when $\mathcal{M}$ is a general matroid: for every element $e \in E$,

$$\mathbf{Pr}[e \text{ is accepted} \mid e \text{ is active}] \ge \mathbf{Pr}[e \notin \mathrm{span}(R(\boldsymbol{x}))].$$

However, unlike in 1-uniform matroids, the probability that an element $e \in E$ is spanned by active elements $R(\boldsymbol{x})$ could be much smaller than $1-b$, even for a scaled $\boldsymbol{x} \in b \cdot \mathcal{P}_{\mathcal{M}}$. In fact, the selectability of the greedy algorithm can be arbitrarily bad for a general matroid $\mathcal{M}$ (see, e.g., [19]).

### Protection

Consider Algorithm 1, a modified greedy algorithm with a *protection set* $S \subsetneq E$ that only handles elements in $E \setminus S$. Intuitively, the algorithm accepts every active element $e \in E \setminus S$ whenever it does not conflict with any element in $S$. As a result, elements in $S$ are "prioritized" over those in $E \setminus S$: regardless of which independent set from $S$ is accepted, it remains an independent set when combined with the accepted elements in $E \setminus S$.

**Algorithm 1** Modified greedy algorithm for $\mathcal{M} = (E, \mathcal{I})$ with a protection set $S \subseteq E$.

```
A ← ∅                                              ▷ the set of accepted elements in E \ S
for each arriving active element e ∈ E \ S do
    if e ∉ span(A ∪ S) then
        A ← A ∪ {e}                                ▷ accepts element e
```

For the modified greedy algorithm, we can similarly lower bound the selectability for $e \in E \setminus S$:

$$\mathbf{Pr}[e \text{ is accepted} \mid e \text{ is active}] \geq \mathbf{Pr}[e \notin \text{span}(R(\boldsymbol{x}) \cup S)].$$

The good news is that, such probabilities can be further lower bounded by $1 - b$ for the $S$ obtained using Algorithm 2, an iterative algorithm that updates $S$ by adding an element $e$ whenever $\mathbf{Pr}[e \in \text{span}(R(\boldsymbol{x}) \cup S)] > b$.

**Algorithm 2** Find a protection set $S$ for $\mathcal{M} = (E, \mathcal{I})$ and $\boldsymbol{x} \in b \cdot \mathcal{P}_{\mathcal{M}}$.

```
function PROTECT(M, x, b)
    S ← ∅
    while ∃e ∈ E \ S,  Pr[e ∈ span(R(x) ∪ S)] > b do
        S ← S ∪ {e}
    return S
```

Note that Algorithm 2 always terminates since $E$ is a finite set, and the modified greedy algorithm with this protection set $S$ guarantees $(1-b)$-selectability for every element $e \in E \setminus S$. More importantly, the protection is non-trivial, i.e., $S$ is a proper subset of $E$.

▶ **Lemma 11** ([12]). *For any matroid $\mathcal{M} = (E, \mathcal{I})$ and $\boldsymbol{x} \in b \cdot \mathcal{P}_{\mathcal{M}}$, $\textsc{Protect}(\mathcal{M}, \boldsymbol{x}, b) \subsetneq E$.*

Therefore, it remains to get a good OCRS for $\mathcal{M}|_S$ and $\boldsymbol{x}|_S$, the restriction of the original matroid and vector to the protection set $S$.

### Chain decomposition

The matroid OCRS in [12] starts with an offline prepossessing that finds the following *chain decomposition* of the elements:

$$\emptyset = N_\ell \subsetneq N_{\ell-1} \subsetneq \cdots \subsetneq N_1 \subsetneq N_0 = E$$

where $N_{i+1} = \textsc{Protect}(\mathcal{M}|_{N_i}, \boldsymbol{x}|_{N_i}, b)$ for every $0 \leq i < \ell$. And the OCRS is then operates by invoking Algorithm 1 on matroid $\mathcal{M}|_{N_i}$ with a protection set $N_{i+1}$ for each $e \in N_i \setminus N_{i+1}$.

It is easy to see that these algorithms together produces an independent set of $\mathcal{M}$, and the selectability for each element $e \in N_i \setminus N_{i+1}$ is

$$\mathbf{Pr}[e \text{ is accepted} \mid e \text{ is active}] \geq 1 - \mathbf{Pr}[e \in \text{span}_{\mathcal{M}|_{N_i}}(R(\boldsymbol{x}|_{N_i}) \cup N_{i+1})] \geq 1 - b$$

where the last inequality holds due to the way $N_{i+1}$ is obtained using Algorithm 2. By setting $b = \frac{1}{2}$, the resulting OCRS is $\frac{1}{2}$-selectable given any matroid $\mathcal{M}$ and $\boldsymbol{x} \in \frac{1}{2} \cdot \mathcal{P}_{\mathcal{M}}$.

## 4.2 Overview of our construction

We now give a high-level overview of our construction and highlight main difficulties. Let us first examine the case when $\mathcal{M}$ is a $k$-uniform matroid and see why the simple greedy algorithm works better for larger $k$.

### Intuition from $k$-uniform matroids

When $\mathcal{M}$ is a $k$-uniform matroid, it turns out that the simple greedy algorithm that always accepts the active element whenever possible yields an OCRS with a selectability of $1 - O(\sqrt{\frac{\log k}{k}})$. To see this, consider the following tighter analysis of selectability for $k$-uniform matroids: for every element $e \in E$,

$$\mathbf{Pr}[e \text{ is accepted} \mid e \text{ is active}] \geq \mathbf{Pr}[e \notin \operatorname{span}(R(\boldsymbol{x}))] = \mathbf{Pr}[|R(\boldsymbol{x})| < k].$$

Intuitively, $|R(\boldsymbol{x})|$ represents the number of slots occupied by active elements, and we know $|R(\boldsymbol{x})| < k$ indicates $e \notin \operatorname{span}(R(\boldsymbol{x}))$. We want the bad event $|R(\boldsymbol{x})| \geq k$ to occur with a small probability.

Note that $|R(\boldsymbol{x})|$ is a sum of Bernoulli random variables and it concentrates very well: if we consider a slightly scaled-down $\boldsymbol{x} \in (1 - O(\sqrt{\frac{\log k}{k}})) \cdot \mathcal{P}_{\mathcal{M}}$, Chernoff bound (Theorem 27) tells us that $\mathbf{Pr}[|R(\boldsymbol{x})| \geq k] \leq \frac{1}{k}$. By Fact 10, one can further derive an OCRS for $k$-uniform matroids with a selectability of $(1 - O(\sqrt{\frac{\log k}{k}}))(1 - \frac{1}{k}) = 1 - O(\sqrt{\frac{\log k}{k}})$.

To summarize, the greedy algorithm performs well on $k$-uniform matroids because of the existence of a **fine-grained occupancy indicator $|R(\boldsymbol{x})|$ that concentrates well**.

### Main idea and challenges

For a $k$-fold matroid union $\mathcal{M}^k = (E, \mathcal{I}^k)$, the simple greedy algorithm could perform very poor due to inherent non-uniformity of $\mathcal{M}^k$. In the matroid OCRS by [12], this is resolved using the idea of chain decomposition. For each level, an iterative procedure (Algorithm 2) is used to find a protection set $S$ that includes all elements that are easily spanned by $R(\boldsymbol{x}) \cup S$. This is done by directly looking at the probability $\mathbf{Pr}[e \in \operatorname{span}(R(\boldsymbol{x}) \cup S)]$.

Our idea is to construct a different chain decomposition based on functions $\omega_e(\cdot) : 2^E \to [0, k]$ that act as a "generalized occupancy indicator" for each element $e$, such that $\omega_e(\emptyset) = 0$, $\omega_e(S) = k$ if $e$ is spanned by the set $S$, and we want $\omega_e(\cdot)$ to be as smooth as possible (i.e., 1-Lipschitz). For each level of the chain decomposition, we will add $e$ to the protection set $S$ whenever the expected occupancy $\mathbf{E}[\omega_e(R(\boldsymbol{x}) \cup S)]$ is large.

