2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)

Maximizing Determinants under Matroid Constraints

Vivek Madan*, Aleksandar Nikolov†, Mohit Singh‡ and Uthaipon (Tao) Tantipongpipat§
*Amazon, New York, USA. Email: vmadan7@gatech.edu.
†University of Toronto, Toronto, Canada. Email: anikolov@cs.toronto.edu.
‡Georgia Institute of Technology. Atlanta, Georgia, USA. Email: mohit.singh@isye.gatech.edu.
§Twitter, San Francisco, USA. Email: tao@gatech.edu.

***Abstract*—Given a set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^d$ and a matroid $\mathcal{M} = ([n], \mathcal{I})$, we study the problem of finding a basis S of $\mathcal{M}$ such that $\det\left(\sum_{i \in S} \mathbf{v}_i \mathbf{v}_i^\top\right)$ is maximized. This problem appears in a diverse set of areas, such as experimental design, fair allocation of goods, network design, and machine learning. The current best results include an e^{2k}-estimation for any matroid of rank k [8] and a $(1+\epsilon)^d$-approximation for a uniform matroid of rank $k \geq d + \frac{d}{\epsilon}$ [30], where the rank $k \geq d$ denotes the desired size of the optimal set. Our main result is a new approximation algorithm for the general problem with an approximation guarantee that depends only on the dimension d of the vectors, and not on the size k of the output set. In particular, we show an $(O(d))^d$-estimation and an $(O(d))^{d^3}$-approximation for any matroid, giving a significant improvement over prior work when $k \gg d$.**

Our result relies on showing that there exists an optimal solution to a convex programming relaxation for the problem which has *sparse support*; in particular, no more than $O(d^2)$ variables of the solution have fractional values. The sparsity results rely on the interplay between the first order optimality conditions for the convex program and matroid theory. We believe that the techniques introduced to show sparsity of optimal solutions to convex programs will be of independent interest. We also give a new randomized rounding algorithm that crucially exploits the sparsity of solutions to the convex program. To show the approximation guarantee, we utilize recent works on strongly log-concave polynomials [8], [4] and show new relationships between different convex programs [33], [6] studied for the problem. Finally, we show how to use the estimation algorithm to give an efficient deterministic approximation algorithm. Once again, the algorithm crucially relies on sparsity of the fractional solution to guarantee that the approximation factor depends solely on the dimension d.

I. Introduction

Choosing a diverse representative set of items from a large corpus is a common problem studied in a variety of areas, including machine learning, information retrieval, statistics, and optimization [27], [17], [16], [34]. For example, consider the problem of choosing a subset from a large data set to train a machine learning algorithm; or of displaying a small set of images out of a large set of relevant images to a search query. In these contexts, one aims to choose a small and diverse representative set of items from a large data set. Diversity here can be modeled in many different ways, and the choice of a diversity measure can significantly affect both practical performance and the algorithmic complexity of finding a diverse set. Both general and application-specific diversity criteria have been proposed in the past [22], [15], [16], [40], [13].

In this work, we focus on a popular geometric model of the problem above. While it naturally captures problems in data retrieval and statistics, we show that it also encompasses problems in fair allocation of goods, network design, counting, and optimization. We assume that data are represented as points in the d-dimensional Euclidean space, so that choosing a subset of items corresponds to selecting a subset of d-dimensional vectors. A number of natural diversity measures can be formulated in terms of functions of the eigenvalues of the matrix given by the sum of outerproducts of the selected vectors. Some examples are the determinant, the trace, the harmonic mean of the eigenvalues, and the minimum eigenvalue. In this work, we focus on the determinant as the diversity measure. We study the determinant maximization problem with general combinatorial constraints which makes the model rich enough to include many of the problems mentioned above. In particular, we consider matroid constraints, which capture cardinality constraints, partition constraints, and many more as special cases. This allows modeling constraints imposed by, e.g., budget, feasibility, or fairness considerations.

In an instance of the Determinant Maximization problem (under a general matroid constraint), we are given a set of n vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^d$ and a matroid $\mathcal{M} = ([n], \mathcal{I})$ with set of bases $\mathcal{B}$, and our goal is to find a set $S \in \mathcal{B}$ that maximizes $\det\left(\sum_{i \in S} \mathbf{v}_i \mathbf{v}_i^\top\right)$, i.e.

$$\max \left\{ \det\left(\sum_{i \in S} \mathbf{v}_i \mathbf{v}_i^\top \right) : S \in \mathcal{B} \right\}. \tag{1}$$

We denote by k the rank of the matroid $\mathcal{M}$, which is the size of all the bases in $\mathcal{B}$. We denote the combinatorial optimization problem (1) by D-OPT and its optimum value by OPT.

A number of special cases of D-OPT have been studied, in which either the choice of vectors or the matroid is restricted [39], [10], [38], [1], [35]. We highlight two illustrative examples. Under cardinality constraints, in which $\mathcal{B}$ consists of all subsets of $[n]$ of size k, the problem is hard to approximate to a factor better than $(1+c)^d$ for some $c > 0$ when $k = d$ [26], [18], [20], and Nikolov [32]

2575-8454/20/$31.00 ©2020 IEEE
DOI 10.1109/FOCS46700.2020.00059

gave an e^d-approximation for $k \leq d$.[1] Interestingly, when $k > d$, improved guarantees are known [38], [1], [35] with the current best $(1+\epsilon)^d$-approximation when $k \geq d+\frac{d}{\epsilon}$ [30].

For general matroids, a series of works [33], [6], [37], [8] have focused on the case when $k \leq d$, and the latest results of Anari, Oveis-Gharan, and Vinzant [8] imply an e^{2k}-estimation algorithm. These results were first proved for the special case when the *generating polynomial* for the matroid is a *real stable polynomial* [6]. Recent and exciting advances on completely log-concave polynomials [8] (and the equivalent notion of Lorentzian polynomials [11]) allow the techniques of [6] to be generalized to all matroids. While these results are not stated when $k > d$, the analysis naturally yields an e^{2k}-estimation algorithm even in that case. Such a dependence on k is often exorbitant since k can be much larger than d in many applications. Moreover, the hardness result mentioned above only shows that the approximation factor needs to depend exponentially on d, but not necessarily on k.[2] A starting point for this work is a result showing that these existing techniques are incapable of removing the dependence on k for general matroid constraints. In the extended version [29], we show that any algorithm which solves a convex relaxation and rounds the fractional solution without using the structure of the vectors yields an approximation factor necessarily dependent on k even when $d = 2$.

A. Our Results and Contributions

Our main result is an algorithm that estimates the objective of the DETERMINANT MAXIMIZATION problem under a general matroid constraint.

Theorem I.1 *There is an efficiently computable convex program whose objective value estimates the objective of the* DETERMINANT MAXIMIZATION *problem under a general matroid constraint within a multiplicative factor of* $(O(d))^d$.

As outlined earlier, an approximation factor depending only on d cannot be obtained by rounding an arbitrary optimal solution to any of the known convex relaxations of the problem. Our work introduces two key ideas to bypass this bottleneck. First, we show that a convex programming relaxation always has an optimal *sparse* fractional solution, and, in particular, one with no more than $O(d^2)$ fractional variables, out of a total of n variables. The proof of this fact relies crucially on the first order optimality conditions of the convex program. A straightforward presentation of the first order optimality conditions leads to a system of (exponentially many) *non-linear* constraints over an exponential number of variables. We interpret these constraints using matroid theory and reformulate them as a system of (exponentially many) *linear* inequalities. Then, we apply combinatorial optimization techniques such as uncrossing in order to show that any basic feasible solution to the system of inequalities must be sparse, again using the inherent matroid structure of the linear constraints.

Second, we give a new randomized algorithm that rounds such a sparse solution for any matroid, giving the desired result. Our algorithm crucially uses the near-integral structure of optimal solutions, and thus differs significantly from previous rounding algorithms, which are oblivious to any such structure. The main challenge in the design of the algorithm is that the non-linearity of the objective function implies that even an integral variable cannot be included in the solution with probability 1. Our rounding proceeds in two phases: we first randomly round the fractional variables, and then we randomly choose which of the integral variables to include in a solution, while maintaining feasibility. We again rely on matroid theory to show that the random solution obtained has large objective value in expectation.

This combination of techniques from convex optimization and matroid theory, which we use in order to find a sparse optimal solution of a convex program with exponentially many constraints, appears to be novel and may be of independent interest.

We also consider the special case of partition matroids due to its significant applications and note that an improved approximation algorithm can be obtained for this case. We observe that the roadblock in achieving an approximation factor independent of k for general matroids does not appear in the case of partition matroids. Thus, the standard randomized rounding algorithm also achieves $e^{O(d)}$-approximation by generalizing the results on Nash Social Welfare in [5].

Deterministic Algorithms.: A challenge for the DETERMINANT MAXIMIZATION problem under a general matroid constraint has been the lack of *true* approximation algorithms that achieve the same guarantees as the estimation algorithms. Most results [33], [6], [8], [4], [37] give randomized algorithms whose guarantees hold in expectation and are not known to hold with high probability or deterministically. The few existing efficient algorithms with high probability or deterministic guarantees either work only for restricted classes of matroids, such as uniform matroids [32], [2], [35] or partition matroids with a constant number of parts [14], or rely on special structure of the input vectors (or both) [7], [19], [15], [9]. Ebrahimi, Straszak and Vishnoi [21] gave the most general algorithmic results that apply to all regular matroids, but the approximation factors they achieved depend on the size of the ground set and not just the dimension of vectors, as aimed in our work.

We utilize the existence of sparse optimal solutions to our convex programming relaxation to give an efficient

[1]For $k < d$, the objective is naturally replaced by the product of the k highest eigenvalues of the matrix, rather than the determinant, which is the product of all d eigenvalues

[2]Since the objective is the determinant of $d \times d$ matrices, and the determinant is homogeneous of degree d, exponential dependence on d is an appropriate scaling.

deterministic algorithm achieving an approximation factor that only depends on the dimension d of the vectors, and not on the size k of the output set or the size n of the input.

