Dependent Randomized Rounding via Exchange Properties of Combinatorial Structures

(Extended Abstract)

Chandra Chekuri*
Dept. of Computer Science
University of Illinois
Urbana, IL 61801, USA
Email: chekuri@cs.illinois.edu

Jan Vondrák
IBM Almaden Research Center
San Jose, CA 95120, USA
Email: jvondrak@us.ibm.com

Rico Zenklusen†
Dept. of Mathematics
MIT
Cambridge, MA 02139, USA
Email: ricoz@math.mit.edu

***Abstract*—We consider the problem of randomly rounding a fractional solution x in an integer polytope $P \subseteq [0,1]^n$ to a vertex X of P, so that $\mathbf{E}[X] = x$. Our goal is to achieve *concentration properties* for linear and submodular functions of the rounded solution. Such dependent rounding techniques, with concentration bounds for linear functions, have been developed in the past for two polytopes: the assignment polytope (that is, bipartite matchings and b-matchings) [32], [19], [23], and more recently for the spanning tree polytope [2]. These schemes have led to a number of new algorithmic results.**

In this paper we describe a new *swap rounding* technique which can be applied in a variety of settings including *matroids* and *matroid intersection*, while providing Chernoff-type concentration bounds for linear and submodular functions of the rounded solution. In addition to existing techniques based on negative correlation, we use a martingale argument to obtain an exponential tail estimate for monotone submodular functions. The rounding scheme explicitly exploits *exchange properties* of the underlying combinatorial structures, and highlights these properties as the basis for concentration bounds.

Matroids and matroid intersection provide a unifying framework for several known applications [19], [23], [7], [22], [2] as well as new ones, and their generality allows a richer set of constraints to be incorporated easily. We give some illustrative examples, with a more comprehensive discussion deferred to a later version of the paper.

I. Introduction

Randomized rounding is a fundamental technique in approximation algorithms. It was introduced by Raghavan and Thompson [28] for the purpose of rounding a fractional solution x to a linear programming (LP) relaxation of a problem, into an integral solution. The original technique from [28] (and several subsequent papers) relies on *independent* rounding of the variables. This allows one to use Chernoff bounds for linear functions of the variables; these bounds are critical for several applications in packing and covering problems. However, there are situations in which independent rounding is not feasible due to the presence of constraints that must not be violated by the rounded solution. Various techniques are used to handle such scenarios. To name just a few: alteration of solutions obtained by independent rounding, careful derandomization or constructive methods when the probability of a feasible solution is non-zero but small (for example using the Lovász Local Lemma), and various forms of dependent randomized rounding schemes. In some dependent randomized techniques, concentration bounds similar to Chernoff bounds are still available. See for example the work of Arora, Frieze and Kaplan [1], Srinivasan [32] (and several subsequent generalizations and applications [19], [23], [29], [5]), Doerr [14], [15], the work of Asadpour and Saberi [3], and Asadpour et al.'s [2] recent breakthrough on the asymmetric traveling salesman problem (ATSP).

Our focus in this paper is on the following broad class of dependent rounding schemes. Given a fractional solution x in an integer polytope $P \subseteq [0,1]^n$, randomly round x to a vertex X of P such that $\mathbf{E}[X] = x$ and concentration bounds hold for linear (or even more general) functions of X. The polytope P, via its *vertices*, captures the constraints that the rounded solution needs to preserve. For example, Srinivasan [32] considered the polytope $P = \{x \in [0,1]^n : \sum_{i=1}^n x_i = k\}$. That is, we wish to round $x = (x_1, \ldots, x_n) \in P$ to a random integral $X = (X_1, \ldots, X_n)$ such that exactly k of the X_i are set to 1, and moreover, for *any* vector $a \in [0,1]^n$, the linear function $\sum_i a_i X_i$ is concentrated around its mean. I.e., we want Chernoff-type concentration bounds as if the X_i's were independent, even though the variables satisfy a deterministic constraint. A dependent randomized rounding technique is employed in [32], and it is shown that the variables are *negatively correlated*.[1] Negative correlation implies Chernoff-type concentration bounds for linear functions of the variables, as shown earlier by Panconesi and Srinivasan [27]. It is notable that this seemingly simple setting already results in a variety of non-trivial applications [32].

Gandhi et al. [19] employed dependent randomized rounding in the more general setting of an assignment polytope

*Partially supported by NSF grant CCF-0728782.

†Partially supported by the Swiss National Science Foundation, grant number: PBEZP2-129524. Partially developed while working at EPFL.

[1] A collection of $\{0,1\}$-random variables $X_1, \ldots, X_r$ are negatively correlated if, for all subsets $T \subseteq [r]$, $\mathbf{E}[\prod_{i \in T} X_i] \leq \prod_{i \in T} \mathbf{E}[X_i]$ and $\mathbf{E}[\prod_{i \in T}(1 - X_i)] \leq \prod_{i \in T}(1 - \mathbf{E}[X_i])$.

(bipartite matchings and b-matchings). They call it randomized *pipage rounding* following a deterministic variant that was proposed earlier by Ageev and Sviridenko [4]. For the assignment polytope it is no longer the case that negative correlation holds for all subsets of variables, but only for those subsets that contain only the edge variables incident to some vertex of the underlying graph. Nonetheless, this is sufficient for several applications [19], [23], [29]. In recent work, Saha and Srinivasan [29] proposed a general scheme for randomized iterated rounding following the ideas in [23]; the scheme applies in several settings but is too broad to directly give concentration bounds.

A different approach to dependent rounding is *maximum entropy sampling*. Asadpour and Saberi [3] used this technique for a resource allocation problem. More recently, this technique was used by Asadpour et al. [2] in their work on the ATSP problem to round a fractional solution in the spanning tree polytope so that the variables are negatively correlated. Unfortunately, a natural generalization of this approach to matroids does not always provide negative correlation — see [31] and a discussion below.

The success of dependent rounding schemes raises several natural questions. For what polytopes can one implement a rounding procedure that guarantees concentration properties? Is there a framework unifying the known results for the assignment polytope and the spanning tree polytope? Can we obtain concentration bounds for non-linear (in particular submodular) functions under such rounding? In this paper we describe a conceptually new approach that leads to a variety of general results.

Our contribution. We propose the following generic approach to rounding a fractional solution $x \in P$.

