On the Correlation Gap of Matroids

Edin Husić[1], Zhuan Khye Koh[2](✉), Georg Loho[3], and László A. Végh[4]

[1] IDSIA, USI-SUPSI, Lugano, Switzerland
edin.husic@supsi.ch
[2] Centrum Wiskunde & Informatica, Amsterdam, The Netherlands
zhuan.koh@cwi.nl
[3] University of Twente, Enschede, The Netherlands
g.loho@utwente.nl
[4] London School of Economics and Political Science, London, UK
l.vegh@lse.ac.uk

Abstract. A set function can be extended to the unit cube in various ways; the correlation gap measures the ratio between two natural extensions. This quantity has been identified as the performance guarantee in a range of approximation algorithms and mechanism design settings. It is known that the correlation gap of a monotone submodular function is at least $1 - 1/e$, and this is tight for simple matroid rank functions.

We initiate a fine-grained study of the correlation gap of matroid rank functions. In particular, we present an improved lower bound on the correlation gap as parametrized by the rank and girth of the matroid. We also show that for any matroid, the correlation gap of its weighted rank function is minimized under uniform weights. Such improved lower bounds have direct applications for submodular maximization under matroid constraints, mechanism design, and contention resolution schemes.

1 Introduction

A continuous function $h\colon [0,1]^E \to \mathbb{R}_+$ is an *extension* of a set function $f\colon 2^E \to \mathbb{R}_+$ if for every $x \in [0,1]^E$, $h(x) = \mathbb{E}_\lambda[f(S)]$ where λ is a probability distribution over 2^E with marginals x, i.e., $\sum_{S:i\in S} \lambda_S = x_i$ for all $i \in E$. Note that this in particular implies $f(S) = h(\chi_S)$ for every $S \subseteq E$, where χ_S denotes the 0-1 indicator vector of S.

Two natural extensions are the following. The first one corresponds to sampling each $i \in E$ independently with probability x_i, i.e., $\lambda_S = \prod_{i\in S} x_i \prod_{i\notin S}(1 - x_i)$. Thus,

$$F(x) \coloneqq \sum_{S\subseteq E} f(S) \prod_{i\in S} x_i \prod_{i\notin S}(1 - x_i) \ . \tag{1}$$

This is an extended abstract. The full version of the paper with all proofs is available on arXiv:2209.09896. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 757481–ScaleOpt).

Z. K. Koh—This work was done while the author was at the London School of Economics.

A. Del Pia and V. Kaibel (Eds.): IPCO 2023, LNCS 13904, pp. 203–216, 2023.
https://doi.org/10.1007/978-3-031-32726-1_15

This is known as the *multilinear extension* in the context of submodular optimization, see [8]. The second extension corresponds to the probability distribution with maximum expectation:

$$\hat{f}(x) := \max_{\lambda} \left\{ \sum_{S\subseteq E} \lambda_S f(S) : \sum_{S\subseteq E: i\in S} \lambda_S = x_i \, \forall i \in E, \sum_{S\subseteq E} \lambda_S = 1, \lambda \geq \mathbb{0} \right\}. \quad (2)$$

Equivalently, $\hat{f}(x)$ is the upper part of the convex hull of the graph of f; we call it the *concave extension* following terminology of discrete convex analysis [20].

Agrawal, Ding, Saberi and Ye [2] introduced the *correlation gap* as the worst case ratio

$$\mathcal{CG}(f) := \min_{x\in[0,1]^E} \frac{F(x)}{\hat{f}(x)} \ . \quad (3)$$

It bounds the maximum loss incurred in the expected value of f by ignoring correlations. This quantity plays a fundamental role in stochastic optimization [2, 22], mechanism design [7,18,28], prophet inequalities [11,15,24], and a variety of submodular optimization problems [3,12].

The focus of this paper is on weighted matroid rank functions. For a matroid $\mathcal{M} = (E, \mathcal{I})$ and a weight vector $w \in \mathbb{R}^E_+$, the corresponding *weighted matroid rank function* is given by

$$r_w(S) := \max \{w(T) : T \subseteq S, T \in \mathcal{I}\} .$$

It is monotone nondecreasing and submodular. Recall that a function $f : 2^E \to \mathbb{R}$ is *monotone* if $f(X) \leq f(Y)$ for all $X \subseteq Y \subseteq E$, and *submodular* if $f(X) + f(Y) \geq f(X \cap Y) + f(X \cup Y)$ for all $X, Y \subseteq E$.

The correlation gap of a weighted matroid rank function has been identified as the performance guarantee in a range of approximation algorithms and mechanism design settings:

Monotone Submodular Maximization. Calinescu et al. [8] considered the problem of maximizing a sum of weighted matroid rank functions $\sum_{i=1}^m f_i$ subject to a matroid constraint. Using an LP relaxation and pipage rounding [1], they gave a $(1 - 1/e)$-approximation algorithm. This was extended by Shioura [26] to the problem of maximizing a sum of monotone $M^\natural$-concave functions [19]. In [9], a $(1 - 1/e)$-approximation algorithm was obtained for maximizing an arbitrary monotone submodular function subject to a matroid constraint.

A fundamental special case of this model is the *maximum coverage* problem. Given m subsets $E_i \subseteq E$, the corresponding *coverage function* is defined as $f(S) = |\{i \in [m] : E_i \cap S \neq \emptyset\}|$. Note that this is a special case of maximizing a sum of matroid rank functions: $f(S) = \sum_{i=1}^m r_i(S)$ where $r_i(S)$ is the rank function of a rank-1 uniform matroid with support E_i. Even for maximization under a cardinality constraint, there is no better than $(1 - 1/e)$-approximation for this problem unless $P = NP$ (see Feige [16]).

Recently, tight approximations have been established for the special case when the function values $f_i(S)$ are determined by the cardinality of the set S.

Barman et al. [5] studied the *maximum concave coverage* problem: given a monotone concave function $\varphi\colon \mathbb{Z}_+ \to \mathbb{R}_+$ and weights $w \in \mathbb{R}^m_+$, the submodular function is defined as $f(S) = \sum_{i=1}^m w_i \varphi(|S \cap E_i|)$.[1] The maximum coverage problem corresponds to $\varphi(x) = \min\{1, x\}$; on the other extreme, for $\varphi(x) = x$ we get the trivial problem $f(S) = \sum_{j\in S} |\{i \in [m] : j \in E_i\}|$. In [5], they present a tight approximation guarantee for maximizing such an objective subject to a matroid constraint, parametrized by the *Poisson curvature* of the function φ.