For $k$-uniform matroids, a simple occupancy indicator would be $\omega_e(S) = \min(k, |S|)$ (since we require its value to be between 0 and $k$). However, extending the definition of an occupancy function to a general $k$-fold matroid union introduces several challenges:

1. (**Compatibility with chain decomposition**) The most crucial part of the chain decomposition in [12] is to show the protection set $S$ is always a proper subset of $E$ (Lemma 11). Similarly, we will need to show that it is always possible to find a protection set $S \subsetneq E$ such that the expected occupancy $\mathbf{E}[\omega_e(R(\boldsymbol{x}) \cup S)]$ for every $e \in E \setminus S$ is smaller than $k$ by a large enough margin.
2. (**Chernoff-strength concentration**) Based on the fact that $\mathbf{E}[\omega_e(R(\boldsymbol{x}) \cup S)]$ is sufficiently smaller than $k$, we ultimately want to show that $\mathbf{Pr}[\omega_e(R(\boldsymbol{x}) \cup S) = k]$ is very small, which would imply a good selectability for $e$. This is simple for $k$-uniform matroids by using Chernoff bound. However, it turns out $\omega_e(\cdot)$ for general $k$-fold matroid unions does not admit a standard Chernoff-strength concentration inequality, and much more efforts are required to achieve a similar selectability guarantee.

## 4.3 An OCRS for $k$-fold matroid unions

In Subsubsection 4.3.1, we define our candidate occupancy functions and show some useful properties. Then, in Subsubsection 4.3.2, we show these functions are compatible with the chain decomposition approach and can be used to get an OCRS for $k$-fold matroid unions.

Finally, in Subsubsection 4.3.3, we prove the selectability of this OCRS by showing these functions concentrates well enough using Theorem 3. Some proofs in the section are deferred to Appendix B for ease of reading.

### 4.3.1 The occupancy function

To define the occupancy function, we will instead work with the following *extended $k$-fold unions* which essentially introduces $k$ parallel copies for each element. They are still matroids, and OCRS for them implies OCRS for $k$-fold matroid unions. Therefore, it suffices for us to give an OCRS for the extended $k$-fold union.

▶ **Definition 12** (Extended $k$-fold union). *Given a matroid $\mathcal{M} = (E, \mathcal{I})$ and an integer $k \geq 1$, let $\mathcal{M}_* = (E_*, \mathcal{I}_*)$ be the matroid that contains $k$ parallel copies $(e, 1), \ldots, (e, k)$ of each element $e \in E$. Formally,*

$$E_* = E \times [k] = \{(e, i) : e \in E, i \in [k]\},$$
$$\mathcal{I}_* = \{\{(e_1, i_1), \ldots, (e_t, i_t)\} : \{e_1, \ldots, e_t\} \in \mathcal{I}, i_1, \ldots, i_t \in [k]\}.$$

*And we define the* extended $k$-fold union $\mathcal{M}_*^k = (E_*, \mathcal{I}_*^k)$ *of $\mathcal{M}$ to be the $k$-fold union of $\mathcal{M}_*$.*

▶ **Lemma 13.** *The extended $k$-fold union $\mathcal{M}_*^k$ of a matroid $\mathcal{M}$ is a matroid. Furthermore, a $c$-selectable OCRS for $\mathcal{M}_*^k$ implies a $c$-selectable OCRS for $\mathcal{M}^k$.*

We are now ready to define the following *occupancy function* on $\mathcal{M}_*^k = (E_* = E \times [k], \mathcal{I}_*^k)$. Intuitively, the function indicates the number of "slots" for elements $(e, \cdot) \in E_*$ that are occupied by the elements in $S$. We then show the occupancy function has good properties: it is monotone and 1-Lipschitz. More importantly, the value of $\omega_e(S)$ can be used to deduce whether $(e, \cdot) \in E_*$ is spanned by other elements in $S$.

▶ **Definition 14** (Occupancy function). *Given an extended $k$-fold union $\mathcal{M}_*^k = (E_* = E \times [k], \mathcal{I}_*^k)$, for every $e \in E$, define its* occupancy function $\omega_e : 2^{E_*} \to [0, k]$ *as the function where for all $S \subseteq E_*$,*[7]

$$\omega_e(S) = k - \mathrm{rank}(S \cup (\{e\} \times [k])) + \mathrm{rank}(S).$$

▶ **Lemma 15.** *For any extended $k$-fold union $\mathcal{M}_*^k = (E_*, \mathcal{I}_*^k)$ and element $(e, i) \in E_*$, $\omega_e$ satisfies*

1. (Monotone) $\omega_e(S) \leq \omega_e(T)$ *for every $S \subseteq T \subseteq E_*$;*
2. (1-Lipschitz) $\omega_e(S \cup \{a\}) - \omega_e(S) \leq 1$ *for every $S \subseteq E_*$ and $a \in E_*$.*

▶ **Lemma 16.** *For any extended $k$-fold union $\mathcal{M}_*^k = (E_*, \mathcal{I}_*^k)$, element $(e, i) \in E_*$, and set $S \subseteq E_*$, $\omega_e(S) < k$ implies $(e, i) \notin \mathrm{span}(S \setminus \{(e, i)\})$.*

▶ **Example 17.** When $\mathcal{M}$ is a 1-uniform matroid of size $n$, its extended $k$-fold union $\mathcal{M}_*^k$ is a $k$-uniform matroid of size $kn$. For every $e \in E$ and $S \subseteq E_*$, we have

$$\mathrm{rank}(S \cup (\{e\} \times [k])) = \min(k, |S \cup (\{e\} \times [k])|) = k,$$
$$\mathrm{rank}(S) = \min(k, |S|).$$

Therefore, $\omega_e(S) = \min(k, |S|)$, i.e., the number of occupied slots by $S$.

Also, note that for any $\boldsymbol{x}_* \in (1 - O(\sqrt{\frac{\log k}{k}})) \cdot \mathcal{P}_{\mathcal{M}_*^k}$, the value $\omega_e(R(\boldsymbol{x}_*))$ concentrates very well as a capped sum over Bernoulli random variables. Therefore, the bad event $\omega_e(R(\boldsymbol{x}_*)) = k$ rarely happens and the simple greedy algorithm without protection works.

[7] When it is clear from context, we will use $\mathrm{rank}(\cdot)/\mathrm{span}(\cdot)$ to denote the rank/span of a set of elements in $\mathcal{M}_*^k$ for the ease of notation.

### 4.3.2 Chain decomposition based on occupancy functions

Similar to the matroid OCRS by [12], our OCRS for extended $k$-fold union $\mathcal{M}_*^k = (E_*, \mathcal{I}_*^k)$ and $\boldsymbol{x}_* \in b \cdot \mathcal{P}_{\mathcal{M}_*^k}$ starts with an offline prepossessing step that finds the following chain decomposition of elements in $E_*$,

$$\emptyset = N_\ell \subsetneq N_{\ell-1} \subsetneq \cdots \subsetneq N_1 \subsetneq N_0 = E_*$$

where $N_{j+1} = \text{KFOLDPROTECT}(\mathcal{M}_*^k|_{N_j}, \boldsymbol{x}_*|_{N_j}, b)$ for every $0 \leq j < \ell$, as described in Algorithm 3. Unlike Algorithm 2, it relies on the occupancy functions which are only defined for extended $k$-fold unions.

**Algorithm 3** Find a protection set $S$ for extended $k$-fold union $\mathcal{M}_*^k = (E_*, \mathcal{I}_*^k)$ and $\boldsymbol{x}_* \in b \cdot \mathcal{P}_{\mathcal{M}_*^k}$.