The Randomized Swap Rounding Scheme:

- Express x as a convex combination of vertices of P, that is, $x = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$ where each $\mathbf{v}_i$ is a vertex of P and $\sum_i \alpha_i = 1$.
- Let $\mathbf{w}_1 = \mathbf{v}_1$, $\beta_1 = \alpha_1$, and in each stage, merge two vertices $\mathbf{w}_i, \mathbf{v}_{i+1}$ into a new vertex $\mathbf{w}_{i+1}$ with coefficient $\beta_{i+1} = \beta_i + \alpha_{i+1}$ in such a way that $\mathbf{E}[\beta_{i+1}\mathbf{w}_{i+1}] = \beta_i \mathbf{w}_i + \alpha_{i+1}\mathbf{v}_{i+1}$ (**the merge operation**).
- After $n-1$ stages, obtain a vertex $X = \mathbf{w}_n$ such that $\mathbf{E}[X] = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i = x$.

This procedure can be viewed as a random walk in the polytope, or a vector-valued martingale. The properties of this procedure naturally depend on how the merge operation is implemented. One solution is to pick either $\mathbf{w}_{i+1} := \mathbf{w}_i$ or $\mathbf{w}_{i+1} := \mathbf{v}_{i+1}$ with probabilities proportional to β_i, α_{i+1}, but this procedure would not have any interesting properties in general; this is equivalent to picking, in a single step, exactly one of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ where the probability of picking $\mathbf{v}_i$ is α_i. Our goal is to introduce suitable randomness in the merge operation, so that concentration bounds can be proved for certain functions of X.

Our main contribution is in identifying a broad class of polytopes where such concentration bounds can be achieved. We show a simple rounding procedure for *matroid polytopes*, and more generally even for *matroid intersection polytopes*, while satisfying exponential concentration properties (with some limitations in the case of matroid intersection). Since the spanning tree polytope is a special case of a matroid polytope, and the assignment polytope is a special case of matroid intersection, several of the aforementioned applications [19], [23], [2] can be derived as special cases of our rounding scheme. Beyond matroid intersection, our technique can be also applied to *non-bipartite* graph matchings and b-matchings, but we defer details to a future work. The common feature of our implementations of the merge operation is that it proceeds in a sequence of random swaps of elements. Therefore we call our procedure *randomized swap rounding*. This swapping is guided by *exchange properties* of the underlying combinatorial structures. Our main technical results are concentration bounds for our rounding procedure in various settings. We present our technical results in Section II.

Polyhedral descriptions, integrality, and total dual integrality of the matroid and matroid intersection polytopes due to Edmonds [13] were fundamental achievements in combinatorial optimization in the 60's and 70's. Matroids and matroid intersection can be used to capture a variety of constraints, and our results apply to any combinatorial structure of this type. For applications involving monotone submodular functions, we employ the continuous greedy algorithm developed in [33] to approximately maximize the multilinear extension F of a monotone submodular set function f over a solvable polytope. We also extend this algorithm to simultaneously handle *multiple* submodular functions. These ingredients provide a powerful toolkit that can be used to derive a number of applications. For instance, we obtain an optimal $(1-1/e-\varepsilon)$-approximation for the Max-Min allocation problem with submodular valuations, for any constant number of agents. See Section III for more details.

Negative correlation. Previous methods for dependent rounding, including pipage rounding methods and maximum entropy sampling, rely on negative correlation. The work of [27] shows that negative correlation implies concentration for linear functions. Hence, negative correlation is a powerful property for a dependent randomized rounding procedure to satisfy. We show that in some sense, matroids are *exactly* the setting where this property can be achieved.

Theorem I.1. *Let P be a polytope with vertices in $\{0,1\}^N$. Then the following two properties are equivalent:*

1) *For any $x \in P$, there exists a probability distribution over vertices of P such that a random vertex X drawn from this distribution satisfies $\mathbf{E}[X] = x$ and the coordinates $\{X_i\}_{i \in N}$ are negatively correlated.*
2) *P is a projection of a matroid base polytope, in the sense that there is a matroid $\mathcal{M} = (N', \mathcal{I})$ such that $N \subseteq N'$ and X is a vertex of P iff $X = \mathbf{1}_{B \cap N}$ for some base B of $\mathcal{M}$.*

The implication 2) $\Rightarrow$ 1) follows from our swap rounding

procedure, as we prove in this paper. We defer the proof of the other implication to a full version of this paper. The above implies that negative correlation cannot be achieved for all subsets of coordinates beyond the setting of matroids. For matroid intersection we can still achieve negative correlation for special subsets of coordinates that we call *equivalence classes* in a matroid. This generalizes the fact that for bipartite matchings and b-matchings, negative correlation can be achieved for subsets of edges incident to a vertex as in [19].

Beyond negative correlation. In order to prove tail estimates for submodular functions, we need to resort to other techniques. We use a martingale argument tailored to our rounding procedure to show a lower-tail bound for monotone submodular functions on matroid polytopes (Theorem II.2). Our concentration bounds are *dimension-free*, i.e. independent of the size of the ground set n, as opposed to bounds with variance growing with n, which can be obtained easily from known martingale results.

Swap rounding, unlike pipage rounding, allows us to apply martingale methods more easily for the following reason. Recall that the scheme first computes an explicit convex decomposition $x = \sum_i \alpha_i \mathbf{v}_i$. The coefficients α_i are fixed throughout the process, and moreover, the merge step between two vertices is combinatorial, and typically has a well-defined number of steps. We are able to exploit these features to bound the total variance of the process, as a function of the expectation rather than the number of elements n. To prove a lower-tail bound for submodular functions, we charge the variance of successive rounding steps in a careful way to the current value of the multilinear extension of the submodular function. Since the random process here is a sub-martingale, we do not expect a similar upper-tail bound.

Discussion of prior techniques. In [7], [34] randomized pipage rounding was generalized to the setting of matroids. Our work here implies that randomized pipage rounding also satisfies negative correlation. However, our tail estimate for submodular functions (Theorem II.2) relies on the structure of swap rounding and we do not know how to prove it for pipage rounding.

In their recent work on an $O(\log n/\log\log n)$ approximation for the ATSP problem, Asadpour et al. [2] used *maximum entropy sampling* to find a "thin spanning tree" via a fractional solution in the spanning tree polytope. Swap rounding for matroid polytopes, when specialized to the spanning tree polytope, can be used as an alternative subroutine for finding a thin spanning tree in [2]. Compared to maximum entropy sampling, swap rounding has several advantages; it is easy to describe and implement, and more efficient. In maximum entropy sampling [2], negative correlation arises from the structure of spanning trees in a way that does not extend to all matroids. For an example of a matroid where the maximum-entropy distribution over bases does not satisfy negative correlations, see [31].