This generalizes previous work by Barman et al. [6] which considered $\varphi(x) = \min\{\ell, x\}$ (for $\ell > 1$), motivated by the list decoding problem in coding theory. It also generalizes the work by Dudycz et al. [14] which considered geometrically dominated concave functions φ, motivated by approval voting rules such as Thiele rules, proportional approval voting, and p-geometric rules. In both cases, the obtained approximation guarantees improve over the $1 - 1/e$ factor.

In the full version, we make the observation that the algorithm of Calinescu et al. [8] and Shioura [26] actually has an approximation ratio of $\min_{i\in[m]} \mathcal{CG}(f_i)$. We also prove that the Poisson curvature of φ is equal to the correlation gap of the functions $\varphi(|S \cap E_i|)$. Hence, the approximation guarantees in [5,6,14] are in fact correlation gap bounds, and they can be obtained via a single unified algorithm, i.e., the one by Calinescu et al. [8] and Shioura [26]. In particular, the result of Barman et al. [6] which concerned $\varphi(x) = \min\{\ell, x\}$ (for $\ell > 1$) boils down to the analysis of uniform matroid correlation gaps.

Sequential Posted-Price Mechanisms. Following Yan [28], consider a seller with a set of identical services (or goods), and a set E of unit-demand agents. Each agent $i \in E$ has a private valuation v_i for winning the service, and 0 otherwise, where v_i is drawn independently from a known distribution F_i with positive smooth density function over $[0, L]$ for some large L. The seller can only service certain subsets of the agents simultaneously; this is captured by a matroid $\mathcal{M} = (E, \mathcal{I})$ where $\mathcal{I}$ represents the feasible subsets.

Mechanisms like Myerson's mechanism [21] or the VCG mechanism [13,17,27] have optimal revenue or welfare guarantees, but suffer from complicated formats [4] or high computational overhead [23]. Hence, simple mechanisms are often favoured in practice, such as sequential posted-price mechanisms (SPM), in which the seller makes take-it-or-leave-it price offers to agents one by one. Yan [28] showed that the greedy SPM of Chawla et al. [10] achieves an approximation ratio of $\inf_{w\in\mathbb{R}^E_+} \mathcal{CG}(r_w)$, where r_w is the weighted rank function of $\mathcal{M}$ with weights w.

Contention Resolution Schemes. Chekuri et al. [12] introduced contention resolution (CR) schemes as a tool for maximizing a (not necessarily monotone) submodular function f subject to downward-closed constraints, such as matroid constraints, knapsack constraints, and their intersections. Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid imposing one of these constraints. Given a fractional solution x with

[1] We note that such functions are exactly the one-dimensional monotone $M^\natural$-concave functions $f_i\colon \mathbb{Z}_+ \to \mathbb{R}_+$.

multilinear extension value $F(x)$, their CR scheme randomly rounds x to an integral solution χ_S where $S \in \mathcal{I}$ such that $\mathbb{E}[\chi_S] \geq \inf_{w \in \mathbb{R}_+^E} \mathcal{CG}(r_w) F(x)$. Here, r_w is again the weighted rank function of $\mathcal{M}$ with weights w.

Motivated by the significance of the correlation gap in algorithmic applications, we study the correlation gap of weighted matroid rank functions. It is well-known that $\mathcal{CG}(f) \geq 1 - 1/e$ for every monotone submodular function f [8]. Moreover, the extreme case $1 - 1/e$ is already achieved by the rank function of a rank-1 uniform matroid as $|E| \to \infty$. More generally, the rank function of a rank-ℓ uniform matroid has correlation gap $1 - e^{-\ell}\ell^\ell/\ell! \geq 1 - 1/e$ [6,28]. Other than for uniform matroids, we are not aware of any previous work that gave better than $1 - 1/e$ bounds on the correlation gap of specific matroids.

First, we show that among all weighted rank functions of a matroid, the smallest correlation gap is realized by its (unweighted) rank function.

Theorem 1. *For any matroid $\mathcal{M} = (E, \mathcal{I})$ with rank function $r = r_{\mathbb{1}}$,*

$$\inf_{w \in \mathbb{R}_+^E} \mathcal{CG}(r_w) = \mathcal{CG}(r).$$

For the purpose of lower bounding $\mathcal{CG}(r_w)$, Theorem 1 allows us to ignore the weights w and just focus on the matroid $\mathcal{M}$. As an application, to bound the approximation ratio of sequential posted-price mechanisms as in [28], it suffices to focus on the underlying matroid. We remark that $\mathcal{M}$ can be assumed to be *connected*, that is, it cannot be written as a direct sum of at least two nonempty matroids. Otherwise, $r = \sum_{i=1}^m r_i$ for matroid rank functions r_i with disjoint supports, and so $\mathcal{CG}(r) = \min_{i \in [m]} \mathcal{CG}(r_i)$. For example, the correlation gap of a partition matroid is equal to the smallest correlation gap of its parts (uniform matroids).

Our goal is to identify the parameters of a matroid which govern its correlation gap. A natural candidate is the rank $r(E)$. However, as pointed out by Yan [28], there exist matroids with arbitrarily high rank whose correlation gap is still $1-1/e$, e.g., partition matroids with rank-1 parts. The $1-e^{-\ell}\ell^\ell/\ell!$ bound for uniform matroids [6,28] is suggestive of girth as another potential candidate. Recall that the *girth* of a matroid is the smallest size of a dependent set. On its own, a large girth does not guarantee improved correlation gap bounds: in the full version, we show that for any $\gamma \in \mathbb{N}$, there exist matroids with girth γ whose correlation gaps are arbitrarily close to $1 - 1/e$.

It turns out that the correlation gap heavily depends on the relative values of the rank and girth of the matroid. Our second result is an improved lower bound on the correlation gap as a function of these two parameters.