Theorem I.2 *There is a polynomial time deterministic algorithm for the* DETERMINANT MAXIMIZATION *problem that gives an* $(O(d))^{d^3}$*-approximation.*

The above result is achieved by using the optimal objective value of the convex program as an estimate of the value of an optimal solution, and reducing the search problem of finding an approximately optimal solution to estimation. We have shown that some optimal solution to the convex program has at most $O(d^2)$ fractional variables, and, therefore, has support of size $k + O(d^2)$. Then, producing a feasible solution (which has size k) requires finding $O(d^2)$ elements of the support of the optimal solution to exclude from the solution: the remaining k elements form the output. Thus, the sparsity allows us to argue that the estimation problem needs to be recursively solved only $O(d^2)$ times, which is crucial in guaranteeing an approximation factor that depends only on d.

We remark the guarantee is worse than is achieved (in expectation) by the randomized algorithm. Obtaining true approximation algorithms that match the performance of the estimation algorithms remains a challenging open problem for the DETERMINANT MAXIMIZATION problem under a general matroid constraint, even in the case of a partition constraint.

B. Applications

As mentioned earlier, DETERMINANT MAXIMIZATION models problems in many different areas and our results imply new approximations for many of these problems. We give details for some of them below.

Experimental Design.: In the optimal experimental design problem for linear models, the goal is to infer an unknown $\boldsymbol{\theta}^\star \in \mathbb{R}^d$ from a possible set of linear measurements of the form $y_i = \mathbf{v}_i^\top \boldsymbol{\theta}^\star + \eta_i$. Here, $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^d$ are known vectors, and $\eta_1, \dots, \eta_n$ are independent Gaussian noises with mean 0 and variance 1. In some settings, performing all of the n measurements might be infeasible, and combinatorial constraints such as matroid constraints can be used to define the feasible sets of measurements. Given a set $S \subseteq [n]$ of measurements, an estimator $\widehat{\boldsymbol{\theta}}$ for $\boldsymbol{\theta}^*$ is obtained via solving the least squares regression problem $\min_{\boldsymbol{\theta}\in\mathbb{R}^d} \sum_{i\in S}(y_i - \mathbf{v}_i^\top \boldsymbol{\theta})^2$. The error $\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star$ is distributed as a d-dimensional Gaussian $N\left(0, \left(\sum_{i\in S} \mathbf{v}_i\mathbf{v}_i^\top\right)^{-1}\right)$. Minimizing the volume of the confidence ellipsoid, or equivalently the determinant of the covariance matrix of the error, is referred to as D-optimal design in statistics [34]. Our results directly imply improved approximability for D-optimal design under a general matroid constraint.

Nash Social Welfare.: In the indivisible goods allocation problem the goal is to allocate, i.e. partition, m goods among d agents so that some notion of social welfare and/or fairness is achieved. Each agent i has utility $u_i(j)$ for good $j \in [m]$, and if S_i are the goods assigned to agent i, then her utility is $u_i(S_i) = \sum_{j\in S_i} u_i(j)$. A well studied objective in this context is Nash social welfare (NSW), which asks to maximize $\left(\prod_{i=1}^d u_i(S_i)\right)^{1/d}$. This objective interpolates between maximally efficient and maximally egalitarian allocations – see [31], [12] for more extensive background. Maximizing the NSW can be formulated as an instance of DETERMINANT MAXIMIZATION under a partition constraint, as observed in [7]. For each agent i and good j, we create a vector $\mathbf{v}_{(i,j)} = \sqrt{u_i(j)}\mathbf{e}_i$, where $\mathbf{e}_i$ is the i-th standard basis vector of $\mathbb{R}^d$, and form a partition matroid $\mathcal{M}$ whose bases $\mathcal{B}$ consist of all sets $S \subseteq [d] \times [m]$ such that $|\{i : (i,j) \in S\}| = 1$ for all $j \in [m]$. Then, a feasible solution $S \in \mathcal{B}$ corresponds to an allocation of the goods, and the determinant $\det\left(\sum_{(i,j)\in S} \mathbf{v}_{(i,j)}\mathbf{v}_{(i,j)}^\top\right)$ is equal to the NSW objective. Our results recover those in [7] and further allow us to give an $O(d)$-estimation algorithm when the allocation $(S_1, \dots, S_d)$ is required to satisfy additional matroid constraints. For example, the works [24], [25], [23] considered allocations such that $\bigcup_{i=1}^d S_i$ is a basis of a matroid $\mathcal{M}'$. We can model this setting by defining our constraint matroid $\mathcal{M}$ so that $S \subseteq [d] \times [m]$ is a basis of $\mathcal{M}$ if and only if $|\{i : (i,j) \in S\}| = 1$ for all $j \in [m]$ and $\{j : \exists i \text{ s.t. } (i,j) \in S\}$ is a basis of $\mathcal{M}'$. Our results then imply an $O(d)$-estimation algorithm and an $O(d)^{d^2}$-approximation algorithm for maximizing NSW subject to these general matroid constraints.

Network Design Problems.: In general, the goal in network design problems is to pick a subset F of the edges of an undirected graph $G = (V, E)$ with non-negative edge weights w such that the subgraph $H = (V, F)$ is *well-connected*. One measure of connectivity is to maximize the total weight of spanning trees in $H = (V, F)$, where the weight of a tree is defined as the product of the weights of its edges (see [28] and references therein for other applications). This natural network design problem is a special case of the DETERMINANT MAXIMIZATION problem. For each $(i,j) \in E$, we introduce a vector $\mathbf{v}_{(i,j)} \in \{0, 1, -1\}^V$ with $(\mathbf{v}_{(i,j)})_i = \sqrt{w_{(i,j)}}$, $(\mathbf{v}_{(i,j)})_j = -\sqrt{w_{(i,j)}}$, and the rest of the coordinates set to zero. Observe that $\sum_{e\in F} \mathbf{v}_e\mathbf{v}_e^\top$ is exactly the Laplacian of $H = (V, F)$, and the determinant of the Laplacian[3] gives the number of spanning trees in H. Our results imply an $O(|V|)^{|V|}$-estimation algorithm, and $O(|V|)^{|V|^3}$-approximation algorithm for this problem under a general matroid constraint.

[3]We remark that the Laplacian is always singular, but we can first project the vectors $\mathbf{v}_e$ orthogonal to the all-ones vector and take the determinant in $d - 1$ dimensions.

C. Technical Overview

Our starting point is a variant of the convex relaxation introduced in [33] for the partition matroid. Let the set of input vectors be $V = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subset \mathbb{R}^d$. For a matroid $\mathcal{M} = ([n], \mathcal{I})$, we denote by $\mathcal{I}_s(\mathcal{M}) := \{S \in \mathcal{I} : |S| = s\}$ the set of all independent sets of size s. We denote by $\mathcal{P}(\mathcal{M})$ the matroid base polytope of $\mathcal{M}$, which is the convex hull of the indicator vectors of the bases. For any vector $\mathbf{z} \in \mathbb{R}^n$ and a subset $S \subseteq [n]$, we let $z(S) := \sum_{i \in S} z_i$. We let $\mathcal{Z} := \{\mathbf{z} \in \mathbb{R}^n : \forall S \in \mathcal{I}_d(\mathcal{M}), z(S) \geq 0\}$. Our convex relaxation is

$$\sup_{\mathbf{x} \in \mathcal{P}(\mathcal{M})} \inf_{\mathbf{z} \in \mathcal{Z}} g(\mathbf{x}, \mathbf{z}) := \log \det \left(\sum_{i \in [n]} x_i e^{z_i} \mathbf{v}_i \mathbf{v}_i^\top \right). \quad (2)$$

For ease of notation, we define $f(\mathbf{x}) := \inf_{\mathbf{z} \in \mathcal{Z}} g(\mathbf{x}, \mathbf{z})$, the inner infimum of (2).

Similar but somewhat different convex programs have been studied by [7], [6], [37], [36]. (The relationship of our convex program to these also plays a crucial role in our analysis: see below.) The estimation algorithms in these works rely on a simple randomized algorithm to round a fractional optimal solution $\mathbf{x}^\star$. The analysis of the algorithm relies on a positive correlation property: the algorithm outputs a random solution such that all elements of an independent set S of size d are included with probability at least $\frac{1}{\alpha} \cdot \prod_{i \in S} x_i^*$, where α is some function of k. This property, combined with inequalities for real stable and completely log-concave polynomials, leads to an $\alpha \cdot e^{O(d)}$-estimation algorithm. We show that there exist fractional optimal solutions $\mathbf{x}^\star$ such that no rounding scheme has this positive correlation property for any α which is a function of d and independent of k. So, the dependence on k is inherent to all the previous algorithms which round an arbitrary optimal solution $\mathbf{x}^\star$ and do not consider the structure of the vectors to obtain some structure on the optimal $\mathbf{x}^\star$.

Our first technical result is to show that there always exists an optimal solution that has at most $O(d^2)$ fractional variables. We briefly describe how to obtain such a sparse optimal solution. Let $\mathbf{x}^\star$ denote an optimal solution to the convex program (similar reasoning works for near optimal solutions as well). We first show that, using a series of careful preprocessing steps, we can assume that there exists a $\mathbf{z}^\star$ attaining the infimum in $f(\mathbf{x}^\star) = \inf_{\mathbf{z} \in \mathcal{Z}} g(\mathbf{x}^\star, \mathbf{z})$. We then use first order optimality conditions that give a sufficient condition for another solution $\mathbf{x}$ to be optimal (i.e., to have $f(\mathbf{x}) = f(\mathbf{x}^\star)$). These conditions, however, present two significant obstacles: first, the conditions are not linear in $\mathbf{x}$, and, second, they ask for the existence of an exponentially sized dual solution as a certificate of optimality. We address the first problem by noticing that insisting that the entire matrix $\left(\sum_{i \in [n]} x_i^\star e^{z_i} \mathbf{v}_i \mathbf{v}_i^\top \right)$ does not change when $\mathbf{x}^\star$ changes to $\mathbf{x}$ leads to the optimality conditions becoming a system of linear equations in exponentially many variables. We then use the simple, yet elegant fact from matroid theory that minimum weight bases of a matroid under a linear weight function form the base set of another matroid. We use this combinatorial fact to observe that the existence of the exponentially sized dual solution is equivalent to insisting that a vector, whose coordinates are linear functions of $\mathbf{x}$, is in the base polytope of a new matroid. Putting all of this together reduces the search for the new optimal solution $\mathbf{x}$ to solving a system of exponentially many linear inequalities. Now, in the familiar territory of matroid polytopes, we apply standard uncrossing methods and show that every extreme point solution of the system of these linear inequalities has only $O(d^2)$ fractional variables.