In the more general setting of matroid intersection, randomized swap rounding is more complicated to describe and analyze. We are not aware of any prior dependent rounding technique for matroid intersection other than for the special case of bipartite matching [19]. The extension to matroid intersection requires several ingredients, in particular, a decomposition property for feasible swaps in two matroids from recent work of Lee et al. [25].

Arora, Frieze and Kaplan [1] described a rounding scheme based on alternating cycles for bipartite matchings. They obtained concentration for *any* linear function of the rounded solution, however in a somewhat weaker form, with variance depending explicitly on n. Since their focus was on approximation of dense instances, these concentration results were sufficient for their purpose. Their scheme also does not maintain the property that $\mathbf{E}[X] = x$, and in particular does not produce a perfect matching even when x is a fractional perfect matching; one can easily show that the expectation cannot be preserved if one wants concentration properties. Nevertheless, there are interesting parallels between the two approaches which we explore in some recent work [11].

Organization. Due to space constraints, and the large number of technical details in our rounding scheme and concentration bounds, we are only able to give a high-level overview of our applications and defer a detailed discussion to a longer version of the paper. An earlier unpublished version of our work [10] considered only matroid polytopes and describes some of these applications in more detail.

The rest of the paper is organized as follows. In Section I-A, we present the necessary definitions. In Section II, we describe our technical results and in Section III, a few illustrative applications. In Section IV the randomized swap rounding procedure for matroids is introduced. In Section V, we extend this procedure to the setting of matroid intersection. In Section VI, we discuss the connection between negative correlation and increasing submodular expectations, and describe a type of random process which satisfies both these properties.

A. Preliminaries

Matroid polytopes. Given a matroid $\mathcal{M} = (N, \mathcal{I})$ with rank function $r : 2^N \to \mathbb{Z}_+$, the matroid polytope $P(\mathcal{M})$ is defined as the convex hull of characteristic vectors of the independent sets of $\mathcal{M}$, or equivalently [13]:

$$P(\mathcal{M}) = \text{conv}\{\mathbf{1}_I : I \in \mathcal{I}\} = \{x \geq 0 : \forall S; \sum_{i \in S} x_i \leq r(S)\}.$$

A related polytope is the *matroid base polytope* $B(\mathcal{M})$, the convex hull of the characteristic vectors of bases, i.e. independent sets of maximum cardinality in $\mathcal{M}$. Equivalently, $B(\mathcal{M}) = \{x \in P(\mathcal{M}) : \sum_{i \in N} x_i = r(N)\}$. Both these polytopes have the property of *total dual integrality* (see [30] for more information).

Matroid exchange properties. To simplify notation, we use $+$ and $-$ for the addition and deletion of single elements from a set, for example $S - i + j$ denotes the set $(S \setminus \{i\}) \cup \{j\}$. The following base exchange property of matroids is crucial in our rounding for matroid polytopes.

Lemma I.2. *Let $\mathcal{M} = (N, \mathcal{I})$ be a matroid and let B_1, B_2 be bases. For any $i \in B_1 \setminus B_2$ there exists $j \in B_2 \setminus B_1$ such that $B_1 - i + j$ and $B_2 - j + i$ are also bases.*

Given i, a suitable element j for the exchange can be found by checking all elements in $B_2 \setminus B_1$. Faster algorithms exist in special cases such as graphic matroids.

Matroid intersection. Let $\mathcal{M}_1 = (N, \mathcal{I}_1)$ and $\mathcal{M}_2 = (N, \mathcal{I}_2)$ be two matroids. It follows from the work of Edmonds [13] that the intersection polytope $P(\mathcal{M}_1) \cap P(\mathcal{M}_2)$ is integral and its vertices correspond exactly to sets $I \in \mathcal{I}_1 \cap \mathcal{I}_2$.[2] The structure of feasible exchanges between two independent sets $I, J \in \mathcal{I}_1 \cap \mathcal{I}_2$ is more complicated here. We give more details in Section V.

Equivalent elements in matroids. In our treatment of matroid intersection, the following notion is important.

Definition I.3. *Two elements $i, j \in N$ are* equivalent *in a matroid $\mathcal{M} = (N, \mathcal{I})$ if for any set $A \in \mathcal{I}$ not containing i, j, $A + i \in \mathcal{I}$ if and only if $A + j \in \mathcal{I}$.*

It can be shown that this is in fact an equivalence relation (we omit the proof). Observe that for a partition matroid, the equivalence classes with respect to this relation are exactly the respective parts. More generally, for a laminar matroid, the equivalence classes are the minimal sets in the laminar family. The special case of bipartite matchings corresponds to the intersection of two partition matroids, where each equivalence class is a set of edges incident to a fixed vertex.

Submodular functions. A function $f : 2^N \to \mathbb{R}$ is submodular if for any $A, B \subseteq N$, $f(A)+f(B) \geq f(A \cup B)+f(A \cap B)$. In addition, f is monotone if $f(S) \leq f(T)$ whenever $S \subseteq T$. We denote by $f_A(i) = f(A + i) - f(A)$ the *marginal value* of i with respect to A. An important concept in recent work on submodular functions [6], [33], [7], [22], [24], [34] is the *multilinear extension* of a submodular function:

$$F(x_1, \ldots, x_n) = \mathbf{E}[f(\hat{x})] = \sum_{S \subseteq N} f(S) \prod_{i \in S} x_i \prod_{i \in N \setminus S} (1 - x_i)$$

where $\hat{x}$ is obtained by independently rounding $x_1, \ldots, x_n$ to $\{0, 1\}^n$. (We use interchangeably the notation $f(S)$ and $f(X_1, \ldots, X_n)$ where $X_i \in \{0, 1\}$ are interpreted as indicator variables of $i \in S$.)

Algorithmic issues. In this paper, submodular functions are assumed to be given by a *value oracle*, returning $f(S)$ for a given query $S \subseteq N$. For a matroid $\mathcal{M} = (N, \mathcal{I})$ we assume that there is a *rank oracle* that given a set $S \subseteq N$ returns $r_{\mathcal{M}}(S)$. We appeal to well-known facts about polynomial time algorithms for optimizing and separating over matroid and matroid intersection polytopes [30].

[2]The same is no longer true for three matroids, and the maximization problem over the intersection of three matroids is NP-hard.