Theorem 2. *Let $\mathcal{M} = (E, \mathcal{I})$ be a loopless matroid with rank function r, rank $r(E) = \rho$, and girth γ. Then,*

$$\mathcal{CG}(r) \geq 1 - \frac{1}{e} + \frac{e^{-\rho}}{\rho} \left(\sum_{i=0}^{\gamma-2} (\gamma - 1 - i) \left[\binom{\rho}{i} (e-1)^i - \frac{\rho^i}{i!} \right] \right) \geq 1 - \frac{1}{e}.$$

Furthermore, the last inequality is strict whenever $\gamma > 2$.

Figure 1 illustrates the behaviour of the expression in Theorem 2. For any fixed girth γ, it is monotone decreasing in ρ. On the other hand, for any fixed rank ρ, it is monotone increasing in γ. In the full version, we also give complementing albeit non-tight upper bounds that behave similarly with respect to these parameters. When $\rho = \gamma - 1$, our lower bound simplifies to $1 - e^{-\rho}\rho^\rho/\rho!$, i.e., the correlation gap of a rank-ρ uniform matroid (proven in the full version).

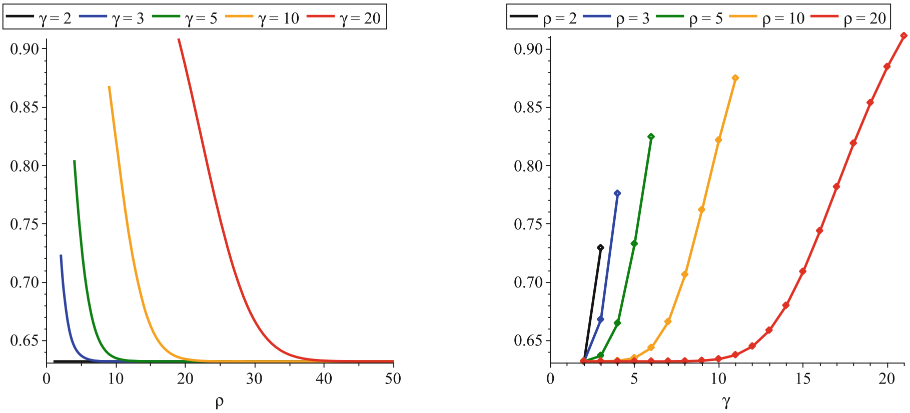

Fig. 1. Our correlation gap bound as a function of the rank ρ and girth γ separately.

The rank and girth have meaningful interpretations in the aforementioned applications. For instance, consider the problem of maximizing a sum of weighted matroid rank functions $\sum_{i=1}^{m} f_i$ under a matroid constraint $(E, \mathcal{J})$. For every $i \in [m]$, let $\mathcal{M}_i$ be the matroid of f_i. In game-theoretic contexts, each f_i usually represents the utility function of agent i. Thus, our goal is to select a bundle of items $S \in \mathcal{J}$ which maximizes the total welfare. If $\mathcal{M}_i$ has girth γ and rank ρ, this means that agent i is interested in $\gamma - 1 \leq k \leq \rho$ items with positive weights. The special case $\rho = \gamma - 1$ (uniform matroids) has already found applications in list decoding [6] and approval voting [14]. On the other hand, for sequential posted-price mechanisms, if the underlying matroid $\mathcal{M}$ has girth γ and rank ρ, this means that the seller can service $\gamma - 1 \leq k \leq \rho$ agents simultaneously.

To the best of our knowledge, our results give the first improvement over the $(1 - 1/e)$ bound on the correlation gap of general matroids. We hope that our paper will motivate further studies into more refined correlation gap bounds, exploring the dependence on further matroid parameters, as well as obtaining tight bounds for special matroid classes.

1.1 Our Techniques

We now give a high-level overview of the proofs of Theorem 1 and Theorem 2.

Weighted Rank Functions. The first step in proving both theorems is to deduce structural properties of the points which realize the correlation gap. In Theorem 4, we show that such a point x can be found in the independent set polytope

$\mathcal{P}$. This implies that $\hat{r}_w(x) = w^\top x$ for any weights $w \in \mathbb{R}^E_+$. Moreover, we deduce that $x(E)$ is integral.

To prove Theorem 1, we fix a matroid $\mathcal{M}$ and derive a contradiction for a non-uniform weighting. More precisely, we consider a weighting $w \in \mathbb{R}^E_+$ and a point $x^* \in [0,1]^E$ which give a smaller ratio $R_w(x^*)/\hat{r}_w(x^*) < \mathcal{CG}(r)$. By the above, we can use the simpler form $R_w(x^*)/\hat{r}_w(x^*) = R_w(x^*)/w^\top x^*$. We pick w such that it has the smallest number of different values. If the number of distinct values is at least 2, then we derive a contradiction by showing that a better solution can be obtained by increasing the weights in a carefully chosen value class until they coincide with the next smallest value. The greedy maximization property of matroids is essential for this argument.

Uniform Matroids. Before outlining our proof of Theorem 2, let us revisit the correlation gap of uniform matroids. Let $\mathcal{M} = (E, \mathcal{I})$ be a uniform matroid on n elements with rank $\rho = r(E)$. If $\rho = 1$, then it is easy to verify that the symmetric point $x = (1/n) \cdot \mathbb{1}$ realizes the correlation gap $1 - 1/e$. Since x lies in the independent set polytope, we have $\hat{r}(x) = \mathbb{1}^\top x = 1$. If one samples each $i \in E$ with probability $1/n$, the probability of selecting at least one element is $R(x) = 1-(1-1/n)^n$. Thus, $\mathcal{CG}(r) = 1-(1-1/n)^n$, which converges to $1-1/e$ as $n \to \infty$. More generally, for $\rho \geq 1$, Yan [28] showed that the symmetric point $x = (\rho/n) \cdot \mathbb{1}$ similarly realizes the correlation gap $1 - e^{-\rho}\rho^\rho/\rho!$.

Poisson Clock Analysis. To obtain the $(1 - 1/e)$ lower bound on the correlation gap of a monotone submodular function, Calinescu et al. [8] introduced an elegant probabilistic analysis. Instead of sampling each $i \in E$ with probability x_i, they consider n independent *Poisson clocks* of rate x_i that are active during the time interval $[0, 1]$. Every clock may send at most one signal from a Poisson process. Let $Q(t)$ be the set of elements whose signal was sent between time 0 and t; the output is $Q(1)$. It is easy to see that $\mathbb{E}[f(Q(1))] \leq F(x)$.