Finally, we give a new randomized algorithm that gives an $O(d)^d$-estimation algorithm in the presence of $O(d^2)$ fractional variables. Since the objective is non-linear, we cannot just pick all variables set to 1 and apply a randomized algorithm to fractional elements. Indeed, the variables set to 1 must also be dropped from the final solution with certain probability. We show that given a solution $\mathbf{x}$ with at most $O(d^2)$ fractional values, our rounding scheme outputs a random solution such that for any independent set S of size d, all elements of S are picked with probability at least $(O(d))^{-d} \prod_{i \in S} x_i$. To show that this property implies the random solution output by the algorithm achieves an $O(d)^d$ approximation in expectation, we utilize recent and exciting work on strongly log-concave polynomials [8], [4] and the equivalent notion of Lorentzian polynomials [11]. While the analysis using strongly log-concave distributions naturally utilizes a different convex programming relaxation introduced in [6], the aforementioned sparsity result is not applicable to these convex programs. To this end, we show that the convex programming relaxation considered in our work is stronger than the convex programming relaxation from [6]. The relationship between the various convex programs for this problem and their respective strengths and weaknesses outlined by our results may be of independent interest.

D. Organization

In Section II, we discuss our convex relaxation, some technical issues in solving the relaxation, our main technical result, and the first order optimality conditions for the relaxation. In Section III, we show the existence of an optimal solution with at most $O(d^2)$ fractional values. In Section IV, we give the randomized algorithm to round a solution of the relaxation with few fractional values. In Section V, we give our deterministic approximation algorithm that gives a guarantee that only depends on d. We refer the reader to an extended version [29] for complete proofs.

E. Related Work

Uniform Matroid:: DETERMINANT MAXIMIZATION is NP-hard even for uniform matroid [39]. Koutis [26] showed that there exists a constant $c > 0$ such that it is NP-hard to approximate better than a factor of $(1+c)^d$ [26]. Let k be the rank of the uniform matroid. Bouhtou et al. [10] gave a $\left(\frac{n}{k}\right)^d$-approximation algorithm based on rounding the solution of a natural concave relaxation. Nikolov [32] improved the result to a e^k-approximation when $k \leq d$. Wang et al. [38] improved the approximation ratio of $(1+\epsilon)^d$ when $k \geq \frac{d^2}{\epsilon}$. Allen-Zhu et al. [3] improved the bound on k to give $(1+\epsilon)^d$-approximation when $k = \Omega\left(\frac{d}{\epsilon^2}\right)$. This was improved by Singh et al. [35] who gave a $(1+\epsilon)^d$-approximation when $k = \Omega\left(\frac{d}{\epsilon} + \frac{1}{\epsilon^2}\log\frac{1}{\epsilon}\right)$. Recently, this was improved by Madanet al. [30] who gave a $(1+\epsilon)^d$-approximation when $k \geq d + \frac{d}{\epsilon}$.

General Matroid:: Nikolov and Singh [33] gave a e^d-approximation for DETERMINANT MAXIMIZATION with partition matroid of rank d. Let k be the rank of a general matroid. Anari and Gharan [6] gave a e^{2k}-approximation when the generating polynomial for the matroid is real-stable. This corresponds to the Strongly Rayleigh matroids which includes both uniform and partition matroids. This was generalized by Anari, Gharan, and Vinzant [8] who gave a e^{2k}-approximation for any matroid. While the result in [8] is not stated for $k > d$, it can be easily deduced from the analysis. Algorithms in [33], [6], [8] are estimation algorithms as they estimate the optimum value up to a certain approximation factor, but do not yield a solution in polynomial time. Straszak and Vishnoi [37] gave a polynomial time algorithm without output a $O(\sqrt{k}e^k)$-approximate solution for partition matroid.

II. CONVEX PROGRAM AND OPTIMALITY CONDITIONS

Our algorithm for DETERMINANT MAXIMIZATION under a general matroid constraint is based on solving a convex relaxation and rounding an optimal solution of the convex relaxation to an integral solution. In this section, we formulate this convex relaxation, show that it is efficiently solvable, and prove some of its properties which are crucial for the rounding algorithm.

A. Formulation of the Convex Program

Let $V = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be input vectors. For a matroid $\mathcal{M} = ([n], \mathcal{I})$, we denote by $\mathcal{I}_s(\mathcal{M}) := \{S \in \mathcal{I} : |S| = s\}$ the set of all independent sets of size s. We denote by $\mathcal{P}(\mathcal{M})$ the matroid base polytope of $\mathcal{M}$, which is the convex hull of all of the bases. For any vector $\mathbf{z} \in \mathbb{R}^n$ of real numbers and a subset $S \subseteq [n]$, we let $z(S) := \sum_{i \in S} z_i$. We let $\mathcal{Z} := \{\mathbf{z} \in \mathbb{R}^n : \forall S \in \mathcal{I}_d(\mathcal{M}), z(S) \geq 0\}$. We introduce the optimization problem

$$\sup_{\mathbf{x} \in \mathcal{P}(\mathcal{M})} \inf_{\mathbf{z} \in \mathcal{Z}} g(\mathbf{x}, \mathbf{z}) := \log\det\left(\sum_{i \in [n]} x_i e^{z_i} \mathbf{v}_i \mathbf{v}_i^\top\right). \quad (3)$$

For ease of notation, we also let $f(\mathbf{x}) := \inf_{\mathbf{z} \in \mathcal{Z}} g(\mathbf{x}, \mathbf{z})$, the inner infimum of (3). The above program is a convex relaxation, as shown in Nikolov and Singh [33]. Unfortunately, it is not clear whether the outer supremum and inner infimum are attained at some $\mathbf{x}^\star$ and finite $\mathbf{z}^\star$. While the supremum over $\mathbf{x}$ can be approximated, our approach relies crucially on the inner infimum being achieved exactly at some finite $\mathbf{z}^\star$. We first show the following technical lemma that gives a sufficient condition for the infimum to be achieved based on KKT conditions and Slater's qualification of constraints. We say that the vectors $\{\mathbf{v}_i : i \in [n]\} \subseteq \mathbb{R}^d$ are in *general position* if any subset of size d is linearly independent.

Lemma II.1 *Let* $\mathbf{x} \in \mathcal{P}(\mathcal{M})$ *be such that* $\max_{i \in [n]} x_i < 1$ *and* $f(\mathbf{x}) = \inf_{\mathbf{z} \in \mathcal{Z}} g(\mathbf{x}, \mathbf{z})$ *is finite, and suppose that the vectors* $\{\mathbf{v}_i : i \in [n]\}$ *are in general position. Then,* $g(\mathbf{x}, \mathbf{z})$ *attains its infimum over* $\mathbf{z} \in \mathcal{Z}$ *at some* $\mathbf{z}^* \in \mathcal{Z}$.

In general, an instance of our problem may not satisfy the conditions of the lemma: the given vectors need not be in general position, and every optimal $\mathbf{x}$ may have value 1 on some coordinates. A preprocessing step is enough to show that both of these assumptions can be made with a slight loss in optimality by modifying the input instance [29]. This is achieved by modifying the matroid by introducing two parallel copies of each element as well as perturbing the vectors slightly to put them in general position. From here on, we assume that these modifications have been carried out, and we use $\mathcal{M}$ and V to denote the resulting matroid and vectors, respectively.

These reductions allow us to formulate the following stronger convex program where we place an additional upper bound on the coordinates of $\mathbf{x}$:

$$\sup_{\mathbf{x} \in \mathcal{P}(\mathcal{M}) \cap \left[0, \frac{1}{2}\right]^n} \inf_{\mathbf{z} \in \mathcal{Z}} g(\mathbf{x}, \mathbf{z}) := \log\det\left(\sum_{i \in [n]} x_i e^{z_i} \mathbf{v}_i \mathbf{v}_i^\top\right) \quad (4)$$

We denote the convex program (4) by CP, its optimum value by OPT_{CP}, and an optimal solution by $(\mathbf{x}^\star, \mathbf{z}^\star)$. We denote by OPT, the optimal value of the DETERMINANT MAXIMIZATION problem. Based on the discussion above, we show the following lemma, which also gives the polynomial time solvability of the convex program.

Lemma II.2 *For any* $\epsilon > 0$*, there is a polynomial time algorithm that returns* $\mathbf{x}^\star \in \mathcal{P}(\mathcal{M}) \cap \left[0, \frac{1}{2}\right]^n$ *such that* $\inf_{\mathbf{z} \in \mathcal{Z}} g(\mathbf{x}^\star, \mathbf{z}) \geq \log(\text{OPT}) - \epsilon$*. Moreover, there exists* $\mathbf{z}^\star$ *attaining the infimum in* $\inf_{\mathbf{z} \in \mathcal{Z}} g(\mathbf{x}^\star, \mathbf{z})$.

Our main algorithmic result is to show that the value of the convex program OPT_{CP} gives a good approximation of the optimal value OPT of the DETERMINANT MAXIMIZATION problem. The theorem below immediately implies Theorem I.1.