II. Our Results

Rounding for matroids. For matroid polytopes, we present a simple rounding procedure, based on the strong exchange property of matroids (see Section IV). We start with a fractional solution $(x_1, \ldots, x_n) \in P(\mathcal{M})$ and produce a random vertex $X = (X_1, \ldots, X_n)$ which corresponds to an independent set in $\mathcal{M}$. The main property of the rounded solution X is that its coordinates have expectations $\mathbf{E}[X_i] = x_i$, and the X_i's are *negatively correlated*. This yields Chernoff-type concentration bounds for *any* linear function of $X_1, \ldots, X_n$ ([27], see also Theorem 3.1 in [19]). We also show an expectation bound for submodular functions, similar to what holds for pipage rounding [7]; this is useful in applications involving submodular functions.

Theorem II.1. *Let $(x_1, \ldots, x_n) \in P(\mathcal{M})$ be a fractional solution in the matroid polytope and $(X_1, \ldots, X_n) \in P(\mathcal{M}) \cap \{0, 1\}^n$ an integral solution obtained using randomized swap rounding. Then for any linear function $Z = \sum a_i X_i$ with $a_i \in [0, 1]$, we have $\mathbf{E}[Z] = \sum a_i x_i$ and*

- *If $\delta \geq 0$ and $\mu \geq \mathbf{E}[Z]$, then*
$$\Pr[Z \geq (1 + \delta)\mu] \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}} \right)^\mu.$$
- *If $\delta \in [0, 1]$, and $\mu \leq \mathbf{E}[Z]$, then*
$$\Pr[Z \leq (1 - \delta)\mu] \leq e^{-\mu\delta^2/2}.$$

Moreover, for any submodular function $f : \{0, 1\}^n \to \mathbb{R}$ and its multilinear extension $F : [0, 1]^n \to \mathbb{R}$,

- $\mathbf{E}[f(X_1, \ldots, X_n)] \geq F(x_1, \ldots, x_n)$.

An interesting technical point is that the property of negative correlations is easily *implied* by the expectation bound for submodular functions. We discuss this relationship in Section VI-A. For monotone submodular functions we also prove an exponential bound for the lower tail. Since the random process here is a sub-martingale rather than a martingale, we do not expect a similar bound for the upper tail.

Theorem II.2. *Let $f : \{0, 1\}^n \to \mathbb{R}_+$ be a monotone submodular function with marginal values in $[0, 1]$, and $F : [0, 1]^n \to \mathbb{R}_+$ its multilinear extension. Let $(x_1, \ldots, x_n) \in P(\mathcal{M})$ be a point in a matroid polytope and $(X_1, \ldots, X_n) \in \{0, 1\}^n$ a random solution obtained from it by randomized swap rounding. Let $\mu_0 = F(x_1, \ldots, x_n)$ and $\delta > 0$. Then*

$$\Pr[f(X_1, \ldots, X_n) \leq (1 - \delta)\mu_0] \leq e^{-\mu_0\delta^2/8}.$$

This result is more involved than Theorem II.1 and does not follow from the property of negative correlations alone. We defer the proof to a full version of the paper. We remark that weaker tail estimates involving a dependence on n follow directly from martingale concentration bounds; the main difficulty here is to obtain a bound which does not depend on n. The tail estimate is with respect to the value of the starting point, $\mu_0 = F(x_1, \ldots, x_n)$, rather than the actual expectation of $f(R)$, which could be larger (it would be equal for a linear function f). For this reason, we do not have an upper tail bound. In applications, μ_0 is the value that we wish to compare with, and hence this is the bound that we need.

Rounding for matroid intersection. We generalize our rounding procedure to the framework of matroid intersection. Here, we have two matroids $\mathcal{M}_1$, $\mathcal{M}_2$ and a fractional solution in the intersection of the respective polytopes, $x \in P(\mathcal{M}_1) \cap P(\mathcal{M}_2)$. We round x to an integral solution X corresponding to a set independent in both matroids, so that $\mathbf{E}[X] = x$.

We develop a more involved version of randomized swap rounding for this setting, which is described in Section V. Our procedure uses a decomposition property for matroid intersection developed recently in [25]. Contrary to the setting of a single matroid, we cannot expect all the coordinates $X_1, \ldots, X_n$ to be negatively correlated here. Even in the special case of bipartite matchings, positive correlation between certain variables is inevitable if we wish to preserve the expectation exactly or even approximately. A simple example that demonstrates this point is that of a cycle of length $2n$ with $x_e = 1/2$ on each edge. We prove that our rounding procedure for matroid intersection has the property of negative correlations for any subset of elements equivalent in either $\mathcal{M}_1$ or $\mathcal{M}_2$ (see Section I-A for a definition). This is a natural generalization of the property of negative correlation for edges incident to a vertex in the setting of bipartite b-matchings [19].

Theorem II.3. *Let $(X_1, \ldots, X_n)$ be obtained by randomized swap rounding for matroid intersection from a starting point $(x_1, \ldots, x_n) \in P(\mathcal{M}_1) \cap P(\mathcal{M}_2)$. Let Q be a set of elements equivalent either in $\mathcal{M}_1$ or $\mathcal{M}_2$. Then for any linear function $Z = \sum_{i \in Q} a_i X_i$ with $a_i \in [0,1]$, $\mathbf{E}[Z] = \sum_{i \in Q} a_i x_i$ and*

- *If $\delta \geq 0$ and $\mu \geq \mathbf{E}[Z]$, then*
$$\Pr[Z \geq (1+\delta)\mu] \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu.$$
- *If $\delta \in [0,1]$, and $\mu \leq \mathbf{E}[Z]$, then*
$$\Pr[Z \leq (1-\delta)\mu] \leq e^{-\mu\delta^2/2}.$$

Moreover, for any submodular function $f : \{0,1\}^Q \to \mathbb{R}$ and its multilinear extension $F : [0,1]^Q \to \mathbb{R}$,

- $\mathbf{E}[f(X_i : i \in Q)] \geq F(x_i : i \in Q)$.

This also means that our rounding procedure can be used in conjunction with a submodular objective function which decomposes into submodular functions on the equivalence classes of $\mathcal{M}_1$ and $\mathcal{M}_2$. It is known that the multilinear optimization problem $\max\{F(x) : x \in P\}$ can be approximated within a factor of $1 - 1/e$ for any monotone submodular function and any down-monotone polytope for which we can optimize linear functions (a "solvable polytope") [33], [7]. The matroid intersection polytope $P = P(\mathcal{M}_1) \cap P(\mathcal{M}_2)$ is solvable and hence we get a $(1 - 1/e)$-approximation for maximizing this restricted type of submodular functions over matroid intersection.