In [8], they show that the derivative of $\mathbb{E}[f(Q(t))]$ can be lower bounded as $f^*(x) - \mathbb{E}[f(Q(t))]$ for every $t \in [0, 1]$, where

$$f^*(x) \coloneqq \min_{S \subseteq E} \left(f(S) + \sum_{i \in E} f_S(i) x_i \right) \tag{4}$$

is an extension of f such that $f^* \geq \hat{f}$. The bound $\mathbb{E}[f(Q(1))] \geq (1 - 1/e)f^*(x)$ is obtained by solving a differential inequality. Thus, $F(x) \geq \mathbb{E}[f(Q(1))] \geq (1 - 1/e)f^*(x) \geq (1 - 1/e)\hat{f}(x)$ follows.

A Two Stage Approach. If f is a matroid rank function, then $f^* = \hat{f}$ (see Theorem 3). Still, the factor $(1 - 1/e)$ in the analysis of [8] cannot be improved: for an integer $x \in \mathcal{P}$, we lose a factor $(1-1/e)$ due to $\mathbb{E}[f(Q(1))] = (1-1/e)F(x)$, even though the extensions coincide: $F(x) = \hat{f}(x)$.

Our analysis in Sect. 4 proceeds in two stages. Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid with rank ρ and girth γ. The basic idea is that up to sets of size $\gamma - 1$, our matroid 'looks like' a uniform matroid. Since the correlation gap of uniform matroids is well-understood, we first extract a uniform matroid of rank $\gamma - 1$ from our matroid, and then analyze the contribution from the remaining part separately. More precisely, we decompose the rank function as $r = g + h$, where $g(S) = \min\{|S|, \ell\}$ is the rank function of a uniform matroid of rank $\ell = \gamma - 1$. Note that the residual function $h := f - g$ is not submodular in general, as $h(S) = 0$ for all $|S| \leq \ell$. We will lower bound the multilinear extensions $G(x)$ and $H(x)$ separately. As g is the rank function of a uniform matroid, similarly as above we can derive a tight lower bound on G in terms of its rank $\ell = \gamma - 1$.

Bounding $H(x)$ is based on a Poisson clock analysis as in [8], but is significantly more involved. Due to the monotonicity of h, directly applying the result in [8] would yield $\mathbb{E}[h(Q(1)] \geq (1 - 1/e)h^*(x)$. However, $h^*(x) = 0$ whenever $\mathcal{M}$ is loopless ($\ell \geq 1$), as $h(\emptyset) = 0$ and $h(\{i\}) = 0$ for all $i \in E$. So, the argument of [8] directly only leads to the trivial $\mathbb{E}[h(Q(1))] \geq 0$. Nevertheless, one can still show that, conditioned on the event $|Q(t)| \geq \ell$, the derivative of $\mathbb{E}[H(Q(t))]$ is at least $r^*(x) - \ell - \mathbb{E}[H(Q(t))]$. Let $T \geq 0$ be the earliest time such that $|Q(T)| \geq \ell$, which we call the *activation time* of Q. Then, solving a differential inequality produces $\mathbb{E}[h(Q(1))|T = t] \geq (1 - e^{-(1-t)})(r^*(x) - \ell)$ for all $t \leq 1$.

To lower bound $\mathbb{E}[h(Q(1))]$, it is left to take the expectation over all possible activation times $T \in [0, 1]$. Let $\bar{h}(x) = (r^*(x) - \ell) \int_0^1 \Pr[T = t](1 - e^{-(1-t)})dt$ be the resulting expression. We prove that $\bar{h}(x)$ is concave in each direction $e_i - e_j$ for $i, j \in E$. This allows us to round x to an integer $x' \in [0, 1]^E$ such that $x'(E) = x(E)$ and $\bar{h}(x') \leq \bar{h}(x)$; recall that $x(E) \in \mathbb{Z}$ by Theorem 4. After substantial simplification of $\bar{h}(x')$, we arrive at the formula in Theorem 2, except that ρ is replaced by $x(E)$. So, the rounding procedure effectively shifts the dependency of the lower bound from the value of x to the value of $x(E)$. Since $x(E) \leq \rho$ by Theorem 4, the final step is to prove that the formula in Theorem 2 is monotone decreasing in ρ.

2 Preliminaries

We denote $\mathbb{Z}_+$ and $\mathbb{R}_+$ as the set of nonnegative integers and nonnegative reals respectively. For $n, k \in \mathbb{Z}_+$, $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ if $n \geq k$, and 0 otherwise. For a set S and $i \in S$, $j \notin S$, we use the shorthand $S - i = S \setminus \{i\}$ and $S + j = S \cup \{j\}$. For a function $f : 2^E \to \mathbb{R}$, a set $S \subseteq E$ and an element $i \in E$, let $f_S(i)$ denote the marginal gain of adding i to S, i.e., $f_S(i) := f(S + i) - f(S)$. For $x \in \mathbb{R}^E$ and $S \subseteq E$, we write $x(S) = \sum_{i \in S} x_i$.

Matroids. Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid with rank function $r : 2^E \to \mathbb{Z}_+$. Its *independent set polytope* $\mathcal{P}(r)$ is the convex hull of incidence vectors of independent sets in $\mathcal{I}$. Equivalently, $\mathcal{P}(r) = \left\{x \in \mathbb{R}_+^E : x(S) \leq r(S)\ \forall S \subseteq E\right\}$, as shown by Edmonds [25, Theorem 40.2]. We need another classical result by Edmonds [25, Theorem 40.3] on intersecting the independent set polytope with a box.

Theorem 3. *For a matroid rank function $r : 2^E \to \mathbb{Z}_+$ and $x \in \mathbb{R}_+^E$,*

$$\max\{y(E) : y \in \mathcal{P}(r), y \leq x\} = \min\{r(T) + x(E \setminus T) : T \subseteq E\}.$$

Probability Distributions. Let $\text{Bin}(n, p)$ denote the binomial distribution with n trials and success probability p. Let $\text{Poi}(\lambda)$ denote the Poisson distribution with rate λ. Recall that $\Pr(\text{Poi}(\lambda) = k) = e^{-\lambda}\lambda^k/k!$ for any $k \in \mathbb{Z}_+$.