Theorem II.3 *The optimum value* $\mathrm{OPT}_{\mathrm{CP}}$ *of the convex program gives a* $(2e^5 d)^d$*-approximation to the value of the optimum, i.e.,*

$$\log(\mathrm{OPT}) - \epsilon \leq \mathrm{OPT}_{\mathrm{CP}} \leq \log(\mathrm{OPT}) + O(d \log d). \quad (5)$$

Moreover, there is a polynomial time algorithm that, given $\mathbf{x}^\star$ *attaining* $\mathrm{OPT}_{\mathrm{CP}}$ *and* $\mathbf{z}^\star$ *attaining the infimum in* $\inf_{\mathbf{z}\in\mathcal{Z}} g(\mathbf{x}^\star, \mathbf{z})$*, returns a random set* $S \in \mathcal{I}$ *such that*

$$\mathbb{E}\left[\det\left(\sum_{i\in S} \mathbf{v}_i \mathbf{v}_i^T\right)\right] \geq (2e^5 d)^{-d} (\mathrm{OPT}).$$

We now outline the ideas behind proving Theorem II.3. First, we obtain the KKT optimality conditions of $\inf_{\mathbf{z}\in\mathcal{Z}} g(\mathbf{x}, \mathbf{z})$ in Section II-B. In Section III, we show that the KKT conditions can be related to a new matroid defined by minimum weight bases of the original matroid under the weight function $\mathbf{z}^\star$. We then apply uncrossing methods on matroids to show that there is always an optimal sparse solution – in particular, one with at most $O(d^2)$ fractional variables. In Section IV, we give a rounding algorithm that uses the fact that number of fractional variables is bounded, and we prove Theorem II.3 building on inequalities proved in [6] and [8] for stable and completely log concave polynomials, respectively.

B. Optimality Conditions

Recall the notation $f(\mathbf{x}) = \inf_{\mathbf{z}\in\mathcal{Z}} g(\mathbf{x}, \mathbf{z})$. In the following result, we state a sufficient condition that some feasible solution $\hat{\mathbf{x}} \in \mathcal{P}(\mathcal{M})$ satisfies $f(\hat{\mathbf{x}}) = f(\mathbf{x}^\star)$, where $\mathbf{x}^\star$ is an (approximately) optimal solution to CP as returned by the algorithm in Lemma II.2. The result is obtained by applying the general KKT conditions to the optimization problem $\inf_{\mathbf{z}\in\mathcal{Z}} g(\mathbf{x}, \mathbf{z})$.

Lemma II.4 *Suppose* $\mathbf{x}^\star \in \mathcal{P}(\mathcal{M}) \cap \left[0, \frac{1}{2}\right]^n$ *is a feasible solution for* CP *such that the infimum over* $\mathcal{Z}$ *in* CP *is achieved, and let* $\mathbf{z}^\star \in \arg\min_{\mathbf{z}\in\mathcal{Z}} g(\mathbf{x}^\star, \mathbf{z})$*. For any* $\hat{\mathbf{x}} \in \mathcal{P}(\mathcal{M})$*, suppose that there exists* $\boldsymbol{\lambda} \in \mathbb{R}_{\geq 0}^{\mathcal{I}_d(\mathcal{M})}$ *such that*

1. *for all* $S \in \mathcal{I}_d(\mathcal{M})$ *with* $\mathbf{z}^\star(S) \neq 0$*, we have* $\lambda_S = 0$*,*
2. *for all* $i \in [n]$*, we have* $\hat{x}_i e^{z_i^\star} \mathbf{v}_i^\top \mathbf{X}^{-1} \mathbf{v}_i = \sum_{S\in\mathcal{I}_d(\mathcal{M}): i\in S} \lambda_S$ *where* $\mathbf{X} = \sum_{i=1}^n x_i^\star e^{z_i^\star} \mathbf{v}_i \mathbf{v}_i^\top$*, and*
3. $\sum_{i=1}^n x_i^\star e^{z_i^\star} \mathbf{v}_i \mathbf{v}_i^\top = \sum_{i=1}^n \hat{x}_i e^{z_i^\star} \mathbf{v}_i \mathbf{v}_i^\top$*.*

Then, $f(\hat{\mathbf{x}}) = f(\mathbf{x}^\star)$*. Moreover, there exists* $\boldsymbol{\lambda} \in \mathbb{R}_{\geq 0}^{\mathcal{I}_d(\mathcal{M})}$ *such that the above three conditions hold with* $\hat{\mathbf{x}} = \mathbf{x}^\star$*.*

We remark that the above criteria ask for the existence of exponentially sized vector $\boldsymbol{\lambda}$ in order to certify that $\hat{\mathbf{x}}$ is optimal. In the next section, we show that the above condition is equivalent to showing a certain vector is in the base polytope of another matroid derived from $\mathcal{M}$.

III. Small Support Solutions to CP

A. Preserving the Value of a Solution

In this section, we show that there is always an optimal solution to CP that has small number of fractional components. Indeed, given any solution $\mathbf{x}$ such that the inner infimum of CP is attained, we show how to obtain a sparse solution whose objective is no worse.

Theorem III.1 (Sparsity of an optimal solution) *Let* $\mathbf{x}^\star$ *be a solution to* CP *such that the inner infimum of* CP *is attained. Then there exists a solution* $\hat{\mathbf{x}} \in \mathcal{P}(\mathcal{M})$ *such that*

1) $f(\hat{\mathbf{x}}) = f(\mathbf{x}^\star)$*, and*
2) $|\{i \in [n] : 0 < \hat{x}_i < 1\}| \leq 2\left(\binom{d+1}{2} + d\right)$*.*

Moreover, such a solution $\hat{\mathbf{x}}$ *can be found in polynomial time.*

Proof: Given $\mathbf{x}^\star$, a solution to CP, we let $\mathbf{z}^\star$ be an optimal solution to $\inf_{\mathbf{z}\in\mathcal{Z}} g(\mathbf{x}^\star, \mathbf{z})$. Also, let $\mathbf{X} = \sum_{i\in[n]} x_i^\star e^{z_i^\star} \mathbf{v}_i \mathbf{v}_i^\top$. We assume that $\mathrm{supp}(\mathbf{x}^\star) = \{1, \ldots, n\}$ since for any i with $x_i^\star = 0$, we can update the instance by deleting these elements. Observe that this does not effect the optimality (restricted to $\mathrm{supp}(\mathbf{x}^\star)$) of $\mathbf{z}^\star$.

We first give a simpler description than Lemma II.4 for a solution $\hat{\mathbf{x}}$ to have an objective better than $f(\mathbf{x}^\star)$. This relies on the following basic lemma.

Lemma III.2 *Let* $\mathcal{B}^\star = \{S \in \mathcal{I}_d : z^\star(S) = 0\}$*. Then,* $\mathcal{B}^\star$ *is a basis of another matroid* $\mathcal{M}^\star = ([n], \mathcal{I}^\star)$*. Additionally, if* $\mathcal{M}$ *admits an independence oracle, then* $\mathcal{M}^\star$ *also admits an independence oracle.*

Proof: Since $z^\star(S) \geq 0$ for all $S \in \mathcal{I}_d$, the basis of $\mathcal{I}_d$ included in $\mathcal{I}^\star$ are the minimum weight bases under the weight function $\mathbf{z}$. Minimum weight bases of a matroid form the bases of another matroid, and the independence oracle can be implemented in polynomial time. ■

We now have the following simpler description for $\hat{\mathbf{x}}$ to be optimal building on Lemma II.4. Let $\mathcal{M}^\star$ be the matroid in Lemma III.2 and let $r^\star : 2^{[n]} \to \mathbb{Z}_+$ denote the rank function of $\mathcal{M}^\star$.

Lemma III.3 *Let* $\mathbf{x}^\star$ *be a solution of* CP *and* $\mathbf{z}^\star \in \arg\min_{\mathbf{z}\in\mathcal{Z}} g(\mathbf{x}^\star, \mathbf{z})$*. Let* $\hat{\mathbf{x}} \in \mathbb{R}^{[n]}$ *be such that*

1) $\hat{\mathbf{x}} \in \mathcal{P}(\mathcal{M})$*,*
2) *the vector* $\mathbf{w} \in \mathbb{R}^{[n]}$ *defined as* $w_i = \hat{x}_i e^{z_i^\star} \mathbf{v}_i^\top \mathbf{X}^{-1} \mathbf{v}_i$ *for each* $i \in [n]$ *satisfies* $\mathbf{w} \in \mathcal{P}(\mathcal{M}^\star)$*, where* $\mathbf{X} = \sum_{i\in[n]} x_i^\star e^{z_i^\star} \mathbf{v}_i \mathbf{v}_i^\top$*,*
3) $\sum_{i\in[n]} \hat{x}_i e^{z_i^\star} \mathbf{v}_i \mathbf{v}_i^\top = \sum_{i\in[n]} x_i^\star e^{z_i^\star} \mathbf{v}_i \mathbf{v}_i^\top$*, and*
4) $\mathrm{supp}(\hat{\mathbf{x}}) \subseteq \mathrm{supp}(\mathbf{x}^\star)$.

Then $f(\hat{\mathbf{x}}) = f(\mathbf{x}^\star)$.