Theorem II.4. *Let $\mathcal{M}_1 = (N, \mathcal{I}_1)$, $\mathcal{M}_2 = (N, \mathcal{I}_2)$ be two matroids and $\mathcal{C}_1, \mathcal{C}_2 \subset 2^N$ their respective families of equivalence classes. Let $f : \{0,1\}^N \to \mathbb{R}_+$ be such that*

$$f(S) = \sum_{C \in \mathcal{C}_1} f_{1,C}(S \cap C) + \sum_{C \in \mathcal{C}_2} f_{2,C}(S \cap C)$$

where $f_{1,C}, f_{2,C}$ are monotone submodular. Then there is a $(1 - 1/e)$-approximation for $\max\{f(S) : S \in \mathcal{I}_1 \cap \mathcal{I}_2\}$.

A useful special case is that one of the matroids, say $\mathcal{M}_1$, is a partition or laminar matroid. Then the bounds in Theorem II.3 hold for any part Q (or minimal set in the laminar family) of this matroid $\mathcal{M}_1$. We can also achieve a $(1-1/e)$-approximation for any sum of monotone submodular functions on the parts of $\mathcal{M}_1$ (Theorem II.4). This is a new result even in the special case of bipartite b-matchings (two partition matroids), where we can handle any sum of monotone submodular functions on the stars incident to different vertices. We also recover the concentration bounds of [19] in this case. However, the matroid $\mathcal{M}_2$ can be arbitrary; this allows us to impose new types of constraints.

Multiple submodular objectives. In our applications, we often deal with submodular objective functions. Here we rely on the notion of multilinear extension and the continuous greedy algorithm [6], [33], [7]. The problem $\max\{F(x) : x \in P\}$ admits a $(1 - 1/e)$-approximation if F is the multilinear extension of a submodular set function $f : \{0,1\}^N \to \mathbb{R}_+$ and $P \subseteq [0,1]^{|N|}$ is a solvable down-monotone polytope [33], [7]. In this work, we extend the continuous greedy algorithm to applications involving multiple submodular functions. This extension might be interesting on its own; we formulate it as follows. In the following statement, n does not have to be constant. We defer the proof to a full version of the paper.

Theorem II.5. *Consider monotone submodular functions $f_1, \ldots, f_n : 2^N \to \mathbb{R}_+$, their multilinear extensions $F_i : [0,1]^N \to \mathbb{R}_+$ and a solvable polytope $P \subseteq [0,1]^N$. There is an algorithm which, given $V_1, \ldots, V_n \in \mathbb{R}_+$,*

- *either finds a point $x \in P$ such that $F_i(x) \geq (1-1/e)V_i$ for each i,*
- *or returns a certificate that there is no point $x \in P$ such that $F_i(x) \geq V_i$ for all i.*

The algorithm is a generalization of the continuous greedy algorithm from [33]. In a nutshell, the algorithm starts from $x(0) = 0$ and at time t follows a direction $\frac{dx}{dt} = v(t) \in P$ such that $v(t) \cdot \nabla F_i(x(t)) \geq V_i - F_i(x(t))$ for all i. Such a direction can be found by linear programming; if it does not exist, we know that the instance is infeasible. Otherwise, the algorithm reaches a solution $x(1)$ such that $F_i(x(1)) \geq (1 - 1/e)V_i$ for all i. We remark that $x(1) \in P$ even though the path of the algorithm might not be contained in P. In [33], [7], P was assumed to be down-closed, but in fact this is not essential for the continuous greedy algorithm.

III. Applications

Our rounding schemes and concentration results allow us to extend the applications of independent rounding to new settings where we can essentially add a matroid constraint "for free". We can even add a matroid intersection constraint, with some restrictions on the concentration bounds that we want to apply (see Theorem II.3). We can handle linear and also submodular objective functions. For example, Kulik, Shachnai and Tamir [22] gave a $(1 - 1/e - \varepsilon)$-approximation for maximizing a monotone submodular function subject to a

constant number of linear constraints. We can obtain the same result with an additional matroid constraint.

Matroids, matroid intersection and submodular functions are abstract tools, and it is not always easy to see that a given application can be viewed as a special case of some of the problems we can handle. However, the advantage of the abstract view is that one can easily see relevant extensions. For instance, throughput maximization in broadcast scheduling, one of the applications in [19], can be cast as a special case of submodular function maximization subject to a simple partition matroid constraint [9]; this allows us to easily obtain a $(1-1/e)$-approximation for a variety of generalizations that are not obtainable by prior techniques.

In this section we give a sample of results without proofs. We defer the proofs and a more detailed discussion to a full version of the paper.

Matroid bases with packing constraints. The multipath routing problem in the work of Srinivasan [32], and the thin spanning trees for ATSP in the work of Asadpour et al. [2] can be viewed in the common framework of rounding a point in the matroid polytope while minimizing congestion for additional packing constraints. These problems were handled by very different dependent rounding schemes while our rounding procedure is the same for both problems. We prove the following theorem.

Theorem III.1. *Given a matroid $\mathcal{M} = (N, \mathcal{I})$ and a matrix $A \in \mathbb{R}^{m \times N}$, there is an $O(\log m / \log\log m)$-approximation for the problem*

$$\min\{\lambda : \exists \text{ base } B \text{ in } \mathcal{M}; A \cdot \mathbf{1}_B \leq \lambda \mathbf{1}\}.$$

Next, we focus on the following concrete problem.

Max-Min Submodular Allocation. In allocation problems, one can consider various objectives to fulfill. The most common one is the social welfare, i.e. the sum of utilities of all agents. A more difficult objective to optimize is the utility of the least satisfied agent (the "Max-Min allocation" problem):

Given m items and n agents with submodular valuation functions $w_i : \{0,1\}^m \to \mathbb{R}_+$, find an allocation of disjoint sets $S_1, \ldots, S_n$ maximizing $\min_{1 \leq i \leq n} w_i(S_i)$.

First consider the special case of linear valuations, i.e. each item j has value v_{ij} for agent i. Maximizing social welfare in this setting is trivial (just give each item to the agent maximizing its value), while the Max-Min problem is much more difficult. For a constant number of agents, this problem admits an FPTAS [35]. When the number of agents is part of the input, the best known algorithm gives for any fixed $\varepsilon > 0$ an $O(n^{\varepsilon})$-approximation (and takes $n^{O(1/\varepsilon)}$ time) [8]. We consider this problem in the more general setting of submodular valuation functions $w_i : \{0,1\}^m \to \mathbb{R}_+$. For this problem, a $(2n-1)$-approximation for n agents (even with subadditive valuations) was presented by Khot and Ponnuswami [21], and an $\tilde{O}(n^{1/4}m^{1/2})$-approximation was shown by Goemans et al. [20]. In this work, we consider Max-Min Submodular Allocation in the regime of a small number of agents.