Definition 1. *Given random variables X and Y, we say that X* is at least Y in the concave order *if for every concave function $\varphi : \mathbb{R} \to \mathbb{R}$, we have $\mathbb{E}[\varphi(X)] \geq \mathbb{E}[\varphi(Y)]$ whenever the expectations exist. It is denoted as $X \geq_{cv} Y$.*

Lemma 1 ([6]). *For any $n \in \mathbb{N}$ and $p \in [0, 1]$, we have $\text{Bin}(n, p) \geq_{cv} \text{Poi}(np)$.*

Properties of the Multilinear Extension. For a set function $f : 2^E \to \mathbb{R}$, let $F : [0, 1]^E \to \mathbb{R}$ denote its multilinear extension. We will use the following well-known properties of F, see e.g. [9].

Proposition 1. *If f is monotone, then $F(x) \geq F(y)$ for all $x \geq y$.*

Proposition 2. *If f is submodular, then for any $x \in [0, 1]^E$ and $i, j \in E$, the function $\phi(t) := F(x + t(e_i - e_j))$ is convex.*

3 Locating the Correlation Gap

In this section, given a weighted matroid rank function r_w, we locate a point $x^* \in [0, 1]^E$ on which the correlation gap $\mathcal{CG}(r_w)$ is realized, and derive some structural properties. Using this, we prove Theorem 1, i.e., the smallest correlation gap over all possible weightings is attained by uniform weights. We start with a more convenient characterization of the concave extension of r_w.

Lemma 2. *Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid with rank function r and weights $w \in \mathbb{R}^E_+$. For any $x \in [0, 1]^E$, we have $\hat{r}_w(x) = \max\{w^\top y : y \in \mathcal{P}(r),\ y \leq x\}$.*

Next, we show that x^* can be chosen to lie in the independent set polytope $\mathcal{P}(r)$; and that $\text{supp}(x^*)$ is a tight set w.r.t. x^*, meaning $x^*(E) = r(\text{supp}(x^*))$.

Theorem 4. *Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid with rank function r. For any weights $w \in \mathbb{R}^E_+ \setminus \{\mathbb{0}\}$, there exists a point $x^* \in \mathcal{P}(r)$ such that $\mathcal{CG}(r_w) = R_w(x^*)/\hat{r}_w(x^*)$ and $x^*(E) = r(\text{supp}(x^*))$.*

Proof (of Theorem 1). For the purpose of contradiction, suppose that there exist weights $w \in \mathbb{R}^E_+$ and a point $x^* \in [0, 1]^E$ such that $R_w(x^*)/\hat{r}_w(x^*) < \mathcal{CG}(r)$. According to Theorem 4, we may assume that $x^* \in \mathcal{P}(r)$. Thus, $\hat{r}_w(x^*) = w^\top x^*$ by Lemma 2.

Let $w^1 > w^2 > \cdots > w^k \geq 0$ denote the distinct values of w. For each $i \in [k]$, let $E_i \subseteq E$ denote the set of elements with weight w_i. Clearly, $k \geq 2$, as otherwise $R_w(x^*)/\hat{r}_w(x^*) = w^1 R(x^*)/(w^1 x^*(E)) = R(x^*)/x^*(E) \geq \mathcal{CG}(r)$. Let us pick a counterexample with k minimal.

First, we claim that $w_k > 0$. Indeed, if the smallest weight is $w_k = 0$, then $R_w(x^*)$ and $\hat{r}_w(x^*)$ remain unchanged after setting $w_e \leftarrow w^1$ and $x^*_e \leftarrow 0$ for all $e \in E_k$; this contradicts the minimal choice of k.

Let X be the random variable for the set obtained by sampling every element $e \in E$ independently with probability x^*_e. Let $I_X \subseteq X$ denote a maximum weight independent subset of X. Recall the well-known property of matroids that a maximum weight independent set can be selected greedily in decreasing order of the weights w_e. We fix an arbitrary tie-breaking rule inside each set E_i.

The correlation gap of r_w is given by

$$\frac{R_w(x^*)}{\hat{r}_w(x^*)} = \frac{\sum_{S\subseteq E} \Pr(X=S) r_w(S)}{w^\top x^*} = \frac{\sum_{i=1}^k w^i \sum_{e\in E_i} \Pr(e \in I_X)}{\sum_{i=1}^k w^i x^*(E_i)}.$$

Consider the set

$$J := \arg\min_{i\in[k]} \frac{\sum_{e\in E_i} \Pr(e\in I_X)}{x^*(E_i)}.$$

We claim that $J \setminus \{1\} \neq \emptyset$. Suppose that $J = \{1\}$ for a contradiction. Define the point $x' \in \mathcal{P}(r)$ as $x'_e := x^*_e$ if $e \in E_1$, and $x'_e := 0$ otherwise. Then, we get a contradiction from

$$\mathcal{CG}(r) \le \frac{R(x')}{\hat{r}(x')} = \frac{w^1 \sum_{e\in E_1} \Pr(e\in I_X)}{w^1 x^*(E_1)} < \frac{\sum_{i=1}^k w^i \sum_{e\in E_i} \Pr(e\in I_X)}{\sum_{i=1}^k w^i x^*(E_i)} = \frac{R_w(x^*)}{\hat{r}_w(x^*)}.$$

The first equality holds because for every $e \in E_1$, $\Pr(e \in I_X)$ only depends on $x^*_{E_1} = x'_{E_1}$. This is by the greedy choice of I_X: elements in E_1 are selected based only on $X \cap E_1$. The strict inequality is due to $J = \{1\}$, $k \ge 2$ and $w_2 > 0$.