Proof: We show that the above conditions imply that the conditions of Lemma II.4 are satisfied. Indeed, we only

$$\min 0 \tag{6}$$

$$s.t. \quad \sum_{i \in S} x_i \le r(S) \;\forall\; \emptyset \subsetneq S \subsetneq [n] \tag{7}$$

$$x([n]) = r([n]) = k \tag{8}$$

$$\sum_{i \in S} x_i e^{z_i^\star} \mathbf{v}_i^\top \mathbf{X}^{-1} \mathbf{v}_i \le r^\star(S) \forall\; \emptyset \subsetneq S \subsetneq [n] \tag{9}$$

$$\sum_{i \in [n]} x_i e^{z_i^\star} \mathbf{v}_i^\top \mathbf{X}^{-1} \mathbf{v}_i = r^\star([n]) = d \tag{10}$$

$$\sum_{i=1}^n x_i e^{z_i^\star} \mathbf{v}_i \mathbf{v}_i^\top = \sum_{i=1}^n x_i^\star e^{z_i^\star} \mathbf{v}_i \mathbf{v}_i^\top \tag{11}$$

$$x_i \ge 0 \qquad \forall i \in [n] \tag{12}$$

Figure 1. Linear program to obtain a sparse solution.

need to show the existence of $\boldsymbol{\lambda} \in \mathbb{R}^{\mathcal{I}_d(\mathcal{M})}$ as claimed. Since $\mathbf{w} \in \mathbb{R}^{[n]}$ is in $\mathcal{P}(\mathcal{M}^\star)$, we have $\mathbf{w} = \sum_{S \in \mathcal{B}(\mathcal{M}^\star)} \mu_S \chi_S$ where $\chi_S \in \mathbb{R}^{[n]}$ is the indicator vector of set S and $\sum_{S \in \mathcal{B}(\mathcal{M}^\star)} \mu_S = 1$. Observe that for each $S \in \mathcal{B}(\mathcal{M}^\star)$, we have $z^\star(S) = 0$. Thus, setting $\lambda_S = \mu_S$ for $S \in \mathcal{B}(\mathcal{M}^\star)$ and $\lambda_S = 0$ for all other sets in $\mathcal{I}_d(\mathcal{M})$ satisfies the conditions of Lemma II.4. ■

Now the above conditions can be formulated as a feasibility system over the following linear constraints as given in Figure 1, and we call the formulated linear program $\text{LP}_{\text{x-OPT}}$. Here, constraints (7)-(8) insist that $\mathbf{x} \in \mathcal{P}(\mathcal{M})$ and (9)-(10) insist that the vector $(x_i e^{z_i^\star} \mathbf{v}_i \mathbf{X}^{-1} \mathbf{v}_i)_{i \in [n]} \in \mathcal{P}(\mathcal{M}^\star)$. Constraints (11) insist that the matrix $\mathbf{X}$ does not change when the solution changes to $\mathbf{x}$ from $\mathbf{x}^\star$. For ease of notation, we let $\mathbf{w}_x$ be the vector $(x_i e^{z_i^\star} \mathbf{v}_i \mathbf{X}^{-1} \mathbf{v}_i)_{i \in [n]}$.

From basic uncrossing methods we obtain the following lemma characterizing any extreme point of the above linear program. Recall that a collection $\mathcal{C}$ of sets is a *chain* if for all $A, B \in \mathcal{C}$, we have $A \subseteq B$ or $B \subseteq A$. Again, we focus on $\text{supp}(\mathbf{x})$ since $\mathbf{x}$ remains extreme after removing coordinates with $x_i = 0$. Thus, we assume that $[n] = \text{supp}(\mathbf{x})$.

Lemma III.4 *If $\mathbf{x}$ is an extreme point of the linear program* $\text{LP}_{\text{x-OPT}}$*, then there exist chains $\mathcal{C}_1, \mathcal{C}_2 \subseteq 2^{[n]}$ and $P \subseteq [d] \times [d]$ such that*

1) *$x(S) = r(S)$ for each $S \in \mathcal{C}_1$, $w_x(S) = r^\star(S)$ for each $S \in \mathcal{C}_2$, and $(\sum_{i=1}^n x_i e^{z_i^\star} \mathbf{v}_i \mathbf{v}_i^\top)_{jk} = (\sum_{i=1}^n x_i^\star e^{z_i^\star} \mathbf{v}_i \mathbf{v}_i^\top)_{jk}$ for each $(j,k) \in P$,*
2) *the linear constraints corresponding to sets in $\mathcal{C}_1, \mathcal{C}_2$ and pairs in P are linearly independent, and*
3) *$|\text{supp}(\mathbf{x})| = |\mathcal{C}_1| + |\mathcal{C}_2| + |P|$.*

Let $\mathbf{x}$ be an extreme point of the linear program $\text{LP}_{\text{x-OPT}}$. Such an $\mathbf{x}$ can be found in polynomial time. Let $\mathcal{C}_1 = \{S_1, \ldots, S_l\}$ where $S_1 \subset S_2 \ldots \subset S_l$. Then, we have $x(S_i) = r(S_i)$. Since $x_i > 0$ for all $i \in [n]$, we have $1 \le r(S_1) < r(S_2) \ldots < r(S_l) \le k$ and from the integrality of the rank function, we obtain that $|\mathcal{C}_1| = l \le k$. Similarly, $|\mathcal{C}_2| \le r^\star([n]) \le d$, and clearly $|P| \le \binom{d+1}{2}$ since $\sum_{i=1}^n x_i e^{z_i} \mathbf{v}_i \mathbf{v}_i^\top$ and $\sum_{i=1}^n x_i^\star e^{z_i^\star} \mathbf{v}_i \mathbf{v}_i^\top$ are $d \times d$ symmetric matrices. Therefore, $\text{supp}(\mathbf{x}) \le k + d + \binom{d+1}{2}$. In what follows we argue all but $2\left(d + \binom{d+1}{2}\right)$ coordinates are set to 1.

For ease of notation, we let $S_0 = \emptyset$. Observe that if $|S_j \setminus S_{j-1}| = 1$ for any $1 \le j \le l$, say $\{i\} = S_j \setminus S_{j-1}$, then $x_i = x(S_j) - x(S_{j-1}) = r(S_j) - r(S_{j-1})$ which is an non-negative integer. Since $x_i > 0$, we obtain that $x_i = 1$. Let $I = \{1 \le j \le k : |S_j \setminus S_{j-1}| = 1\}$. Observe that there are at least $|I|$ variables set to 1. But since every set S_j with $j \notin I$ contains at least two elements in $S_j \setminus S_{j-1}$, we have

$$|\text{supp}(\mathbf{x})| \ge |I| + 2(l - |I|).$$

But from Lemma III.4, we have

$$|\text{supp}(\mathbf{x})| \le l + d + \binom{d+1}{2}.$$

Combining the two inequalities, we get $|I| \ge l - d - \binom{d+1}{2} \ge |\text{supp}(\mathbf{x})| - 2(d + \binom{d+1}{2})$. This implies that the number of fractional variables is at most $\text{supp}(x) - |I| \le 2(d + \binom{d+1}{2})$. ■

IV. Randomized Rounding Algorithm

In this section, we give our randomized rounding algorithm and prove the guarantee on its performance claimed in Theorem II.3.

Throughout this section, we assume that the algorithm receives an input $\mathbf{x} \in \mathcal{P}(\mathcal{M})$ such that

$$|\{i : 0 < x_i < 1\}| \le 2\left(\binom{d+1}{2} + d\right).$$

We first describe the rounding algorithm, presented in Algorithm 1. It is obvious that Algorithm 1 runs in polynomial time.

For ease of notation we denote $\gamma = (2e^3 d)^{-d}$ and $\mathcal{I}_d = \mathcal{I}_d(\mathcal{M})$. We first claim that every independent subset S of $R_1 \cup R_2$ of size d is contained in the output set with probability at least γ. The claim can only be true if the ground set $R_1 \cup R_2$, which has been restricted to the support of $\mathbf{x}$, is small.

Lemma IV.1 *Let T denote the random set returned by Algorithm 1. Then, for any set $S \subseteq R_1 \cup R_2$ such that $S \in \mathcal{I}_d$, we have*

$$\mathbb{P}[S \subseteq T] \ge \gamma.$$

Lemma IV.1 implies a lower bound on the expected objective value of the solution returned.

Algorithm 1 Rounding Algorithm

Input: a matroid $\mathcal{M} = ([n], \mathcal{I})$, $\mathbf{x} \in \mathcal{P}(\mathcal{M})$.
Output: a set $S \in \mathcal{I}$.
procedure ROUNDING($x, \mathcal{I}$)
  $R_1 \leftarrow \{i : 0 < x_i < 1\}, R_2 \leftarrow \{i : x_i = 1\}$
  $T \leftarrow \emptyset$
  for i in R_1 **do**
    if $T \cup \{i\} \in \mathcal{I}$ **then**
      $T \leftarrow T \cup \{i\}$ with probability $\frac{1}{d}$
  for i in R_2 **do**
    if $T \cup \{i\} \in \mathcal{I}$ **then**
      $T \leftarrow T \cup \{i\}$ with probability $\frac{1}{2}$
  if T is not a basis **then**
    Extend T to a basis (e.g. by going through each element in $[n] \setminus T$ and add it to T if T remains independent until T is a basis)
  return T

Lemma IV.2 *Algorithm 1 returns an independent set $T \in \mathcal{I}$ with expected objective value*

$$\mathbb{E}\left[\det\left(\sum_{i\in T}\mathbf{v}_i\mathbf{v}_i^\top\right)\right] \geq \gamma \sum_{S\in\mathcal{I}_d}\det\left(\sum_{i\in S}x_i\mathbf{v}_i\mathbf{v}_i^\top\right).$$

Next, we relate this lower bound to the objective of the convex relaxation CP in a two-step procedure. Building on results by [8], the lower bound on the expected objective of the algorithm can be bounded in terms of objective of a different convex relaxation as described in Lemma IV.3. Proof of the lemma is inferred from the inequality proved in [8] on completely log-concave polynomials by observing that the polynomials (in $\mathbf{y}$ and $\mathbf{z}$ variables) $\det\left(\sum_{i=1}^n x_i^\star y_i\mathbf{v}_i\mathbf{v}_i^\top\right)$ and $\sum_{S\in\mathcal{I}_d} z^{[n]\setminus S}$ are completely log-concave. Here we use the notation $\left(\frac{\mathbf{yw}}{\boldsymbol{\alpha}}\right)^{\boldsymbol{\alpha}} := \prod_{i=1}^n\left(\frac{y_iw_i}{\alpha_i}\right)^{\alpha_i}$.