**function** KFOLDPROTECT($\mathcal{M}_*^k, \boldsymbol{x}_*, b$)
  $S_0, S \leftarrow \emptyset$
  **while** $\exists e \in E \setminus S_0,\ \mathbf{E}[\omega_e(R(\boldsymbol{x}_*) \cup S)] > bk$ **do**
    $S_0 \leftarrow S_0 \cup \{e\}$
    $S \leftarrow S \cup (\{e\} \times [k])$
  **return** $S$

Before introducing our OCRS, we need to make sure the chain decomposition above is well-defined, i.e., Algorithm 3 will always returns a proper subset $S$ of elements, and $\mathcal{M}_*^k|_S$ remains an extended $k$-fold union. This is formally stated in Lemma 18, which resembles Lemma 11 in [12].

▶ **Lemma 18.** *For any $b \in (0, 1)$, any extended $k$-fold union $\mathcal{M}_*^k = (E_*, \mathcal{I}_*^k)$ and $\boldsymbol{x}_* \in b \cdot \mathcal{P}_{\mathcal{M}_*^k}$, $S \subsetneq E_*$ for $S = \text{KFOLDPROTECT}(\mathcal{M}_*^k, \boldsymbol{x}_*, b)$. Moreover, $\mathcal{M}_*^k|_S$ remains an extended $k$-fold union.*

Having obtained such a chain decomposition for $\mathcal{M}_*^k$ and $\boldsymbol{x}_* \in b \cdot \mathcal{P}_{\mathcal{M}_*^k}$, our OCRS is simply running the modified greedy algorithm, Algorithm 1, for each submatroid $\mathcal{M}_*^k|_{N_j}$ with a protection set $N_{j+1}$ for all $0 \leq j < \ell$ together. Note that although the chain decomposition is constructed with $\boldsymbol{x}_*$, an extra scaling factor of $e^{-(1-b)}$ will be applied before invoking Algorithm 1. This will be useful later when we apply the bicriterion concentration inequality.

**Algorithm 4** OCRS for extended $k$-fold union $\mathcal{M}_*^k = (E_*, \mathcal{I}_*^k)$ and $\boldsymbol{x}_* \in b \cdot \mathcal{P}_{\mathcal{M}_*^k}$.

Construct the chain decomposition $\emptyset = N_\ell \subsetneq \cdots \subsetneq N_1 \subsetneq N_0 = E_*$ for $\mathcal{M}_*^k$ and $\boldsymbol{x}_*$
**for** each arriving active element $(e, i) \in E_*$ **do**
  Sample $r \sim \text{Ber}(e^{-(1-b)})$
  **if** $r = 0$ **then**
    Reject $(e, i)$
  **else**
    ▷ *The set of remaining active elements follows the same distribution as $R(e^{-(1-b)}\boldsymbol{x}_*)$.* ◁
    Find $0 \leq j < \ell$ such that $(e, i) \in N_j \setminus N_{i+1}$
    Invoke Algorithm 1 for $\mathcal{M}_*^k|_{N_j}$ with protection set $N_{j+1}$ for $(e, i)$

The feasibility of such a scheme follows exactly from [12] as running Algorithm 1 on any chain decomposition always produces an independent set. We are left to show the OCRS guarantees a good selectability for any $\mathcal{M}_*^k$ and $\boldsymbol{x}_* \in b \cdot \mathcal{P}_{\mathcal{M}_*^k}$ for some parameter $b$. In fact, we will set $b = 1 - \sqrt{\frac{\log k}{k}}$ and show the selectability is at least $1 - O(\sqrt{\frac{\log k}{k}})$, proving Theorem 9.

### 4.3.3 Analyzing the selectability

Without loss of generality, let us focus on the selectability of elements in the first layer $E_* \setminus N_1$, since a same proof would work for all submatroid $\mathcal{M}_*^k|_{N_j}$ as they remains to be extended $k$-fold unions.

By Lemma 16, for every element $(e,i) \in E_* \setminus N_1$, its selectability can be lower bounded as

$$\begin{aligned}\mathbf{Pr}[(e,i) \text{ is accepted} \mid (e,i) \text{ is active}] &\geq \mathbf{Pr}[(e,i) \notin \operatorname{span}((R(e^{-(1-b)}\boldsymbol{x_*}) \setminus \{(e,i)\}) \cup N_1)] \\ &\geq \mathbf{Pr}[\omega_e(R(e^{-(1-b)}\boldsymbol{x_*}) \cup N_1) < k].\end{aligned}$$

On the other hand, by the way chain decomposition is obtained using Algorithm 3, we know even without the extra scaling of $e^{-(1-b)}$, the expected value of $\omega_e(R(\boldsymbol{x_*}) \cup N_1)$ is not too close to $k$:

$$\mathbf{E}[\omega_e(R(\boldsymbol{x_*}) \cup N_1)] \leq bk.$$

For the ease of notation, denote $X = R(\boldsymbol{x_*})$ and $X' = R(e^{-(1-b)}\boldsymbol{x_*})$. Fixing an element $(e,i) \in E_*$, define the function $f : 2^{E_*} \to [0,k]$ where for every $S \subseteq E_*$,

$$f(S) = \omega_e(S \cup N_1).$$

Then, to lower bound selectability for $(e,i)$, it is equivalent to upper bound $\mathbf{Pr}[f(X') = k]$ given that $\mathbf{E}[f(X)] \leq bk$. Specifically, to get a selectability of $1 - O(\sqrt{\frac{\log k}{k}})$, we will set $b = 1 - \sqrt{\frac{\log k}{k}}$, and it suffices to show the following *bicriterion concentration inequality*:

$$\mathbf{E}[f(X)] \leq k - \sqrt{k \log k} \implies \mathbf{Pr}\left[f(X') \geq \mathbf{E}[f(X)] + \sqrt{k \log k}\right] \leq O\left(\frac{1}{k}\right). \qquad (*)$$

By Lemma 15, we know $f$ is always monotone and 1-Lipschitz. Then, using Theorem 3 (and recall that $X' = R(e^{-\sqrt{\log k/k}}\boldsymbol{x_*})$), we have

$$\mathbf{Pr}\left[f(X') \geq \mathbf{E}[f(X)] + \sqrt{k \log k}\right] \leq \exp\left(-\sqrt{\frac{\log k}{k}} \cdot \sqrt{k \log k}\right) = \frac{1}{k}.$$

Therefore, for extended $k$-fold union $\mathcal{M}_*^k$ and $\boldsymbol{x_*} \in b \cdot \mathcal{P}_{\mathcal{M}_*^k}$, running Algorithm 4 yields

$$\begin{aligned}\mathbf{Pr}[(e,i) \text{ is accepted} \mid (e,i) \text{ is active}] &\geq 1 - \mathbf{Pr}\left[f(X') \geq k\right] \\ &\geq 1 - \mathbf{Pr}\left[f(X') \geq \mathbf{E}[f(X)] + \sqrt{k \log k}\right] \geq 1 - \frac{1}{k}.\end{aligned}$$

Together with Fact 10 and Lemma 13, we prove Theorem 9 by showing the existence of an OCRS for all $k$-fold union $\mathcal{M}^k$ and $\boldsymbol{x_*} \in \mathcal{P}_{\mathcal{M}^k}$ with a selectability of

$$\left(1 - \frac{1}{k}\right) \cdot b \cdot e^{-(1-b)} = \left(1 - \frac{1}{k}\right) \cdot \left(1 - \sqrt{\frac{\log k}{k}}\right) \cdot e^{-\sqrt{\frac{\log k}{k}}} = 1 - O\left(\sqrt{\frac{\log k}{k}}\right).$$

▶ Remark 19. It might seems bizarre and unnecessary to consider $f(X')$ instead of $f(X)$. Indeed, since $f$ is monotone non-decreasing, the following claim that only contains $f(X)$ would imply $(*)$, and it looks more like a standard concentration inequality:

$$\mathbf{E}[f(X)] \leq k - \sqrt{k \log k} \implies \mathbf{Pr}\left[f(X) \geq \mathbf{E}[f(X)] + \sqrt{k \log k}\right] \leq O\left(\frac{1}{k}\right). \qquad (**)$$

We know (∗∗) is true when $f$ is a sum over Bernoulli random variables by Chernoff bound, and it is tempting to use more powerful concentration inequalities to prove (∗∗) for general 1-Lipschitz $f$. Unfortunately, such a bound does not exist for general monotone and 1-Lipschitz set functions (see Section 5 for details), and it turns out to be impossible even for the specific $f$ we use here, as Example 29 shown.