Now, pick any index $j \in J \setminus \{1\}$. Since $w^j > 0$, we have

$$\frac{w^j \sum_{e\in E_j} \Pr(e\in I_X)}{w^j x^*(E_j)} \le \frac{\sum_{i=1}^k w^i \sum_{e\in E_i} \Pr(e\in I_X)}{\sum_{i=1}^k w^i x^*(E_i)}.$$

So, we can increase w^j to w^{j-1} without increasing the correlation gap. That is, defining $\bar{w} \in \mathbb{R}^E_+$ as $\bar{w}_e := w^{j-1}$ if $e \in E_j$ and $\bar{w}_e := w_e$ otherwise, we get

$$\begin{aligned}\frac{R_w(x^*)}{\hat{r}_w(x^*)} &\ge \frac{\sum_{i\neq j} w^i \sum_{e\in E_i} \Pr(e\in I_X) + w^{j-1}\sum_{e\in E_j} \Pr(e\in I_X)}{\sum_{i\neq j} w^i x^*(E_i) + w^{j-1} x^*(E_j)} \\ &= \frac{\sum_{S\subseteq E} \Pr(X=S) r_{\bar{w}}(S)}{\bar{w}^\top x^*} \ge \min_{x\in[0,1]^E} \frac{R_{\bar{w}}(x)}{\hat{r}_{\bar{w}}(x)}\ .\end{aligned}$$

The equality holds because for every $S \subseteq E$, I_S remains a max-weight independent set with the new weights $\bar{w}$. This contradicts the minimal choice of k. □

4 Lower Bounding the Correlation Gap

This section is dedicated to the proof of Theorem 2. Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid with rank function r, rank $\rho = r(E)$ and girth $\gamma > 1$. By Theorem 4, there exists a point $x^* \in \mathcal{P}(r)$ such that $\mathcal{CG}(r) = R(x^*)/r(x^*)$ and $x^*(E) = r(\mathrm{supp}(x^*))$. For the sake of brevity, we denote $\ell = \gamma - 1$ and $\lambda = x^*(E) \in \mathbb{Z}_+$. Note that if

$\lambda < \ell$, then $\text{supp}(x^*)$ is independent. As $x^*(E) = r(\text{supp}(x^*)) = |\text{supp}(x^*)|$, we have $x_i^* = 1$ for all $i \in \text{supp}(x^*)$. Since x^* is integral, the correlation gap is 1 because $R(x^*) = \hat{r}(x^*)$. Henceforth, we will assume that $\lambda \geq \ell$.

From Lemma 2, we already know that $\hat{r}(x^*) = \mathbb{1}^\top x^* = \lambda$. So, it remains to analyze $R(x^*)$. Let g be the rank function of a rank-ℓ uniform matroid on ground set E, and define the function $h := r - g \geq 0$. By linearity of expectation, $R(x^*) = G(x^*) + H(x^*)$. We lower bound $G(x^*)$ and $H(x^*)$ separately.

4.1 Lower Bounding $G(x^*)$

As g is the rank function of a uniform matroid, the arguments of Yan [28] and Barman et al. [6] apply. In particular, since G is a symmetric polynomial, and convex along $e_i - e_j$ for all $i, j \in E$ by Proposition 2, we have

$$G(x^*) \geq G\left(\frac{\lambda}{n} \cdot \mathbb{1}\right) = \mathbb{E}\left[\min\left\{\text{Bin}\left(n, \frac{\lambda}{n}\right), \ell\right\}\right] \geq \mathbb{E}\left[\min\left\{\text{Poi}(\lambda), \ell\right\}\right] \quad . \tag{5}$$

The last inequality follows from Lemma 1. The latter expectation is equal to

$$\sum_{j=1}^{\ell} \Pr(\text{Poi}(\lambda) \geq j) = \sum_{j=1}^{\ell}\left(1 - \sum_{k=0}^{j-1} \frac{\lambda^k e^{-\lambda}}{k!}\right) = \ell - \sum_{k=0}^{\ell-1} (\ell - k) \frac{\lambda^k e^{-\lambda}}{k!}. \tag{6}$$

4.2 Lower Bounding $H(x^*)$

Our analysis of $H(x^*)$ uses the Poisson clock setup of Calinescu et al. [8], which incrementally builds a set $Q(1)$ as follows. Each element $i \in E$ is assigned a Poisson clock of rate x_i^*. We start all the clocks simultaneously at time $t = 0$, and begin with the initial set $Q(0) = \emptyset$. For $t \in [0, 1]$, if the clock on an element i rings at time t, then we add i to our current set $Q(t)$. We stop at time $t = 1$.

Clearly, $\Pr(i \in Q(1)) = 1 - e^{-x_i^*} \leq x_i^*$ for all $i \in E$. Since h is monotone, Proposition 1 yields $H(x^*) \geq H(1 - e^{-x^*}) = \mathbb{E}[h(Q(1))]$, where equality is due to independence of the Poisson clocks. So, it suffices to lower bound $\mathbb{E}[h(Q(1))]$.

Let $t \in [0, 1)$ and consider an infinitesimally small interval $[t, t + dt]$. For each $i \in E$, the probability of adding i during this interval is $\Pr(\text{Poi}(x_i^* dt) \geq 1) = x_i^* dt + O(dt^2)$. Note that the probability of adding two or more elements is also $O(dt^2)$. Since dt is very small, we can effectively neglect all $O(dt^2)$ terms.

Definition 2. *We say that Q is* activated at time T *if $|Q(t)| < \ell$ for all $t < T$ and $|Q(t)| \geq \ell$ for all $t \geq T$. We call T the* activation time *of Q.*

Let $S \subseteq E$ where $|S| \geq \ell$ and let $t \geq t' \geq 0$. Conditioning on the events $Q(t) = S$ and $T = t'$, the expected increase of $h(Q(t))$ (up to $O(dt^2)$ terms) is

$$\mathbb{E}[h(Q(t + dt)) - h(Q(t)) | Q(t) = S \wedge T = t'] = \sum_{i \in E} r_S(i) x_i^* dt \geq (\lambda - \ell - h(S))dt,$$

where the inequality is due to

$$h(S) + \sum_{i \in E} r_S(i) x_i^* = r(S) - \ell + \sum_{i \in E} r_S(i) x_i^* \geq r^*(x^*) - \ell = \hat{r}(x^*) - \ell = \lambda - \ell.$$

The inequality follows from the definition of r^* in (4), the second equality is by Theorem 3, while the third equality is due to Lemma 2 because $x^* \in \mathcal{P}(r)$. Dividing by dt and taking expectation over S, we obtain for all $t \geq t' \geq 0$,

$$\frac{1}{dt}\mathbb{E}[h(Q(t+dt)) - h(Q(t))|T = t'] \geq \lambda - \ell - \mathbb{E}[h(Q(t))|T = t']. \tag{7}$$