Lemma IV.3 *For any $\mathbf{x}^\star \geq 0$,*

$$\sum_{S\in\mathcal{I}_d}\det\left(\sum_{i\in S}x_i^\star\mathbf{v}_i\mathbf{v}_i^\top\right) \geq e^{-2d}\sup_{\boldsymbol{\alpha}\in P(\mathcal{I}_d)}\inf_{\mathbf{y},\mathbf{w}>0}\frac{\det\left(\sum_{i=1}^n x_i^\star y_i\mathbf{v}_i\mathbf{v}_i^\top\right)\left(\sum_{S\in\mathcal{I}_d}w^S\right)}{\left(\frac{\mathbf{yw}}{\boldsymbol{\alpha}}\right)^{\boldsymbol{\alpha}}}.$$

To finish the proof of Theorem II.3, we show that the convex relaxation CP is stronger than the convex relaxation studied in [8]. The proof of the following lemma is deferred to the full version.

Lemma IV.4 *For any $\mathbf{x}^\star \geq 0$,*

$$\sup_{\boldsymbol{\alpha}\in P(\mathcal{I}_d)}\inf_{\mathbf{y},\mathbf{w}>0}\frac{\det\left(\sum_{i=1}^n x_i^\star y_i\mathbf{v}_i\mathbf{v}_i^\top\right)\left(\sum_{S\in\mathcal{I}_d}w^S\right)}{\left(\frac{\mathbf{yw}}{\boldsymbol{\alpha}}\right)^{\boldsymbol{\alpha}}} \geq \inf_{\mathbf{z}\in\mathcal{Z}}\det\left(\sum_{i=1}^n x_i^\star e^{z_i}\mathbf{v}_i\mathbf{v}_i^\top\right).$$

Note that we cannot directly use the convex relaxation of [8] and avoid the two-step procedure for our problem. Algorithm 1 and the proof of Lemma IV.1 require that the solution $\mathbf{x}$ is sparse, and we do not know if such property holds true for the convex relaxation of [8].

Before we prove Lemmas IV.1 and IV.2, we present the proof of the main result of our paper.

Proof of Theorem II.3: We first show (5). Recall that $f(\mathbf{x}) = \inf_{\mathbf{z}\in\mathcal{Z}}\log\det\left(\sum_{i\in[n]}x_ie^{z_i}\mathbf{v}_i\mathbf{v}_i^\top\right)$. Let $(\mathbf{x}^\star, \mathbf{z}^\star)$ be an optimal solution to CP, so we have $f(\mathbf{x}^\star) = \text{OPT}_{\text{CP}}$. The first inequality of (5) follows from Lemma II.2. It remains to show the second inequality.

By Theorem III.1, there exists $\hat{\mathbf{x}} \in \mathcal{P}(\mathcal{M})$ such that $f(\hat{\mathbf{x}}) = f(\mathbf{x}^\star)$ and $|\{i \in [n] \mid 0 < \hat{x}_i < 1\}| \leq 2\left(\binom{d+1}{2} + d\right)$. Let $T \in \mathcal{I}$ be the random solution returned by Algorithm 1 given an input $\mathbf{x} = \hat{\mathbf{x}}$. We apply Lemmas IV.2, IV.3, and IV.4 successively and in this order to get

$$\mathbb{E}\left[\det\left(\sum_{i\in T}\mathbf{v}_i\mathbf{v}_i^\top\right)\right] \geq (2e^3d)^{-d}\sum_{S\in\mathcal{I}_d}\det\left(\sum_{i\in S}\hat{x}_i\mathbf{v}_i\mathbf{v}_i^\top\right)$$
$$\geq (2e^3d)^{-d}e^{-2d}. \tag{13}$$
$$\sup_{\boldsymbol{\alpha}\in P(\mathcal{I}_d)}\inf_{\mathbf{y},\mathbf{w}>0}\frac{\det\left(\sum_{i=1}^n \hat{x}_i y_i\mathbf{v}_i\mathbf{v}_i^\top\right)\left(\sum_{S\in\mathcal{I}_d}w^S\right)}{\left(\frac{\mathbf{yw}}{\boldsymbol{\alpha}}\right)^{\boldsymbol{\alpha}}}$$
$$\geq (2e^5d)^{-d}\inf_{\mathbf{z}\in\mathcal{Z}}\det\left(\sum_{i=1}^n\hat{x}_ie^{z_i}\mathbf{v}_i\mathbf{v}_i^\top\right)$$
$$= (2e^5d)^{-d}\inf_{\mathbf{z}\in\mathcal{Z}}\det\left(\sum_{i=1}^n x_i^\star e^{z_i}\mathbf{v}_i\mathbf{v}_i^\top\right) = (2e^5d)^{-d}\cdot\text{OPT}_{\text{CP}} \tag{14}$$

where the first of the two equalities follows from Theorem III.1.

On the other hand, for any $T \in \mathcal{I}$, we have $\det\left(\sum_{i\in T}\mathbf{v}_i\mathbf{v}_i^\top\right) \leq \text{OPT}$, and therefore

$$\mathbb{E}\left[\det\left(\sum_{i\in T}\mathbf{v}_i\mathbf{v}_i^\top\right)\right] \leq \text{OPT}. \tag{15}$$

Combining (14) and (15) proves the second inequality of (5).

Given a solution $\mathbf{x}^\star$ and $\mathbf{z}^\star$ attaining the infimum in $\inf_{\mathbf{z}\in\mathcal{Z}}\det\left(\sum_{i=1}^n x_i^\star e^{z_i}\mathbf{v}_i\mathbf{v}_i^\top\right)$, the efficiency of the randomized algorithm that satisfies (14) follows from the efficiency of obtaining a sparse solution (by Theorem III.1) and of the rounding Algorithm 1.

Now we prove Lemmas IV.1 and IV.2.

Proof of Lemma IV.1: We need to prove that for any $S \subseteq R_1 \cup R_2$ such that $S \in \mathcal{I}_d$,

$$\mathbb{P}[S \subseteq T] \geq (2e^3 d)^{-d}.$$

Let $S_1 = S \cap R_1$ and $S_2 = S \cap R_2$. Since $x_i = 1$ for any $i \in R_2$ and $\mathbf{x} \in \mathcal{P}(\mathcal{M})$, we have $R_2 \in \mathcal{I}$. Since $S \in \mathcal{I}$ and $S_1 \subseteq S$, we have $S_1 \in \mathcal{I}$. We first claim the following.

Claim IV.5 *There exists $Y \subseteq R_2 \setminus S_2$ such that $|Y| \leq |S_1|$ and $S_1 \cup (R_2 \setminus Y) \in \mathcal{I}$.*

Proof: Recall that $S = S_1 \cup S_2 \in \mathcal{I}$ and $R_2 \in \mathcal{I}$. If $|R_2| \leq |S_1 \cup S_2|$, then $Y = R_2$ satisfies the condition. Else, by the definition of matroids, there exists an element $i \in R_2 \setminus (S_1 \cup S_2)$ such that $S_1 \cup S_2 \cup \{i\} \in \mathcal{I}$. Since $S_1 \cap R_2 = \emptyset$, we get that $i \in R_2 \setminus S_2$. Repeating this process for $|R_2| - |S_1 \cup S_2|$ times, we obtain a set $W \subseteq R_2 \setminus S_2$ of size $|R_2| - |S_1 \cup S_2|$ such that $S_1 \cup S_2 \cup W \in \mathcal{I}$. If $Y = R_2 \setminus (S_2 \cup W)$, then $S_1 \cup (R_2 \setminus Y) = S_1 \cup S_2 \cup W \in \mathcal{I}$. Since W has size $|R_2| - |S_1 \cup S_2|$, Y has size $|R_2| - (|R_2| - |S_1 \cup S_2|) - |S_2| = |S_1 \cup S_2| - |S_2| \leq |S_1|$. ■

Let $Y \subseteq R_2 \setminus S_2$ be a set such that $S_1 \cup (R_2 \setminus Y) \in \mathcal{I}$. Next, we prove a lower bound on $\mathbb{P}[S \subseteq T]$. Note that S is a disjoint union of S_1 and S_2. Hence,

$$\begin{aligned}
\mathbb{P}[S \subseteq T] &= \mathbb{P}[S_1 \subseteq T \text{ and } S_2 \subseteq T] \\
&\geq \mathbb{P}[T \cap R_1 = S_1 \text{ and } S_2 \subseteq T] \\
&\geq \mathbb{P}[T \cap Y = \emptyset \text{ and } T \cap R_1 = S_1 \text{ and } S_2 \subseteq T] \\
&= \mathbb{P}[T \cap Y = \emptyset] \cdot \mathbb{P}[T \cap R_1 = S_1 \mid T \cap Y = \emptyset] \cdot \\
&\quad \mathbb{P}[S_2 \subseteq T \mid T \cap Y = \emptyset, T \cap R_1 = S_1]
\end{aligned}$$

Next, we lower bound each of the probabilities.