## 5 A Bicriterion Concentration Inequality

In this section, we assume the ground set $E = [n]$ and consider a function $f : \{0,1\}^n \to \mathbb{R}$ that satisfies the following properties:[8]

1. (**Monotone**) $f(\boldsymbol{x}) \le f(\boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^n$ where $\boldsymbol{x} \le \boldsymbol{y}$ (element-wise).
2. (**1-Lipschitz**) $|f(\boldsymbol{x}) - f(\boldsymbol{y})| \le \|\boldsymbol{x} - \boldsymbol{y}\|_1$ for all $\boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^n$.

Also, let $\boldsymbol{X} = (X_1, X_2, \dots, X_n)$ be a vector of $n$ independent Bernoulli random variables where $X_i \sim \mathrm{Ber}(p_i)$ for each $i \in [n]$ and $\boldsymbol{p} \in [0,1]^n$. For simplicity, we denote this as $\boldsymbol{X} \sim \mathrm{Ber}(\boldsymbol{p})$.

We are interested in how well $f(\boldsymbol{X})$ concentrates on its upper tail. By McDiarmid's inequality (Theorem 28), for every $t > 0$,

$$\mathbf{Pr}\left[f(\boldsymbol{X}) \ge \mathbf{E}[f(\boldsymbol{X})] + t\right] \le e^{-\frac{2t^2}{n}},$$

Unfortunately, the bound depends on the dimension $n$, whereas our application in Section 4 requires a *dimension-free* bound that is independent from $n$. In fact, it is known that dimension-free concentration inequality does not exist for $f$ in general (see, e.g., [27]).

The good news is that, for our application, it suffices to consider another $\boldsymbol{X'} \sim \mathrm{Ber}(\boldsymbol{p'})$ with slightly smaller parameters $\boldsymbol{p'} < \boldsymbol{p}$ and show $f(\boldsymbol{X'})$ does not exceed $\mathbf{E}[f(\boldsymbol{X})]$ by much, with high probability. Formally, we define $\boldsymbol{X}^{(s)}$ with a scaling factor $s$ as follows:

▶ **Definition 20** (Scaling). *Given $n$ independent Bernoulli random variables $\boldsymbol{X} \sim \mathrm{Ber}(\boldsymbol{p})$, for any* scaling factor *$s \ge 0$, define $\boldsymbol{X}^{(s)} \sim \mathrm{Ber}(e^{-s}\boldsymbol{p})$. In other words, $X_i^{(s)} \sim \mathrm{Ber}(e^{-s}p_i)$ for all $i \in [n]$.*

And we prove Theorem 3, a *bicriterion concentration inequality*, where the bound depends on both the scaling factor $s$ and the deviation size $t$.

▶ **Theorem 3.** *Let $f : \{0,1\}^E \to \mathbb{R}$ be a monotone 1-Lipschitz function. For any $s \in (0,1]$, $t > 0$:*

$$\mathbf{Pr}\left[f(\boldsymbol{X}^{(s)}) \ge \mathbf{E}[f(\boldsymbol{X})] + t\right] \le e^{-st}.$$

For our application in Section 4, we basically set $s = \sqrt{\frac{\log k}{k}}, t = \sqrt{k \log k}$ for some $k \approx \mathbf{E}[f(\boldsymbol{X})]$ and the inequality gives us $\mathbf{Pr}[f(\boldsymbol{X}^{(s)}) \ge k + \sqrt{k \log k}] \le \frac{1}{k}$. Note that this bound is sharp up to a constant factor in the exponent: even in the case where $f(\boldsymbol{x}) = \sum_{i=1}^n x_i$, the Chernoff bound of $f(\boldsymbol{X}^{(s)})$ only yields $\mathbf{Pr}[f(\boldsymbol{X}^{(s)}) \ge k + \sqrt{k \log k}] \le O(\frac{1}{k^c})$ for some constant $c$.

[8] Note that $f$ can be equivalently viewed as a function over subsets of a ground set of size $n$, as we did in Section 4.

## 5.1 Technical overview

Before getting into the proof, let us first outline our approach and highlight the main difficulty. Our proof utilizes the entropy method for self-bounding functions [5, 6, 22, 7]. Roughly speaking, to prove a exponential concentration inequality for some $Z = f(\boldsymbol{X})$, the plan is to establish a differential inequality for the moment-generating function $\mathbf{E}[e^{\lambda Z}]$ based on the following modified logarithmic Sobolev inequality. If this differential inequality implies strong bounds for $\mathbf{E}[e^{\lambda Z}]$, a concentration inequality can be subsequently obtained.

▶ **Lemma 21** (A modified logarithmic Sobolev inequality [20]). *Given $n$ independent Bernoulli random variables $\boldsymbol{X}$ and a function $f : \{0,1\}^n \to \mathbb{R}$. Let $Z = f(\boldsymbol{X})$ and $Z_i = f_i(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$ for an arbitrary function $f_i : \{0,1\}^{n-1} \to \mathbb{R}$. For any $\lambda \in \mathbb{R}$,*

$$\lambda\,\mathbf{E}\left[Ze^{\lambda Z}\right] - \mathbf{E}\left[e^{\lambda Z}\right]\log\mathbf{E}\left[e^{\lambda Z}\right] \le \sum_{i=1}^{n}\mathbf{E}\left[e^{\lambda Z}\phi(-\lambda(Z - Z_i))\right]$$

*where $\phi(x) = e^x - x - 1$.*

Whether Lemma 21 can be effectively converted into a useful differential inequality for $\mathbf{E}[e^{\lambda Z}]$ depends on the choice of $\{Z_i\}_{i\in[n]}$. For a monotone function $f$, a typical choice is $Z_i = f(X_1, \ldots, X_{i-1}, 0, X_{i+1}, \ldots, X_n)$, and previous works have demonstrated that such a conversion is possible if $f$ is 1-Lipschitz and the following condition holds almost surely for some constants $a, b \ge 0$: [9]

$$\sum_{i=1}^{n} Z - Z_i \le aZ + b. \tag{†}$$

Now, given $Z^{(s)} = f(\boldsymbol{X}^{(s)})$ under a scaling factor $s > 0$, one might attempt to similarly derive a differential inequality of $\mathbf{E}[e^{\lambda Z^{(s)}}]$ based on Lemma 21 if the condition (†) can be satisfied. In fact, if we define $Z_i^{(s)} = f(X_1^{(s)}, \ldots, X_{i-1}^{(s)}, 0, X_{i+1}^{(s)}, \ldots, X_n^{(s)})$, the following holds:

$$\mathbf{E}\left[\sum_{i=1}^{n} Z^{(s)} - Z_i^{(s)}\right] = -\frac{\mathrm{d}}{\mathrm{d}s}\,\mathbf{E}\left[Z^{(s)}\right].$$

Thus, if $-\frac{\mathrm{d}}{\mathrm{d}s}\,\mathbf{E}[Z^{(s)}] \le aZ^{(s)} + b$, then (†) holds in expectation for $Z^{(s)}$; otherwise, $E[Z^{(s)}]$ is decreasing rapidly with respect to $s$ at that point.

As a result, either there exists some $s^* \in (0, s)$ such that (†) holds *in expectation* for $Z^{(s^*)}$, or $\mathbf{E}[Z^{(s)}]$ becomes significantly smaller than $\mathbf{E}[Z^{(0)}]$. Intuitively, the latter case should directly imply a bicriterion concentration result, leaving only the former case to be addressed.[10] However, it turns out that such a use of Lemma 21 crucially depends on (†) holding *almost surely*, which is not applicable to such $Z^{(s^*)}$ in the former case.[11]

Given this limitation, rather than working with moment-generating functions directly, we propose an alternative approach. Our key idea is to relate Lemma 21 with the following unconventional function, defined for every $\lambda \ge 0$:

$$F(\lambda) = \mathbf{E}\left[e^{\lambda Z^{(\lambda)}}\right].$$

---

[9] In this case, $f$ is a so-called $(a, b)$-*self-bounding* function [22, 7].