Let $\phi(t) := \mathbb{E}[h(Q(t))|T = t']$. Then, (7) can be written as $\frac{d\phi}{dt} \geq \lambda - \ell - \phi(t)$. To solve this differential inequality, let $\psi(t) := e^t \phi(t)$ and consider $\frac{d\psi}{dt} = e^t(\frac{d\phi}{dt} + \phi(t)) \geq e^t(\lambda - \ell)$. Since $\psi(t') = \phi(t') = 0$, we get

$$\psi(t) = \int_{t'}^{t} \frac{d\psi}{ds} ds \geq \int_{t'}^{t} e^s(\lambda - \ell) ds = (e^t - e^{t'})(\lambda - \ell)$$

for all $t \geq t'$. It follows that $\mathbb{E}[h(Q(t))|T = t'] = \phi(t) = e^{-t}\psi(t) \geq (1 - e^{t'-t})(\lambda - \ell)$ for all $t \geq t'$. In particular, at time $t = 1$, we have $\mathbb{E}[h(Q(1))|T = t'] \geq (1 - e^{t'-1})(\lambda - \ell)$ for all $t' \leq 1$. By the law of total expectation,

$$\mathbb{E}[h(Q(1))] \geq (\lambda - \ell) \int_0^1 \Pr(T = t)(1 - e^{t-1}) dt. \tag{8}$$

Now, the cumulative distribution function of T is given by

$$\begin{aligned}\Pr(T \leq t) &= 1 - \sum_{\substack{S \subseteq E: \\ |S| < \ell}} \prod_{i \in S}(1 - e^{-x_i^* t}) \prod_{i \notin S} e^{-x_i^* t} \\ &\stackrel{\star}{=} 1 - \sum_{S \subseteq E} (-1)^{|S|+\ell-n-1} \binom{|S|-1}{n-\ell} e^{-x^*(S)t}.\end{aligned}$$

Any marked equality $\stackrel{\star}{=}$ indicates that several derivation steps have been skipped, whose details can be found in the full version. Differentiating with respect to t yields the probability density function of T

$$\Pr(T = t) = \frac{d}{dt}\Pr(T \leq t) = \sum_{S \subseteq E} (-1)^{|S|+\ell-n-1} \binom{|S|-1}{n-\ell} x^*(S) e^{-x^*(S)t}.$$

Plugging this back into (8) gives us

$$\begin{aligned}\mathbb{E}[h(Q(1))] &\geq (\lambda - \ell) \sum_{S \subseteq E} (-1)^{|S|+\ell-n-1} \binom{|S|-1}{n-\ell} x^*(S) \int_0^1 e^{-x^*(S)t}(1 - e^{t-1}) dt \\ &= (\lambda - \ell) \sum_{S \subseteq E} (-1)^{|S|+\ell-n-1} \binom{|S|-1}{n-\ell} \left(1 - \frac{1}{e} - \frac{e^{-1} - e^{-x^*(S)}}{x^*(S) - 1}\right) \\ &= (\lambda - \ell) \left[1 - \frac{1}{e} + \sum_{S \subseteq E} (-1)^{|S|+\ell-n} \binom{|S|-1}{n-\ell} \frac{e^{-1} - e^{-x^*(S)}}{x^*(S) - 1}\right]\end{aligned} \tag{9}$$

In the full version, we prove that (9) is concave along $e_i - e_j$ for all $i, j \in E$, when viewed as a function of x^*. This allows us to round x^* to an integral vector $x' \in \{0,1\}^E$ such that $x'(E) = x^*(E)$ without increasing the value of (9). Note that x' has exactly λ ones and $n - \lambda$ zeroes because $\lambda \in \mathbb{Z}_+$ by Theorem 4. Hence, (9) is lower bounded by

$$(\lambda - \ell)\left[1 - \frac{1}{e} + \sum_{i=0}^{\lambda}\sum_{j=0}^{n-\lambda}\binom{\lambda}{i}\binom{n-\lambda}{j}(-1)^{i+j+\ell-n}\binom{i+j-1}{n-\ell}\frac{e^{-1}-e^{-i}}{i-1}\right]$$
$$\stackrel{\star}{=} (\lambda - \ell)\left[1 - \frac{1}{e} + \sum_{i=0}^{\ell-1}(-1)^{\ell-i}\binom{\lambda}{i}\binom{\lambda-i-1}{\ell-i-1}\frac{e^{-1}-e^{-(\lambda-i)}}{\lambda-i-1}\right] . \quad (10)$$

Since (10) evaluates to 0 when $\lambda = \ell$, let us assume that $\lambda > \ell$. Then, using $\frac{1}{\lambda-i-1}\binom{\lambda-i-1}{\ell-i-1} = \frac{1}{\lambda-\ell}\binom{\lambda-i-2}{\ell-i-1}$, we can simplify (10) as

$$(\lambda - \ell)\left(1 - \frac{1}{e}\right) + \sum_{i=0}^{\ell-1}(-1)^{\ell-i}\binom{\lambda}{i}\binom{\lambda-i-2}{\ell-i-1}\left(e^{-1} - e^{-(\lambda-i)}\right)$$
$$\stackrel{\star}{=} \lambda\left(1 - \frac{1}{e}\right) - \ell + e^{-\lambda}\sum_{i=0}^{\ell-1}(-1)^{\ell-i-1}\binom{\lambda}{i}\binom{\lambda-i-2}{\ell-i-1}e^{i} . \quad (11)$$

The sum in (11) can be viewed as a univariate polynomial of degree $\ell - 1$ in $\alpha \in \mathbb{R}$ for $\alpha = e$. Taking its Taylor expansion at $\alpha = 1$, we can rewrite (11) as

$$\lambda\left(1 - \frac{1}{e}\right) - \ell + e^{-\lambda}\sum_{i=0}^{\ell-1}\binom{\lambda}{i}(\ell - i)(e-1)^{i} . \quad (12)$$

4.3 Putting Everything Together

We are finally ready to lower bound the correlation gap of the matroid rank function r. Recall that we assumed $\lambda > \ell$ in the previous subsection. Combining the lower bounds (6) and (12) gives us

$$\mathcal{CG}(r) = \frac{G(x^*) + H(x^*)}{\mathbb{1}^\top x^*} = 1 - \frac{1}{e} + \frac{e^{-\lambda}}{\lambda}\sum_{i=0}^{\ell-1}(\ell - i)\left[\binom{\lambda}{i}(e-1)^{i} - \frac{\lambda^i}{i!}\right] . \quad (13)$$

On the other hand, if $\lambda = \ell$, then $h = 0$. By (6), we obtain

$$\mathcal{CG}(r) = \frac{G(x^*)}{\mathbb{1}^\top x^*} = \frac{G(x^*)}{\ell} \geq 1 - \sum_{k=0}^{\ell-1}\left(1 - \frac{k}{\ell}\right)\frac{\ell^k e^{-\ell}}{k!} = 1 - \frac{\ell^{\ell-1}e^{-\ell}}{(\ell-1)!} , \quad (14)$$

which agrees with (13) when $\lambda = \ell$ (proven in full version).