1) $\mathbb{P}[T \cap Y = \emptyset]$**:** Consider the event that $T \cap Y = \emptyset$. It happens if for each $i \in Y$, i is not added to T during the execution of the algorithm. Let T' be the set T before the iteration considering i. If $T' \cup \{i\} \notin \mathcal{I}$, i is not added to T with probability 1. If $T' \cup \{i\} \in \mathcal{I}$, i is not added to T with probability $1/2$. Hence, for each $i \in T$, i is not added to T with probability at least $1/2$. Probability that none of the elements of Y are added to T is therefore at least $\left(\frac{1}{2}\right)^{|Y|}$. Since $|Y| \leq |S_1|$,

$$\mathbb{P}[Y \cap T = \emptyset] \geq \left(\frac{1}{2}\right)^{|S_1|}.$$

2) $\mathbb{P}[T \cap R_1 = S_1 \mid T \cap Y = \emptyset]$**:** Since all elements of R_1 are considered before the elements of R_2 (and hence Y), we have

$$\mathbb{P}[T \cap R_1 = S_1 \mid T \cap Y = \emptyset] = \mathbb{P}[T \cap R_1 = S_1]. \quad (16)$$

To get $T \cap R_1 = S_1$, we must have $(R_1 \setminus S_1) \cap T = \emptyset$ and $S_1 \subseteq T$. Hence,

$$\begin{aligned}
\mathbb{P}[T \cap R_1 = S_1] = \mathbb{P}[T \cap (R_1 \setminus S_1) = \emptyset] \cdot \\
\mathbb{P}[S_1 \subset T \mid T \cap (R_1 \setminus S_1) = \emptyset]. \quad (17)
\end{aligned}$$

As argued above, for any element $i \in R_1 \setminus S_1$, the probability that i is not in T (regardless of other elements) is at least $1 - \frac{1}{d}$. Hence, $\mathbb{P}[T \cap (R_1 \setminus S_1)] \geq \left(1 - \frac{1}{d}\right)^{|R_1 \setminus S_1|}$ which is equal to $\left(1 - \frac{1}{d}\right)^{|R_1| - |S_1|}$ since $S_1 \subseteq R_1$. Since $S \in \mathcal{I}$ and $S_1 \subseteq S$, we have $S_1 \in \mathcal{I}$. Consider an element $i \in S_1$ and the set T' being the set T before the algorithm processes the element i. If no element of $R_1 \setminus S_1$ is picked, then $T' \cup \{i\} \subset S_1$. Hence, $T' \cup \{i\} \in \mathcal{I}$, and the probability that the element i is picked is $\frac{1}{d}$. Hence, if no element of $R_1 \setminus S_1$ is picked, then every element of S_1 is picked with probability $\frac{1}{d}$. This implies that

$$\mathbb{P}[S_1 \subseteq T \mid T \cap (R_1 \setminus S_1) = \emptyset] = \left(\frac{1}{d}\right)^{|S_1|}. \quad (18)$$

Combining (16)-(18), we get

$$\mathbb{P}[T \cap R_1 = S_1 \mid T \cap Y = \emptyset] \geq \left(1 - \frac{1}{d}\right)^{|R_1| - |S_1|} \left(\frac{1}{d}\right)^{|S_1|}.$$

3) $\mathbb{P}[S_2 \subseteq T \mid T \cap Y = \emptyset, T \cap R_1 = S_1]$**:** Consider an element $i \in S_2$. Let T' be the set T just before the algorithm considers the element i. If $T' \cap R_1 = S_1$ and $T' \cap Y = \emptyset$, then $T' \cup \{i\} \subseteq S_1 \cup (R_2 \setminus Y)$. By Claim IV.5, $S_1 \cup (R_2 \setminus Y) \in \mathcal{I}$. Hence, if $T' \cap R_1 = S_1$ and $T' \cap Y = \emptyset$. Then, $T' \cup \{i\} \in \mathcal{I}$, and i is added to T with probability $1/2$. Therefore,

$$\mathbb{P}[S_2 \subseteq T \mid T \cap Y = \emptyset, T \cap R_1 = S_1] = \left(\frac{1}{2}\right)^{|S_2|}.$$

Combining the bounds on the three probabilities, we get

$$\mathbb{P}[S \subseteq T] \geq \left(\frac{1}{2}\right)^{|S_1|} \left(1 - \frac{1}{d}\right)^{|R_1| - |S_1|} \left(\frac{1}{d}\right)^{|S_1|} \left(\frac{1}{2}\right)^{|S_2|}.$$

Since $|S| = d$ and S is a disjoint union of S_1 and S_2, we have $|S_1| + |S_2| = d$. Also, by the assumption of the theorem, $|R_1| \leq 2\left(\binom{d+1}{2} + d\right)$. Hence,

$$\mathbb{P}[S \subseteq T] \geq \left(\frac{1}{2}\right)^{d} \left(1 - \frac{1}{d}\right)^{2\left(\binom{d+1}{2} + d\right)} \left(1 - \frac{1}{d}\right)^{-|S_1|} \left(\frac{1}{d}\right)^{|S_1|}.$$

Since $|S_1| \leq d$, we have

$$\begin{aligned}
\mathbb{P}[S \subseteq T] &\geq \left(\frac{1}{2}\right)^{d} \left(1 - \frac{1}{d}\right)^{2\left(\binom{d+1}{2} + d\right)} \left(1 - \frac{1}{d}\right)^{-d} \left(\frac{1}{d}\right)^{d} \\
&= \left(\frac{1}{2d}\right)^{d} \left(1 - \frac{1}{d}\right)^{d(d+1)+d}
\end{aligned}$$

For $d \geq 2$, we have $1 - \frac{1}{d} \geq e^{-\frac{1.5}{d}} \geq e^{-\frac{3}{d+2}}$. Hence,

$$\mathbb{P}[S \subseteq T] \geq (2d)^{-d} e^{-3d} = \left(2e^3 d\right)^{-d}$$

finishing the proof of Lemma IV.1 □

Proof of Lemma IV.2: By Lemma IV.1, for any $S \subseteq R_1 \cup R_2$ such that $S \in \mathcal{I}_d$, we have $\mathbb{P}[S \subset T] \geq \left(2e^3 d\right)^{-d}$. The rounding Algorithm 1 returns a solution T of expected value

$$\mathbb{E}\left[\det\left(\sum_{i \in T} \mathbf{v}_i \mathbf{v}_i^\top\right)\right] = \mathbb{E}\left[\sum_{S \in \binom{T}{d}} \det\left(\sum_{i \in S} \mathbf{v}_i \mathbf{v}_i^\top\right)\right]$$
$$= \sum_{S \subseteq [n]: |S| = d} \mathbb{P}\left[S \subseteq T\right] \det\left(\sum_{i \in S} \mathbf{v}_i \mathbf{v}_i^\top\right)$$

where we apply the Cauchy-Binet formula to obtain the first equality. Since we only pick elements of $R_1 \cup R_2$ which form an independent set, we have

$$\mathbb{E}\left[\det\left(\sum_{i \in T} \mathbf{v}_i \mathbf{v}_i^\top\right)\right] =$$
$$\sum_{S \subseteq R_1 \cup R_2 : S \in \mathcal{I}_d} \mathbb{P}\left[S \subseteq T\right] \det\left(\sum_{i \in S} \mathbf{v}_i \mathbf{v}_i^\top\right)$$
$$\geq \left(2e^3 d\right)^{-d} \sum_{S \subseteq R_1 \cup R_2 : S \in \mathcal{I}_d} \det\left(\sum_{i \in S} \mathbf{v}_i \mathbf{v}_i^\top\right).$$

For each $i \in [n]$, we have $0 \leq x_i \leq 1$. Hence,

$$\mathbb{E}\left[\det\left(\sum_{i \in T} \mathbf{v}_i \mathbf{v}_i^\top\right)\right] \geq$$
$$\left(2e^3 d\right)^{-d} \sum_{S \subseteq R_1 \cup R_2 : S \in \mathcal{I}_d} \det\left(\sum_{i \in S} x_i \mathbf{v}_i \mathbf{v}_i^\top\right).$$

For $S \in \mathcal{I}_d$ such that $S \not\subseteq R_1 \cup R_2$, there exists $i \in S$ such that $x_i = 0$. Hence, for $S \in \mathcal{I}_d$ such that $S \not\subseteq R_1 \cup R_2$, $\sum_{i \in S} x_i \mathbf{v}_i \mathbf{v}_i^\top$ has rank at most $d - 1$ and $\det\left(\sum_{i \in S} x_i \mathbf{v}_i \mathbf{v}_i^\top\right) = 0$. Therefore,

$$\mathbb{E}\left[\det\left(\sum_{i \in T} \mathbf{v}_i \mathbf{v}_i^\top\right)\right] \geq \left(2e^3 d\right)^{-d} \sum_{S \in \mathcal{I}_d} \det\left(\sum_{i \in S} x_i \mathbf{v}_i \mathbf{v}_i^\top\right)$$

finishing the proof of Lemma IV.2. □

V. Deterministic Algorithm

In this section, we prove Theorem I.2 and give the deterministic algorithm achieving the claimed guarantee. The algorithm reduces the ground set in each iteration until the ground set is itself an independent set. Given any $V \subseteq [n]$, we let $\mathcal{M}_{|V} = (V, \mathcal{I}_{|V})$ denote the matroid obtained by deleting all elements not in V from $\mathcal{M}$. Moreover, we let CP(V) denote the convex program when the ground set and the matroid are V and $\mathcal{M}_{|V}$, respectively, and we consider only vectors indexed by V. We let OPT_{CP}(V) denote optimal value of the convex program CP(V). We denote by $r(V)$ the rank of the matroid $\mathcal{M}_{|V}$.

We first describe the deterministic rounding algorithm, presented in Algorithm 2.

Algorithm 2 Deterministic Algorithm

```
1: Input: a matroid M = ([n], I).
2: Output: a basis S ∈ I.
3: procedure ROUNDING
4:     Let x be optimal solution to CP such that |{i ∈ [n] : 0 < x_i}| ≤ k + 2(C(d+1,2) + d) as returned by Theorem III.1.
5:     Let V ← {i ∈ [n] : 0 < x_i}.
6:     while V ∉ I do
7:         i ← arg max_{j∈V : r(V\{j})=r(V)} OPT_CP(V \ {j}) (breaking a tie arbitrarily)
8:         V ← V \ {i}
       return V
```

Observe that V is initialized to a set of size at most $k + 2\left(\binom{d+1}{2} + d\right)$ along with $r(V) = k$. Moreover, $\text{OPT}_{\text{CP}}(V) = \text{OPT}_{\text{CP}}$ initially, since we just remove all elements with $x_i = 0$ from the ground set.