[10] If we do not aim for an exponential tail bound, these observations indeed suffice to get a Chebyshev-type bicriterion concentration inequality for $Z^{(s)}$, by using Efron-Stein inequality to bound its variance in the former case.

[11] Specifically, applying the entropy method for $\mathbf{E}[e^{\lambda Z}]$ requires $\sum_{i=1}^{n}\mathbf{E}[e^{\lambda Z}(Z - Z_i)] \le \mathbf{E}[e^{\lambda Z}(aZ + b)]$ for every $\lambda$, which might be false even if (†) holds with very high probability.

Note that this is not a moment-generating function, as $\lambda$ here also serves as the scaling factor of $Z$, causing the random variable $Z^{(\lambda)}$ to change with it. Surprisingly, we can obtain the following upper bound for the derivative of $F(\lambda)$ that aligns well with Lemma 21.

▶ **Lemma 22.** *Given $n$ independent Bernoulli random variables $\boldsymbol{X}$ and a monotone 1-Lipschitz function $f : \{0,1\}^n \to \mathbb{R}$. For any $\lambda \in (0,1]$,*

$$F'(\lambda) \le \mathbf{E}\left[Z^{(\lambda)} e^{\lambda Z^{(\lambda)}}\right] - \frac{1}{\lambda}\sum_{i=1}^{n} \mathbf{E}\left[e^{\lambda Z^{(\lambda)}} \phi(-\lambda(Z^{(\lambda)} - Z_i^{(\lambda)}))\right]$$

*where $Z^{(\lambda)} = f(\boldsymbol{X}^{(\lambda)})$, $Z_i^{(\lambda)} = f(X_1^{(\lambda)}, \ldots, X_{i-1}^{(\lambda)}, 0, X_{i+1}^{(\lambda)}, \ldots, X_n^{(\lambda)})$, and $F(\lambda) = \mathbf{E}[e^{\lambda Z^{(\lambda)}}]$.*

By combining Lemma 22 with Lemma 21, we can conclude that for all $\lambda \in (0,1]$,

$$\lambda F'(\lambda) \le F(\lambda) \log F(\lambda).$$

Solving this differential inequality provides an upper bound for $F(\lambda)$. Theorem 3 then follows by applying Markov's inequality to the random variable $e^{sZ^{(s)}}$.

## 5.2 Proof of Theorem 3

Let $Z^{(\lambda)} = f(\boldsymbol{X}^{(\lambda)})$ and $Z_i^{(\lambda)} = f(X_1^{(\lambda)}, \ldots, X_{i-1}^{(\lambda)}, 0, X_{i+1}^{(\lambda)}, \ldots, X_n^{(\lambda)})$ throughout the proof. Given Lemma 21 and Lemma 22, it is not hard to show the bicriterion concentration inequality.

**Proof of Theorem 3.** For any $\lambda > 0$, we apply Lemma 21 to $Z^{(\lambda)}$ and $\{Z_i^{(\lambda)}\}_{i\in[n]}$ and obtain

$$\lambda\,\mathbf{E}\left[Z^{(\lambda)} e^{\lambda Z^{(\lambda)}}\right] - \mathbf{E}\left[e^{\lambda Z^{(\lambda)}}\right] \log \mathbf{E}\left[e^{\lambda Z^{(\lambda)}}\right] \le \sum_{i=1}^{n} \mathbf{E}\left[e^{\lambda Z^{(\lambda)}} \phi(-\lambda(Z^{\lambda} - Z_i^{(\lambda)}))\right]$$

where $\phi(x) = e^x - x - 1$. Rearranging the inequality, we have

$$\lambda\left(\mathbf{E}\left[Z^{(\lambda)} e^{\lambda Z^{(\lambda)}}\right] - \frac{1}{\lambda}\sum_{i=1}^{n} \mathbf{E}\left[e^{\lambda Z^{(\lambda)}} \phi(-\lambda(Z^{\lambda} - Z_i^{(\lambda)}))\right]\right) \le \mathbf{E}\left[e^{\lambda Z^{(\lambda)}}\right] \log \mathbf{E}\left[e^{\lambda Z^{(\lambda)}}\right].$$

Together with Lemma 22, this gives us the following differential inequality for $F(\lambda) = \mathbf{E}[e^{\lambda Z^{(\lambda)}}]$:

$$\lambda F'(\lambda) \le F(\lambda) \log F(\lambda), \quad \forall \lambda \in (0,1].$$

And by letting $G(\lambda) = \log F(\lambda)$, we can rewrite the inequality as

$$\lambda G'(\lambda) \le G(\lambda), \quad \forall \lambda \in (0,1].$$

Note that $G_0(\lambda) = \lambda\,\mathbf{E}[Z^{(0)}]$ is a solution to $\lambda G'(\lambda) = G(\lambda)$ for $\lambda \in (0,1]$. Define $g(\lambda) = \frac{G(\lambda) - G_0(\lambda)}{\lambda}$ and we have

$$g'(\lambda) = \frac{G'(\lambda) - G_0'(\lambda)}{\lambda} - \frac{G(\lambda) - G_0(\lambda)}{\lambda^2} = \frac{(\lambda G'(\lambda) - G(\lambda)) - (\lambda G_0'(\lambda) - G_0(\lambda))}{\lambda^2} \le 0.$$

Also note that $\lim_{\lambda\to 0^+} \frac{G(\lambda)}{\lambda} = G'(0) = \frac{F'(0)}{F(0)} = \mathbf{E}[Z^{(0)}]$ (where the last equality holds by Lemma 23), therefore $\lim_{\lambda\to 0^+} g(\lambda) = 0$. Combining this with $g' \le 0$, we conclude that $g$ is non-positive on $(0,1]$. In other words, for $\lambda \in (0,1]$,

$$G(\lambda) \le G_0(\lambda) = \lambda\,\mathbf{E}[Z^{(0)}].$$

Finally, by Markov's inequality, we conclude that for any $\lambda \in (0,1]$ and $t > 0$,

$$\mathbf{Pr}\left[Z^{(\lambda)} \geq \mathbf{E}[Z^{(0)}] + t\right] = \mathbf{Pr}\left[e^{\lambda Z^{(\lambda)}} \geq e^{\lambda(\mathbf{E}[Z^{(0)}]+t)}\right] \leq \frac{\mathbf{E}[e^{\lambda Z^{(\lambda)}}]}{e^{\lambda(\mathbf{E}[Z^{(0)}]+t)}} \leq e^{-\lambda t}. \qquad \blacktriangleleft$$

We are left to prove Lemma 22. Let us first compute $F'(\lambda)$ by definition.

▶ **Lemma 23.** *For any $\lambda \geq 0$, $F'(\lambda) = \mathbf{E}[Z^{(\lambda)} e^{\lambda Z^{(\lambda)}}] - \sum_{i=1}^n \mathbf{E}[e^{\lambda Z^{(\lambda)}} - e^{\lambda Z_i^{(\lambda)}}]$.*

**Proof.** Define a function $h : \mathbb{R} \times [0,1]^n \to \mathbb{R}$ as

$$h(t, \boldsymbol{q}) = \mathop{\mathbf{E}}_{\boldsymbol{Y} \sim \mathrm{Ber}(\boldsymbol{q})}\left[e^{tf(\boldsymbol{Y})}\right].$$