Theorem III.2. *For any $\varepsilon > 0$ and any constant number of agents $n \geq 2$, there is a $(1-1/e-\varepsilon)$-approximation for Max-Min Submodular Allocation (in the value oracle model). This is the best possible factor independent of n, since a $(1-(1-1/n)^n+\varepsilon)$-approximation for any fixed $n \geq 2$ and $\varepsilon > 0$ would require exponentially many queries.*

An important ingredient in our algorithm is the generalization of the continuous greedy algorithm to multiple submodular objectives (Theorem II.5), and the lower-tail bound for monotone submodular functions (Theorem II.2). Given a valid guess V for the optimum value, the continuous greedy algorithm finds a fractional solution which satisfies *every* agent to within $(1-1/e)V$. We then round the fractional solution using swap rounding for (partition) matroids, and the lower-tail bound in the rounding stage allows us to conclude that with high probability, no agent loses too much compared to the fractional solution. We remark that we can also achieve a constant factor approximation even when the number of agents is part of the input, provided that all items are "small" compared to the optimum; by small we mean that the value of an item is at most $\mathrm{OPT}/(c \log n)$ for some constant c.

The statement of optimality follows from the oracle hardness construction for Maximum Submodular Welfare in [26], where in the YES case all agents receive the same value V, while in the NO case all the agents together cannot receive more than $(1-(1-1/n)^n+\varepsilon)nV$.

We finally describe a general class of problems that can be handled by our tools.

General framework for constrained allocation problems. *Suppose we have m items and n agents with submodular utility functions $w_i : 2^{[m]} \to \mathbb{R}_+$. Assume in addition that we have the following constraints:*

- *The set of allocated items is required to be independent in a matroid $\mathcal{M}$.*
- *There is a laminar family $\mathcal{F}$ of constraints on the agents, such that for each $S \in \mathcal{F}$ the agents in S can together receive at most k_S items.*

Given these constraints, we want to find an allocation $(S_1, \ldots, S_n)$ maximizing the social welfare $\sum_{i=1}^n w_i(S_i)$.

Using our techniques in the framework of *matroid intersection*, we can prove the following.

Theorem III.3. *There is a $(1-1/e)$-approximation for any allocation problem in the above framework.*

First we observe that the simplest case is when there are no constraints on the items to be allocated and no constraints on the number of items that agents can receive. This special case is the submodular welfare problem and an optimal $(1-1/e)$ approximation for this problem was given in [33]. Even in this setting, there is a natural implicit matroid constraint present - in essence, one makes copies of an item, one for each agent, and this allows one to view the separate submodular functions of the agents as a single submodular function. A matroid constraint then requires that at most 1 copy of each item is allocated.

In our framework, we are allowed to impose any matroid constraint on the set of items that are allocated. For example, items can be partitioned into certain groups and we can only allocate a certain number of items from each group. A more exotic matroid constraint would be that the allocated items should be vertices matchable in a given bipartite graph, or that they should be edges forming a spanning tree in a given graph. In addition, we have constraints on the number of items that the agents can receive, for example agent i can receive at most k_i items and overall we allocate at most m' items.

IV. Randomized swap rounding for matroids

Let $\mathcal{M} = (N, \mathcal{I})$ be a matroid of rank $d = r(N)$ and let $n = |N|$. Randomized swap rounding is a randomized procedure that rounds a point $x \in P(\mathcal{M})$ to an independent set. We present the procedure for points in the base polytope. It can easily be generalized to round any point in the matroid polytope.

Assume that $x \in B(\mathcal{M})$ is the point we want to round. The procedure needs a representation of x as a convex combination of bases, i.e., $x = \sum_{\ell=1}^{m} \beta_\ell \mathbf{1}_{B_\ell}$ with $\sum_{\ell=1}^{m} \beta_\ell = 1, \beta_\ell \geq 0$. We can find such a convex combination using standard techniques (see [30]). The rounding proceeds in $n-1$ stages, where in the first stage we merge the bases B_1, B_2 (randomly) into a new base C_2, and replace $\beta_1 \mathbf{1}_{B_1} + \beta_2 \mathbf{1}_{B_2}$ in the linear combination by $(\beta_1 + \beta_2)\mathbf{1}_{C_2}$. In the k-th stage, C_k and B_{k+1} are merged into a new base C_{k+1}, and $(\sum_{\ell=1}^{k} \beta_\ell)\mathbf{1}_{C_k} + \beta_{k+1}\mathbf{1}_{B_{k+1}}$ is replaced in the linear combination by $(\sum_{\ell=1}^{k+1} \beta_\ell)\mathbf{1}_{C_{k+1}}$. After $n-1$ stages, we obtain a linear combination $(\sum_{\ell=1}^{n} \beta_\ell)\mathbf{1}_{C_n} = \mathbf{1}_{C_n}$, and the base C_n is returned.

Algorithm **MergeBases**$(\beta_1, B_1, \beta_2, B_2)$:
  While $(B_1 \neq B_2)$ do
    Pick $i \in B_1 \setminus B_2$ and find $j \in B_2 \setminus B_1$ such that
      $B_1 - i + j \in \mathcal{I}$ and $B_2 - j + i \in \mathcal{I}$;
    With probability $\beta_1/(\beta_1 + \beta_2)$, $\{B_2 \leftarrow B_2 - j + i\}$;
      Else $\{B_1 \leftarrow B_1 - i + j\}$;
  EndWhile
  Output B_1.

The procedure we use to merge two bases, called **MergeBases**, takes as input two bases B_1 and B_2 and two positive scalars β_1 and β_2. It is described in the figure below. Notice that the procedure relies heavily on the basis exchange property given by Lemma I.2 to guarantee the existence of the elements j in the while loop. As discussed in Section I-A, j can be found by checking all elements in $B_2 \setminus B_1$. Furthermore, since the cardinality of $B_1 \Delta B_2$ decreases at each iteration, the total number of iterations is at most $|B_1| = d$.

Algorithm **SwapRound**$(x = \sum_{\ell=1}^{n} \beta_\ell \mathbf{1}_{B_\ell})$:
  $C_1 = B_1$;
  For $(k = 1$ to $n-1)$ do
    $C_{k+1} =$**MergeBases**$(\sum_{\ell=1}^{k} \beta_\ell, C_k, \beta_{k+1}, B_{k+1})$;
  EndFor
  Output C_n.