In each iteration of the while loop, we decrease the size of V by one, and thus there can be at most $2\left(\binom{d+1}{2} + d\right)$ iterations of the while loop. In each iteration, we do not decrease the rank of V from k, so the final output, by construction, is an independent set of size k and hence feasible. To prove the guarantee, we show that in each iteration,

$$\text{OPT}_{\text{CP}}(V \setminus \{i\}) \geq \beta \cdot \text{OPT}_{\text{CP}}(V) \tag{19}$$

where $\beta = (2e^5 d)^{-d} = O(d)^{-d}$. Also, the relaxation is exact after the last iteration because V is a basis after the while loop terminates. Thus, the objective value of the returned solution is at least

$$\beta^{2(\binom{d+1}{2}+d)} \cdot \text{OPT}_{\text{CP}},$$

giving an approximation factor $O(d)^{2d(\binom{d+1}{2}+d)} = O(d)^{d^3} \cdot O(1)^{3d^2 \log d} = O(d)^{d^3}$, as claimed.

It only remains to prove (19). From the guarantee of the randomized algorithm given in Theorem II.3, there *exists* a basis $S \in \mathcal{I}_{|V}$ with $S \subseteq V$ such that

$$\det\left(\sum_{j \in S} \mathbf{v}_j \mathbf{v}_j^\top\right) \geq \beta \cdot \text{OPT}_{\text{CP}}(V). \tag{20}$$

Let $j \in V \setminus S$ where j must exist since $V \notin \mathcal{I}$. Then $r(V \setminus \{j\}) = r(S) = k$ since S is a basis. We have $\text{OPT}_{\text{CP}}(V \setminus \{j\}) \geq \det\left(\sum_{e \in S} \mathbf{v}_e \mathbf{v}_e^\top\right)$, because the indicator vector $\mathbf{x}$ of $S \subseteq V \setminus \{j\}$ is a solution to CP($V \setminus \{j\}$) of value $\det\left(\sum_{e\in S} \mathbf{v}_e \mathbf{v}_e^\top\right)$. Together with (20), and because i is chosen to maximize $\mathrm{OPT}_{\mathrm{CP}}(V \setminus \{j\})$ over j s.t. $r(V \setminus \{j\}) = k$, we have established (19). This completes the proof of Theorem I.2.

Acknowledgements

Part of the work was done when Vivek Madan and Uthaipon (Tao) Tantipongpipat were at Georgia Institute of Technology. Uthaipon (Tao) Tantipongpipat and Mohit Singh were supported by Supported by NSF-AF:1910423 and NSF-AF:1717947. Aleksandar Nikolov is supported by NSERC Discovery Grant.

References

[1] Z. Allen-Zhu, Y. Li, A. Singh, and Y. Wang. Near-optimal design of experiments via regret minimization. In D. Precup and Y. W. Teh, editors, *Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6-11 August 2017*, volume 70 of *Proceedings of Machine Learning Research*, pages 126–135. PMLR, 2017.

[2] Z. Allen-Zhu, Y. Li, A. Singh, and Y. Wang. Near-optimal discrete optimization for experimental design: A regret minimization approach. *arXiv preprint arXiv:1711.05174*, 2017.

[3] Z. Allen-Zhu, Y. Li, A. Singh, and Y. Wang. Near-optimal discrete optimization for experimental design: A regret minimization approach. *arXiv preprint arXiv:1711.05174*, 2017.

[4] N. Anari, K. Liu, S. Oveis Gharan, and C. Vinzant. Log-concave polynomials ii: high-dimensional walks and an fpras for counting bases of a matroid. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 1–12. ACM, 2019.

[5] N. Anari, T. Mai, S. Oveis Gharan, and V. V. Vazirani. Nash social welfare for indivisible items under separable, piecewise-linear concave utilities. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2274–2290. SIAM, 2018.

[6] N. Anari and S. Oveis Gharan. A generalization of permanent inequalities and applications in counting and optimization. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 384–396. ACM, 2017.

[7] N. Anari, S. Oveis Gharan, A. Saberi, and M. Singh. Nash social welfare, matrix permanent, and stable polynomials. *In Proceedings of Conference on Innovations in Theoretical Computer Science*, 2016.

[8] N. Anari, S. Oveis Gharan, and C. Vinzant. Log-concave polynomials, entropy, and a deterministic approximation algorithm for counting bases of matroids. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 35–46. IEEE, 2018.

[9] S. Barman, S. K. Krishnamurthy, and R. Vaish. Finding fair and efficient allocations. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 557–574. ACM, 2018.

[10] M. Bouhtou, S. Gaubert, and G. Sagnol. Submodularity and randomized rounding techniques for optimal experimental design. *Electronic Notes in Discrete Mathematics*, 36:679–686, 2010.

[11] P. Brändén and J. Huh. Lorentzian polynomials. *arXiv preprint arXiv:1902.03719*, 2019.

[12] I. Caragiannis, D. Kurokawa, H. Moulin, A. D. Procaccia, N. Shah, and J. Wang. The unreasonable fairness of maximum nash welfare. In *EC*, 2016.

[13] J. Carbonell and J. Goldstein. The use of mmr, diversity-based reranking for reordering documents and producing summaries. In *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 335–336. ACM, 1998.

[14] L. E. Celis, A. Deshpande, T. Kathuria, D. Straszak, and N. K. Vishnoi. On the complexity of constrained determinantal point processes. In *APPROX-RANDOM*, volume 81 of *LIPIcs*, pages 36:1–36:22. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017.

[15] L. E. Celis, V. Keswani, D. Straszak, A. Deshpande, T. Kathuria, and N. K. Vishnoi. Fair and diverse dpp-based data summarization. In *ICML*, volume 80 of *Proceedings of Machine Learning Research*, pages 715–724. PMLR, 2018.

[16] A. Cevallos, F. Eisenbrand, and R. Zenklusen. Local search for max-sum diversification. In P. N. Klein, editor, *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, Hotel Porta Fira, January 16-19*, pages 130–142. SIAM, 2017.

[17] H. Chen and D. R. Karger. Less is more: probabilistic models for retrieving fewer relevant documents. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 429–436. ACM, 2006.

[18] A. Çivril and M. Magdon-Ismail. Exponential inapproximability of selecting a maximum volume sub-matrix. *Algorithmica*, 65(1):159–176, 2013.

[19] R. Cole, N. Devanur, V. Gkatzelis, K. Jain, T. Mai, V. V. Vazirani, and S. Yazdanbod. Convex program duality, fisher markets, and nash social welfare. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pages 459–460. ACM, 2017.

[20] M. Di Summa, F. Eisenbrand, Y. Faenza, and C. Moldenhauer. On largest volume simplices and sub-determinants. In *SODA*, pages 315–323. SIAM, 2015.

[21] J. B. Ebrahimi, D. Straszak, and N. K. Vishnoi. Subdeterminant maximization via nonconvex relaxations and anti-concentration. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1020–1031. Ieee, 2017.

[22] B. Gong, W.-L. Chao, K. Grauman, and F. Sha. Diverse sequential subset selection for supervised video summarization. In *Advances in Neural Information Processing Systems*, pages 2069–2077, 2014.

[23] L. Gourvès and J. Monnot. On maximin share allocations in matroids. *Theor. Comput. Sci.*, 754:50–64, 2019.

[24] L. Gourvès, J. Monnot, and L. Tlilane. A matroid approach to the worst case allocation of indivisible goods. In *IJCAI*, pages 136–142. IJCAI/AAAI, 2013.

[25] L. Gourvès, J. Monnot, and L. Tlilane. Near fairness in matroids. In *ECAI*, volume 263 of *Frontiers in Artificial Intelligence and Applications*, pages 393–398. IOS Press, 2014.

[26] I. Koutis. Parameterized complexity and improved inapproximability for computing the largest j-simplex in a v-polytope. *Information Processing Letters*, 100(1):8–13, 2006.

[27] A. Kulesza and B. Taskar. Determinantal point processes for machine learning. *Foundations and Trends® in Machine Learning*, 5(2–3):123–286, 2012.

[28] H. Li, S. Patterson, Y. Yi, and Z. Zhang. Maximizing the number of spanning trees in a connected graph. *IEEE Transactions on Information Theory*, 2019.

[29] V. Madan, A. Nikolov, M. Singh, and U. Tantipongpipat. Maximizing determinants under matroid constraints. *arXiv preprint arXiv:2004.07886*, 2020.

[30] V. Madan, M. Singh, U. Tantipongpipat, and W. Xie. Combinatorial algorithms for optimal design. In *Conference on Learning Theory*, pages 2210–2258, 2019.

[31] H. Moulin. *Fair division and collective welfare*. MIT press, 2004.

[32] A. Nikolov. Randomized rounding for the largest simplex problem. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, pages 861–870. ACM, 2015.

[33] A. Nikolov and M. Singh. Maximizing determinants under partition constraints. In *ACM symposium on Theory of computing*, pages 192–201, 2016.

[34] F. Pukelsheim. *Optimal design of experiments*. SIAM, 2006.

[35] M. Singh and W. Xie. Approximate positive correlated distributions and approximation algorithms for D-optimal design. *In Proceedings of SODA*, 2018.

[36] D. Straszak and N. K. Vishnoi. Belief propagation, bethe approximation and polynomials. In *Communication, Control, and Computing (Allerton), 2017 55th Annual Allerton Conference on*, pages 666–671. IEEE, 2017.

[37] D. Straszak and N. K. Vishnoi. Real stable polynomials and matroids: optimization and counting. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 370–383. ACM, 2017.

[38] Y. Wang, A. W. Yu, and A. Singh. On computationally tractable selection of experiments in regression models. *arXiv preprint arXiv:1601.02068*, 2016.

[39] W. J. Welch. Algorithmic complexity: three np-hard problems in computational statistics. *Journal of Statistical Computation and Simulation*, 15(1):17–25, 1982.

[40] C. Zhai, W. W. Cohen, and J. D. Lafferty. Beyond independent relevance: Methods and evaluation metrics for subtopic retrieval. *SIGIR Forum*, 49(1):2–9, 2015.