For each $i \in [n]$ and $b \in \{0,1\}$, denote $f_{i,b}(\boldsymbol{Y}) = f(Y_1, \ldots, Y_{i-1}, b, Y_{i+1}, \ldots, Y_n)$ and we can compute the partial derivative of $h$ with respect to $q_i$ as

$$\begin{aligned}\frac{\partial}{\partial q_i} h(t, \boldsymbol{q}) &= \frac{\partial}{\partial q_i}\left(q_i \mathop{\mathbf{E}}_{\boldsymbol{Y} \sim \mathrm{Ber}(\boldsymbol{q})}[e^{tf_{i,1}(\boldsymbol{Y})}] + (1 - q_i) \mathop{\mathbf{E}}_{\boldsymbol{Y} \sim \mathrm{Ber}(\boldsymbol{q})}[e^{tf_{i,0}(\boldsymbol{Y})}]\right) \\ &= \mathop{\mathbf{E}}_{\boldsymbol{Y} \sim \mathrm{Ber}(\boldsymbol{q})}[e^{tf_{i,1}(\boldsymbol{Y})} - e^{tf_{i,0}(\boldsymbol{Y})}].\end{aligned}$$

Recall that $\boldsymbol{X} \sim \mathrm{Ber}(\boldsymbol{p})$ and $F(\lambda) = \mathbf{E}[e^{\lambda Z^{(\lambda)}}] = h(\lambda, e^{-\lambda}\boldsymbol{p})$. Therefore,

$$\begin{aligned}F'(\lambda) &= \frac{\mathrm{d}t}{\mathrm{d}\lambda} \cdot \frac{\partial}{\partial t} h(\lambda, e^{-\lambda}\boldsymbol{p}) + \sum_{i=1}^n \frac{\mathrm{d}q_i}{\mathrm{d}\lambda} \cdot \frac{\partial}{\partial q_i} h(\lambda, e^{-\lambda}\boldsymbol{p}) \\ &= 1 \cdot \mathbf{E}\left[f(\boldsymbol{X}^{(\lambda)}) e^{\lambda f(\boldsymbol{X}^{(\lambda)})}\right] + \sum_{i=1}^n (-e^{-\lambda} p_i) \cdot \mathop{\mathbf{E}}_{\boldsymbol{Y} \sim \mathrm{Ber}(e^{-\lambda}\boldsymbol{p})}\left[e^{\lambda f_{i,1}(\boldsymbol{Y})} - e^{\lambda f_{i,0}(\boldsymbol{Y})}\right] \\ &= \mathbf{E}\left[f(\boldsymbol{X}^{(\lambda)}) e^{\lambda f(\boldsymbol{X}^{(\lambda)})}\right] - \sum_{i=1}^n \mathop{\mathbf{E}}_{\boldsymbol{Y} \sim \mathrm{Ber}(e^{-\lambda}\boldsymbol{p})}\left[e^{\lambda f(\boldsymbol{Y})} - e^{\lambda f_{i,0}(\boldsymbol{Y})}\right] \\ &= \mathbf{E}\left[Z^{(\lambda)} e^{\lambda Z^{(\lambda)}}\right] - \sum_{i=1}^n \mathbf{E}\left[e^{\lambda Z^{(\lambda)}} - e^{\lambda Z_i^{(\lambda)}}\right]. \qquad \blacktriangleleft\end{aligned}$$

Then we further derive a lower bound to the latter term, $\sum_{i=1}^n \mathbf{E}[e^{\lambda Z^{(\lambda)}} - e^{\lambda Z_i^{(\lambda)}}]$.

▶ **Lemma 24.** *For any $\lambda \geq 0$, $\sum_{i=1}^n \mathbf{E}[e^{\lambda Z^{(\lambda)}} - e^{\lambda Z_i^{(\lambda)}}] \geq \lambda e^{-\lambda} \sum_{i=1}^n \mathbf{E}[e^{\lambda Z^{(\lambda)}}(Z^{(\lambda)} - Z_i^{(\lambda)})]$.*

**Proof.** We prove the inequality for each term separately and without expectation. For any $i \in [n]$, note that

$$e^{\lambda Z^{(\lambda)}} - e^{\lambda Z_i^{(\lambda)}} = e^{\lambda Z_i^{(\lambda)}}(e^{\lambda(Z^{(\lambda)} - Z_i^{(\lambda)})} - 1) \geq e^{\lambda Z_i^{(\lambda)}} \cdot \lambda(Z^{(\lambda)} - Z_i^{(\lambda)})$$

since $e^x - 1 \geq x$. Meanwhile, we know $Z_i^{(\lambda)} \geq Z^{(\lambda)} - 1$ by 1-Lipschitzness of $f$. Therefore,

$$e^{\lambda Z^{(\lambda)}} - e^{\lambda Z_i^{(\lambda)}} \geq \lambda e^{-\lambda} \cdot e^{\lambda Z^{(\lambda)}}(Z^{(\lambda)} - Z_i^{(\lambda)}). \qquad \blacktriangleleft$$

The following two facts of the function $\phi(x) = e^x - x - 1$ will also be used.

▶ **Fact 25.** *For any $\lambda \in (0,1]$, $\frac{\phi(-\lambda)}{\lambda} \leq \lambda e^{-\lambda}$.*

▶ **Fact 26.** *For any $\lambda \in \mathbb{R}$ and $x \in [0,1]$, $\phi(-\lambda x) \leq \phi(-\lambda) x$.*

Now we are ready to prove the lemma.

**Proof of Lemma 22.** We upper bound $F'(\lambda)$ step-by-step as follows:

$$\begin{aligned}
F'(\lambda) &= \mathbf{E}[Z^{(\lambda)} e^{\lambda Z^{(\lambda)}}] - \sum_{i=1}^{n} \mathbf{E}\left[e^{\lambda Z^{(\lambda)}} - e^{\lambda Z_i^{(\lambda)}}\right] && \text{(Lemma 23)}\\
&\le \mathbf{E}[Z^{(\lambda)} e^{\lambda Z^{(\lambda)}}] - \lambda e^{-\lambda} \sum_{i=1}^{n} \mathbf{E}\left[e^{\lambda Z^{(\lambda)}} (Z^{(\lambda)} - Z_i^{(\lambda)})\right] && \text{(Lemma 24)}\\
&\le \mathbf{E}[Z^{(\lambda)} e^{\lambda Z^{(\lambda)}}] - \frac{\phi(-\lambda)}{\lambda} \sum_{i=1}^{n} \mathbf{E}\left[e^{\lambda Z^{(\lambda)}} (Z^{(\lambda)} - Z_i^{(\lambda)})\right] && \text{(Fact 25)}\\
&\le \mathbf{E}[Z^{(\lambda)} e^{\lambda Z^{(\lambda)}}] - \frac{1}{\lambda} \sum_{i=1}^{n} \mathbf{E}\left[e^{\lambda Z^{(\lambda)}} \phi(-\lambda(Z^{(\lambda)} - Z_i^{(\lambda)}))\right] && \text{(Fact 26)}
\end{aligned}$$

where in the last step we also use the fact that $Z^{(\lambda)} - Z_i^{(\lambda)} \in [0, 1]$, as $f$ is monotone and 1-Lipschitz. ◀

## References


1 Saeed Alaei. Bayesian combinatorial auctions: Expanding single buyer mechanisms to many buyers. *SIAM Journal on Computing*, 43(2):930–972, 2014. doi:10.1137/120878422.

2 Saeed Alaei, Hu Fu, Nima Haghpanah, Jason Hartline, and Azarakhsh Malekian. Bayesian optimal auctions via multi- to single-agent reduction. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, EC '12, page 17, New York, NY, USA, 2012. Association for Computing Machinery. doi:10.1145/2229012.2229017.

3 C. J. Argue, Anupam Gupta, Marco Molinaro, and Sahil Singla. Robust secretary and prophet algorithms for packing integer programs. In Joseph (Seffi) Naor and Niv Buchbinder, editors, *Proceedings of the 2022 ACM-SIAM Symposium on Discrete Algorithms, SODA 2022, Virtual Conference / Alexandria, VA, USA, January 9 - 12, 2022*, pages 1273–1297. SIAM, SIAM, 2022. doi:10.1137/1.9781611977073.53.

4 Pablo Daniel Azar, Robert Kleinberg, and S. Matthew Weinberg. Prophet inequalities with limited information. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 1358–1377, 2014. doi:10.1137/1.9781611973402.100.