The main algorithm **SwapRound** is described in the figure. It uses **MergeBases** to repeatedly merge bases in the convex decomposition of x. After $n-1$ merge operations, we obtain one random base which is our rounded solution. We claim that this solution satisfies the property of negative correlations and increasing expectations for submodular functions. We present further discussion and proofs in Section VI.

V. Randomized swap rounding for matroid intersection

Here we work with two matroids, $\mathcal{M}_1 = (N, \mathcal{I}_1)$ and $\mathcal{M}_2 = (N, \mathcal{I}_2)$. The high-level structure of our rounding procedure for matroid intersection is the same as in Section IV (see $Algorithm$ $\mathbf{SwapRound}(x)$). Again, we need to start with a convex combination $\sum \alpha_I \mathbf{1}_I$ where $I \in \mathcal{I}_1 \cap \mathcal{I}_2$. Any fractional solution $x \in P(\mathcal{M}_1) \cap P(\mathcal{M}_2)$ can be efficiently decomposed in this manner, see [30].

We focus on the implementation of the merge procedure, which is more involved here. We define the *exchange digraph* which is a standard tool in the study of matroid intersection (see [30, Chapter 41]). For $I \in \mathcal{I}_1 \cap \mathcal{I}_2$, we define two digraphs $D_{\mathcal{M}_1}(I)$ and $D_{\mathcal{M}_2}(I)$ as follows.

- For each $i \in I, j \in N \setminus I$ with $I + j - i \in \mathcal{I}_1$, we have an arc $(i, j) \in D_{\mathcal{M}_1}(I)$;
- For each $i \in I, j \in N \setminus I$ with $I + j - i \in \mathcal{I}_2$, we have an arc $(j, i) \in D_{\mathcal{M}_2}(I)$.

We define a digraph $D_{\mathcal{M}_1, \mathcal{M}_2}(I)$ as the union of $D_{\mathcal{M}_1}(I)$ and $D_{\mathcal{M}_2}(I)$. A directed cycle in $D_{\mathcal{M}_1, \mathcal{M}_2}(I)$ corresponds to a chain of feasible swaps. However, it is not necessarily the case that the entire cycle gives a valid exchange in both matroids. Nonetheless, it is known that if a cycle decomposes into two matchings which are unique on their set of vertices respectively in $D_{\mathcal{M}_1}(I)$ and $D_{\mathcal{M}_2}(I)$, then the cycle corresponds to a feasible swap. This motivates the following definition.

Definition V.1. *A directed cycle C in $D_{\mathcal{M}_1, \mathcal{M}_2}(I)$ is* irreducible *if $C \cap D_{\mathcal{M}_1}(I)$ is the unique perfect matching in $D_{\mathcal{M}_1}(I)$ and $C \cap D_{\mathcal{M}_2}(I)$ is the unique perfect matching in $D_{\mathcal{M}_2}(I)$ on the vertex set $V(C)$. Otherwise, C is* reducible*.*

Let us assume that we have two sets $I, J \in \mathcal{I}_1 \cap \mathcal{I}_2$ and $|I| = |J|$. (This assumption can be easily avoided.) The following lemma, building on the ideas of matroid intersection [30, Section 41.3], was proved in [25].

Lemma V.2. *Let $\mathcal{M}_\ell = (N, \mathcal{I}_1)$, $\ell = 1, 2$, be matroids on ground set N. Suppose that $I, J \in \mathcal{I}_1 \cap \mathcal{I}_2$ and $|I| = |J|$. Then there is an integer $s \geq 0$ and a collection of irreducible directed cycles $\{C_1, \ldots, C_m\}$ (allowing repetition) in $D_{\mathcal{M}_1, \mathcal{M}_2}(I)$, using only elements of $I \Delta J$, so that each element of $I \Delta J$ appears in exactly 2^s of the directed cycles.*

We remark that this lemma is existential and there are examples where the collection of irreducible cycles is exponentially large. This is not an issue in [25], but here we need an efficient variant of this decomposition. We prove the following constructive variant of this lemma (we defer the proof to a full version).

Lemma V.3. *Let $\mathcal{M}_\ell = (N, \mathcal{I}_1)$, $\ell = 1, 2$, be matroids on ground set N. Suppose that $I, J \in \mathcal{I}_1 \cap \mathcal{I}_2$ and $|I| = |J|$. Then we can find in polynomial time a collection of irreducible cycles $\{C_1, \ldots, C_m\}, m \leq |I \Delta J|$, in $D_{\mathcal{M}_1,\mathcal{M}_2}(I)$, with coefficients $\gamma_i \geq 0$, $\sum_{i=1}^m \gamma_i = 1$, such that for some $\gamma > 0$, $\sum_{i=1}^m \gamma_i \mathbf{1}_{V(C_i)} = \gamma \mathbf{1}_{I \Delta J}$.*

Swap procedure overview. Let us explain how we use this to implement randomized swap rounding. Suppose for simplicity we have a linear combination $x = \frac{1}{2}\mathbf{1}_I + \frac{1}{2}\mathbf{1}_J$, where $I, J \in \mathcal{I}_1 \cap \mathcal{I}_2$ and $|I| = |J|$. We would like to perform the merge operation on I and J. To that end, we apply Lemma V.3 twice, to obtain (1) a convex combination of irreducible cycles $\sum \gamma_i \mathbf{1}_{V(C_i)}$ in $D_{\mathcal{M}_1,\mathcal{M}_2}(I)$, representing feasible swaps from I to J, and (2) a convex combination of irreducible cycles $\sum \delta_i \mathbf{1}_{V(C'_i)}$ in $D_{\mathcal{M}_1,\mathcal{M}_2}(J)$, representing feasible swaps from J to I.

We have $\sum \gamma_i \mathbf{1}_{V(C_i)} = \gamma \mathbf{1}_{I \Delta J}$ and $\sum \delta_i \mathbf{1}_{V(C'_i)} = \delta \mathbf{1}_{I \Delta J}$, where $\gamma, \delta > 0$ are not necessarily equal. We would like to choose a random swap cycle such that each element in $I \Delta J$ has an equal probability of being added or removed. Therefore, we first decide randomly whether we want to pick a cycle from $D_{\mathcal{M}_1,\mathcal{M}_2}(I)$ (with probability $\frac{\delta}{\gamma+\delta}$), or from $D_{\mathcal{M}_1,\mathcal{M}_2}(J)$ (with probability $\frac{\gamma}{\gamma+\delta}$). Then, we pick a specific cycle: in the first case, we choose C_i with probability γ_i, and replace I by $I \Delta V(C_i)$. In the second case, we choose C'_i with probability δ_i, and replace J by $J \Delta V(C'_i)$. Summing up over all cycles containing a given element in I, we get that the probability of being removed from I is equal to the probability of being added to J, both being equal to $\frac{\gamma\delta}{\gamma+\delta}$. This implies that our fractional solution is preserved in expectation, and it forms a vector-valued martingale.