To finish the proof of Theorem 2, it is left to show that (13) is a decreasing function of λ because $\lambda \leq \rho$. We also need to prove that the final expression is strictly greater than $1 - 1/e$ whenever $\ell \geq 2$. These are done in the full version.

References

1. Ageev, A.A., Sviridenko, M.: Pipage rounding: a new method of constructing algorithms with proven performance guarantee. J. Comb. Optim. **8**(3), 307–328 (2004)
2. Agrawal, S., Ding, Y., Saberi, A., Ye, Y.: Price of correlations in stochastic optimization. Oper. Res. **60**(1), 150–162 (2012)
3. Asadpour, A., Niazadeh, R., Saberi, A., Shameli, A.: Sequential submodular maximization and applications to ranking an assortment of products. In: EC 2022: The 23rd ACM Conference on Economics and Computation, p. 817 (2022)
4. Ausubel, L.M., Milgrom, P.: The lovely but lonely Vickrey auction. In: Cramton, P., Shoham, Y., Steinberg, R. (eds.) Combinatorial Auctions, chap. 1. MIT Press (2006)
5. Barman, S., Fawzi, O., Fermé, P.: Tight approximation guarantees for concave coverage problems. In: 38th International Symposium on Theoretical Aspects of Computer Science (STACS). LIPIcs, vol. 187, pp. 1–17 (2021)
6. Barman, S., Fawzi, O., Ghoshal, S., Gürpinar, E.: Tight approximation bounds for maximum multi-coverage. Math. Program. **192**(1), 443–476 (2022)
7. Bhalgat, A., Chakraborty, T., Khanna, S.: Mechanism design for a risk averse seller. In: Goldberg, P.W. (ed.) WINE 2012. LNCS, vol. 7695, pp. 198–211. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-35311-6_15
8. Calinescu, G., Chekuri, C., Pál, M., Vondrák, J.: Maximizing a submodular set function subject to a matroid constraint (Extended Abstract). In: Fischetti, M., Williamson, D.P. (eds.) IPCO 2007. LNCS, vol. 4513, pp. 182–196. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-72792-7_15
9. Călinescu, G., Chekuri, C., Pál, M., Vondrák, J.: Maximizing a monotone submodular function subject to a matroid constraint. SIAM J. Comput. **40**(6), 1740–1766 (2011)
10. Chawla, S., Hartline, J.D., Malec, D.L., Sivan, B.: Multi-parameter mechanism design and sequential posted pricing. In: Schulman, L.J. (ed.) Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC 2010, Cambridge, Massachusetts, USA, 5–8 June 2010, pp. 311–320. ACM (2010)
11. Chekuri, C., Livanos, V.: On submodular prophet inequalities and correlation gap. In: 14th International Symposium on Algorithmic Game Theory, SAGT. Lecture Notes in Computer Science, vol. 12885, p. 410 (2021)
12. Chekuri, C., Vondrák, J., Zenklusen, R.: Submodular function maximization via the multilinear relaxation and contention resolution schemes. SIAM J. Comput. **43**(6), 1831–1879 (2014)
13. Clarke, E.H.: Multipart pricing of public goods. Public choice, pp. 17–33 (1971)
14. Dudycz, S., Manurangsi, P., Marcinkowski, J., Sornat, K.: Tight approximation for proportional approval voting. In: Bessiere, C. (ed.) Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020, pp. 276–282. ijcai.org (2020)
15. Dughmi, S.: Matroid secretary is equivalent to contention resolution. In: 13th Innovations in Theoretical Computer Science Conference, ITCS. LIPIcs, vol. 215, pp. 1–23 (2022)
16. Feige, U.: A threshold of $\ln n$ for approximating set cover. J. ACM (JACM) **45**(4), 634–652 (1998)
17. Groves, T.: Incentives in teams. Econometrica: J. Econometric Soc. **41**, 617–631 (1973)
18. Hartline, J.D.: Mechanism design and approximation (2013)

19. Leme, R.P.: Gross substitutability: an algorithmic survey. Games Econom. Behav. **106**, 294–316 (2017)
20. Murota, K.: On basic operations related to network induction of discrete convex functions. Optim. Methods Softw. **36**(2–3), 519–559 (2021)
21. Myerson, R.B.: Optimal auction design. Math. Oper. Res. **6**(1), 58–73 (1981)
22. Nikolova, E.: Approximation algorithms for reliable stochastic combinatorial optimization. In: Serna, M., Shaltiel, R., Jansen, K., Rolim, J. (eds.) APPROX/RANDOM -2010. LNCS, vol. 6302, pp. 338–351. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-15369-3_26
23. Nisan, N., Ronen, A.: Computationally feasible VCG mechanisms. J. Arti. Intell. Res. **29**, 19–47 (2007)
24. Rubinstein, A., Singla, S.: Combinatorial prophet inequalities. In: Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 1671–1687. SIAM (2017)
25. Schrijver, A.: Combinatorial optimization: polyhedra and efficiency, vol. 24. Springer (2003)
26. Shioura, A.: On the Pipage rounding algorithm for submodular function maximization - a view from discrete convex analysis. Discret. Math. Algorithms Appl. **1**(1), 1–24 (2009)
27. Vickrey, W.: Counterspeculation, auctions, and competitive sealed tenders. J. Financ. **16**(1), 8–37 (1961)
28. Yan, Q.: Mechanism design via correlation gap. In: Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms, pp. 710–719. SIAM (2011)