5 Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. A sharp concentration inequality with applications. *Random Structures & Algorithms*, 16(3):277–292, 2000. doi:10.1002/(SICI)1098-2418(200005)16:3\%3C277::AID-RSA4\%3E3.0.CO;2-1.

6 Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. Concentration inequalities using the entropy method. *The Annals of Probability*, 31(3):1583–1614, 2003.

7 Stephane Boucheron, Gabor Lugosi, and Pascal Massart. On concentration of self-bounding functions. *Electronic Journal of Probability*, 14:1884–1899, 2009.

8 Shuchi Chawla, Jason D. Hartline, David L. Malec, and Balasubramanian Sivan. Multi-parameter mechanism design and sequential posted pricing. In Leonard J. Schulman, editor, *Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC 2010, Cambridge, Massachusetts, USA, 5-8 June 2010*, pages 311–320. ACM, 2010. doi:10.1145/1806689.1806733.

9 Chandra Chekuri, Junkai Song, and Weizhong Zhang. Contention resolution for the l-fold union of a matroid via the correlation gap. In *2024 Symposium on Simplicity in Algorithms (SOSA)*, pages 396–405. SIAM, 2024. doi:10.1137/1.9781611977936.36.

10 Atanas Dinev and S. Matthew Weinberg. Simple and Optimal Online Contention Resolution Schemes for k-Uniform Matroids. In Venkatesan Guruswami, editor, *15th Innovations in Theoretical Computer Science Conference (ITCS 2024)*, volume 287 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 39:1–39:23, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ITCS.2024.39.

11 Paul Dütting, Michal Feldman, Thomas Kesselheim, and Brendan Lucier. Prophet inequalities made easy: Stochastic optimization by pricing nonstochastic inputs. *SIAM J. Comput.*, 49(3):540–582, 2020. doi:10.1137/20M1323850.

12 Moran Feldman, Ola Svensson, and Rico Zenklusen. Online contention resolution schemes. In Robert Krauthgamer, editor, *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016*, pages 1014–1033. SIAM, 2016. doi:10.1137/1.9781611974331.ch72.

13 Mohammad Taghi Hajiaghayi, Robert Kleinberg, and Tuomas Sandholm. Automated online mechanism design and prophet inequalities. In *AAAI*, volume 7, pages 58–65, 2007. URL: http://www.aaai.org/Library/AAAI/2007/aaai07-009.php.

14 Jiashuo Jiang, Will Ma, and Jiawei Zhang. Tight guarantees for multi-unit prophet inequalities and online stochastic knapsack. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1221–1246. SIAM, 2022. doi:10.1137/1.9781611977073.51.

15 Thomas Kesselheim, Andreas Tönnis, Klaus Radke, and Berthold Vöcking. Primal beats dual on online packing lps in the random-order model. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 303–312, 2014. doi:10.1145/2591796.2591810.

16 Robert Kleinberg and S. Matthew Weinberg. Matroid prophet inequalities. In *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 123–136, 2012. doi:10.1145/2213977.2213991.

17 Ulrich Krengel and Louis Sucheston. On semiamarts, amarts, and processes with finite value. *Advances in Probability and Related Topics*, 4:197–266, 1978.

18 Felix Lazebnik, Vasiliy A Ustimenko, and Andrew J Woldar. A new series of dense graphs of high girth. *Bulletin of the American mathematical society*, 32(1):73–79, 1995.

19 Euiwoong Lee and Sahil Singla. Optimal Online Contention Resolution Schemes via Ex-Ante Prophet Inequalities. In *26th Annual European Symposium on Algorithms (ESA 2018)*, volume 112, pages 57:1–57:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPIcs.ESA.2018.57.

20 Pascal Massart. About the constants in talagrand's concentration inequalities for empirical processes. *The Annals of Probability*, 28(2):863–884, 2000.

21 Colin McDiarmid et al. On the method of bounded differences. *Surveys in combinatorics*, 141(1):148–188, 1989.

22 Colin McDiarmid and Bruce Reed. Concentration for self-bounding functions and an inequality of talagrand. *Random Structures & Algorithms*, 29(4):549–557, 2006. doi:10.1002/rsa.20145.

23 Aviad Rubinstein and Sahil Singla. Combinatorial prophet inequalities. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, Hotel Porta Fira, January 16-19*, pages 1671–1687, 2017. doi:10.1137/1.9781611974782.110.

24 Aviad Rubinstein and S. Matthew Weinberg. Simple mechanisms for a subadditive buyer and applications to revenue monotonicity. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation, EC '15, Portland, OR, USA, June 15-19, 2015*, pages 377–394, 2015. doi:10.1145/2764468.2764510.

25 Ester Samuel-Cahn. Comparison of threshold stop rules and maximum for independent nonnegative random variables. *Annals of Probability*, 12(4):1213–1216, 1984.

26 Gideon Schechtman. Concentration, results and applications. In *Handbook of the geometry of Banach spaces*, volume 2, pages 1603–1634. Elsevier, 2003.

27 Jan Vondrák. A note on concentration of submodular functions. *CoRR*, abs/1005.2791, 2010. doi:10.48550/arXiv.1005.2791.

# A Useful Concentration Inequalities

▶ **Theorem 27** (Multiplicative Chernoff bound). *Given $n$ independent Bernoulli random variables $X_1, X_2, \ldots, X_n$, let $X = \sum_{i=1}^n X_i$ denote their sum. For any $\delta > 0$, we have*

$$\mathbf{Pr}[X \geq (1+\delta)\,\mathbf{E}[X]] \leq \exp\left(-\frac{\delta^2\,\mathbf{E}[X]}{2+\delta}\right).$$

▶ **Theorem 28** (McDiarmid's inequality). *Given $n$ independent random variables $X_1, X_2, \ldots, X_n \in \mathcal{X}$ and a function $f : \mathcal{X}^n \to \mathbb{R}$. If for every $i \in [n]$ and $x_1, x_2, \ldots, x_n, x_i' \in \mathcal{X}$, the function $f$ satisfies*

$$|f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n)| \leq c_i,$$

*then for any $t > 0$, we have*

$$\mathbf{Pr}[f(X) \geq \mathbf{E}[f(X)] + t] \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n c_i^2}\right).$$

# B Missing Proofs and Examples

**Proof of Lemma 13.** It is straightforward to check $\mathcal{M}_*$ in Definition 12 is a matroid, and hence its $k$-fold union $\mathcal{M}_*^k$ remains a matroid by the closure property of matroid union. Also, note that the restriction of $\mathcal{M}_*^k$ to $E \times \{1\}$, $\mathcal{M}_*^k\,|_{E\times\{1\}}$, is isomorphic to the $k$-fold union $\mathcal{M}^k$ of $\mathcal{M}$, as there exists a simple bijection $(e,1) \mapsto e$ between $E \times \{1\}, \mathcal{I}_*|_{E\times\{1\}}$ and $E, \mathcal{I}$. Therefore, an $\alpha$-selectable OCRS for $\mathcal{M}_*^k$ can also be used as an $\alpha$-selectable OCRS for $\mathcal{M}^k$. ◀

**Proof of Lemma 15.** Note that the rank function for any matroid is a submodular function. Therefore, $\mathrm{rank}(S \cup (\{e\} \times [k])) - \mathrm{rank}(S) \geq \mathrm{rank}(T \cup (\{e\} \times [k])) - \mathrm{rank}(T)$ for every $S \subseteq T$ by a simple induction, and thus $\omega_e(\cdot)$ is monotone.

Also, we know the rank function is monotone, and the rank of a set can increase by at most 1 after adding an element. Therefore, $\mathrm{rank}(S \cup \{a\}) \geq \mathrm{rank}(S)$ and $\mathrm{rank}(S \cup \{a\} \cup (\{e\} \times [k])) \leq \mathrm{rank}(S \cup (\{e\} \times [k])) + 1$ for every $a \in E_*$ and $S \subseteq E_*$. As a result, $\omega_e(\cdot)$ is 1-Lipschitz. ◀

**Proof of Lemma 16.** When $\omega_e(S) < k$, we have $\mathrm{rank}(S \cup (\{e\} \times [k])) > \mathrm{rank}(S)$ and there exists (at least) one element $(e,j) \in \{e\} \times [k]$ such that $(e,j) \notin \mathrm{span}(S)$. By definition of extended $k$-fold union, it further implies $(e,i) \notin \mathrm{span}(S \setminus \{(e,i)\})$. ◀

**Proof of Lemma 18.** It is straightforward to see $S = S_0 \times [k]$ when the algorithm terminates, and thus $\mathcal{M}_*^k|_{S_0\times[k]}$ is the extended $k$-fold union of $\mathcal{M}|_{S_0}$ by definition. It remains to prove $S \subsetneq E_*$. Since $S$ must be a subset of the universe $E_*$, it suffices to show $S \neq E_*$. Our plan is to show that $S$ is not full rank in $\mathcal{M}_*^k$, even after combined with all active elements $R(\boldsymbol{x_*})$ and take the expectation, i.e.,

$$\mathbf{E}[\mathrm{rank}_{\mathcal{M}_*^k}(R(\boldsymbol{x_*}) \cup S)] < \mathrm{rank}_{\mathcal{M}_*^k}(E_*).$$

This would directly imply $S \neq E_*$ by the monotonicity of the rank function.

Denote $r = \mathrm{rank}_{\mathcal{M}}(S_0)$. Let $e_1, e_2, \ldots, e_r \in S_0$ be the elements from $\mathcal{M}$ that increase the rank of $S_0$ in $\mathcal{M}$ during the execution of Algorithm 3, and denote $e_i$ (for $1 \leq i \leq r$) as the specific element that increases $\mathrm{rank}_{\mathcal{M}}(S_0)$ from $i-1$ to $i$. By definition, $\mathrm{span}_{\mathcal{M}}(\{e_1, e_2, \ldots, e_r\}) = S_0$. In fact, we also have

$$\mathrm{span}_{\mathcal{M}_*^k}(\{e_1, e_2 \ldots, e_r\} \times [k]) = S.$$

This is because $\{e_1, e_2, \ldots, e_r\} \times [k] \subseteq S$ is an independent set of size $kr$ in $\mathcal{M}_*^k$ by definition of the extended $k$-fold union, and we can further show it is a basis of $S$. Suppose it is not, then there must be another independent set $T \subseteq S$ of size larger than $kr$. Since one can partition $T$ into $k$ disjoint independent sets $T_1, T_2, \ldots, T_k$ in $\mathcal{M}_*$ where $\sum_{j\in[k]} |T_j| > kr$, we know there exists some $T_j$ of size larger than $r$, which leads to a contradiction as $T_j \subseteq S_0 \times [k]$ and $\mathrm{rank}_{\mathcal{M}_*}(S_0 \times [k]) = r$.

Also, by the way the algorithm picks elements to be added to $S_0$, for every $1 \leq i \leq r$ we have

$$\mathbf{E}[\omega_{e_i}(R(\boldsymbol{x_*}) \cup (\{e_1, e_2, \ldots, e_{i-1}\} \times [k]))] > bk.$$

Equivalently, we have

$$\mathbf{E}[\mathrm{rank}_{\mathcal{M}_*^k}(R(\boldsymbol{x_*}) \cup (\{e_1, e_2, \ldots, e_i\} \times [k])) - \mathrm{rank}_{\mathcal{M}_*^k}(R(\boldsymbol{x_*}) \cup (\{e_1, e_2, \ldots, e_{i-1}\} \times [k]))] < (1-b)k.$$

Together with these observations, we can upper bound $\mathbf{E}[\mathrm{rank}_{\mathcal{M}_*^k}(R(\boldsymbol{x_*}) \cup S)]$ by a telescoping sum as follows:

$$\begin{aligned}
\mathbf{E}[\mathrm{rank}_{\mathcal{M}_*^k}(R(\boldsymbol{x_*}) \cup S)] &= \mathbf{E}[\mathrm{rank}_{\mathcal{M}_*^k}(R(\boldsymbol{x_*}) \cup \{e_1, e_2, \ldots, e_r\} \times [k])] \\
&= \mathbf{E}[\mathrm{rank}_{\mathcal{M}_*^k}(R(\boldsymbol{x_*}))] + \sum_{i=1}^{r} \mathbf{E}[\mathrm{rank}_{\mathcal{M}_*^k}(R(\boldsymbol{x_*}) \cup (\{e_1, e_2, \ldots, e_i\} \times [k])) \\
&\qquad\qquad - \mathrm{rank}_{\mathcal{M}_*^k}(R(\boldsymbol{x_*}) \cup (\{e_1, e_2, \ldots, e_{i-1}\}) \times [k])] \\
&< \mathbf{E}[\mathrm{rank}_{\mathcal{M}_*^k}(R(\boldsymbol{x_*}))] + (1-b)kr.
\end{aligned}$$

The former term $\mathbf{E}[\mathrm{rank}_{\mathcal{M}_*^k}(R(\boldsymbol{x_*}))]$ can be trivially upper bounded by $\mathbf{E}[|R(\boldsymbol{x_*})|]$ and further by $b\,\mathrm{rank}_{\mathcal{M}_*^k}(E_*)$ due to $\boldsymbol{x_*} \in b \cdot \mathcal{P}_{\mathcal{M}_*^k}$. For the latter term involving $kr$, we already know $kr = \mathrm{rank}_{\mathcal{M}_*^k}(S) \leq \mathrm{rank}_{\mathcal{M}_*^k}(E_*)$. In conclusion, we have

$$\mathbf{E}[\mathrm{rank}_{\mathcal{M}_*^k}(R(\boldsymbol{x_*}) \cup S)] < b\,\mathrm{rank}_{\mathcal{M}_*^k}(E_*) + (1-b)\,\mathrm{rank}_{\mathcal{M}_*^k}(E_*) = \mathrm{rank}_{\mathcal{M}_*^k}(E_*). \qquad \blacktriangleleft$$

▶ **Example 29** (A counterexample to $(**)$)**.** Fix parameters $n, k$ where $n \gg k$, and consider the case when $\mathcal{M}$ is an $n$-uniform matroid of size $2n$. Its extended $k$-fold union $\mathcal{M}_*^k$ is a $kn$-uniform matroid of size $2kn$. Similar to Example 17, for every $e \in E$ and $S \subseteq E_*$ we can derive

$$\omega_e(S) = \begin{cases} 0, & |S| \leq kn - k \\ |S| - (kn - k), & kn - k < |S| < kn \\ k, & |S| \geq kn. \end{cases}$$

Since no protection is needed for uniform matroids, let $f(\cdot) = \omega_e(\cdot)$ for some fixed $e \in E$. When $\boldsymbol{x_*} = (\frac{1}{2} - \frac{1}{2n}) \cdot \mathbf{1}_{E_*}$ (namely, every element in $E_*$ is active with probability $\frac{1}{2} - \frac{1}{2n}$), $|X|$ will follow a binomial distribution with $kn - k$ as both its mean and median. As a result,

$$\mathbf{E}[f(X)] \leq k\,\mathbf{Pr}[f(X) > 0] = k\,\mathbf{Pr}\left[|X| > kn - k\right] \leq \frac{k}{2},$$

$$\text{while} \quad \mathbf{Pr}[f(X) \geq k] = \mathbf{Pr}\left[|X| \geq kn\right] \geq \Omega(1), \qquad (n \gg k)$$

which is a counterexample to the claim $(**)$.

Note that this is not an actual counterexample to Algorithm 4 (even without the extra scaling) since $\boldsymbol{x_*} \notin (1 - O(\sqrt{\frac{\log k}{k}})) \cdot \mathcal{P}_{\mathcal{M}_*^k}$. But it shows that the condition $\mathbf{E}[f(X)] \leq k - O(\sqrt{k \log k})$ alone is not enough to derive a good enough upper bound for $\mathbf{Pr}[f(X) \geq k]$, and it is crucial to also rely on the scaling applied to $\boldsymbol{x_*}$.