Once we perform this swap, the size of $I \Delta J$ shrinks. We repeat the same procedure, until I and J become identical. This finishes the merge operation for two vertex solutions $\mathbf{1}_I, \mathbf{1}_J$. It is quite straightforward to generalize from $\frac{1}{2}\mathbf{1}_I + \frac{1}{2}\mathbf{1}_J$ to an arbitrary linear combination. We give a concise description of the merge operation in the box.

Algorithm **MergeIntersectionBases**(α, I, β, J):
While $(I \neq J)$ do
Generate a collection of irreducible cycles in $D_{\mathcal{M}_1,\mathcal{M}_2}(I)$ such that $\sum \gamma_i \mathbf{1}_{V(C_i)} = \gamma \mathbf{1}_{I \Delta J}$, $\sum \gamma_i = 1$, $\gamma_i \geq 0$,
and a collection of irreducible cycles in $D_{\mathcal{M}_1,\mathcal{M}_2}(J)$, such that $\sum \delta_i \mathbf{1}_{V(C'_i)} = \delta \mathbf{1}_{I \Delta J}$, $\sum \delta_i = 1$, $\delta_i \geq 0$.
With probability $\frac{\beta\delta\gamma_i}{\alpha\gamma+\beta\delta}$ for each i, let $I := I \Delta V(C_i)$.
else with probability $\frac{\alpha\gamma\delta_i}{\alpha\gamma+\beta\delta}$ for each i, let $J := J \Delta V(C'_i)$.
EndWhile
Output I.

VI. Some proofs

A. Submodular functions and negative correlation

In this section, we show a connection between the property of negative correlation of random variables $X_1, \ldots, X_n$ and a property of submodular functions of these random variables, $f(X_1, \ldots, X_n)$. Then we prove that both properties hold for the ensembles of random variables arising from our rounding procedures for matroids (the proof for matroid intersection is deferred). The connection we present is interesting in its own right; we remark that negative correlation for matroids under swap rounding (and also randomized pipage rounding) can be shown via a more elementary and direct proof — see our earlier manuscript [10].

Lemma VI.1. *Let $X_1, \ldots, X_n$ be $\{0,1\}$ random variables with expectations $\mathbf{E}[X_i] = x_i$. The following two properties are equivalent:*

- *$X_1, \ldots, X_n$ are negatively correlated, i.e. for any subset $T \subseteq [n]$, we have $\mathbf{E}[\prod_{i \in T} X_i] \leq \prod_{i \in T} x_i$ and $\mathbf{E}[\prod_{i \in T}(1 - X_i)] \leq \prod_{i \in T}(1 - x_i)$.*
- *$X_1, \ldots, X_n$ have the "property of increasing expectations" for any submodular function $f : \{0,1\}^n \to \mathbb{R}$ of coverage type or complementary coverage type[3]: We have $\mathbf{E}[f(X_1, \ldots, X_n)] \geq F(x_1, \ldots, x_n)$ where $F : [0,1]^n \to \mathbb{R}$ is the multilinear extension of f.*

Proof: First, let us assume that $X_1, \ldots, X_n$ are negatively correlated. Let $f(X_1, \ldots, X_n) = |\bigcup_{i:X_i=1} A_i|$ be a coverage-type submodular function. Then $f(X_1, \ldots, X_n)$ for $X_i \in \{0,1\}$ can be written as

$$f(X_1, \ldots, X_n) = \sum_j \left(1 - \prod_{i:j \in A_i} (1 - X_i)\right).$$

By the property of negative correlations,

$$\begin{aligned} \mathbf{E}[f(X_1, \ldots, X_n)] &= \sum_j \left(1 - \mathbf{E}\left[\prod_{i:j \in A_i} (1 - X_i)\right]\right) \\ &\geq \sum_j \left(1 - \prod_{i:j \in A_i} (1 - x_i)\right) \end{aligned}$$

which is exactly the multilinear extension $F(x_1, \ldots, x_n)$. Similarly, a complementary coverage-type function can be written as $f(X_1, \ldots, X_n) = \sum_j (1 - \prod_{i:j \in A_i} X_i)$ and the same argument shows that $\mathbf{E}[f(X_1, \ldots, X_n)] \geq \sum_j (1 - \prod_{i:j \in A_i} x_i)$, which is the respective multilinear extension of f.

Conversely, assume that we have the property of increasing expectations for all coverage-type and complementary coverage-type functions. For a given $T \subseteq [n]$, take the coverage function induced by $A_i = \{1\}$ for $i \in T$ and $A_i = \emptyset$ otherwise. Then we have $f(X_1, \ldots, X_n) = 1 - \prod_{i \in T}(1 - X_i)$ and the multilinear extension of this function is $F(x_1, \ldots, x_n) = 1 - \prod_{i \in T}(1 - x_i)$. The property of increasing expectations, $\mathbf{E}[f(X_1, \ldots, X_n)] \geq F(x_1, \ldots, x_n)$, gives exactly the second condition of negative correlation:

$$\mathbf{E}[\prod_{i \in T}(1 - X_i)] \leq \prod_{i \in T}(1 - x_i).$$

[3] A coverage-type function is $f(S) = |\bigcup_{i \in S} A_i|$ for some set system $\{A_i\}_{i=1}^n$. By complementary coverage type, we mean a function $f(S) = g(\bar{S})$ such that g is of coverage type.

The first condition, $\mathbf{E}[\prod_{i\in T} X_i] \le \prod_{i\in T} x_i$, follows from the property of increasing expectations applied to the complementary coverage function $f(X_1,\dots,X_n) = 1 - \prod_{i\in T} X_i$. ∎

We remark that we have found a counterexample to the above equivalence for general monotone submodular functions (which we omit here). I.e., the property of increasing expectations for all submodular functions is strictly stronger than the property of negative correlation. This also implies that the lower-tail bound in Theorem II.2 cannot follow from the property of negative correlations, because even the expectation $\mathbf{E}[f(X_1,\dots,X_n)]$ could be lower than $F(x_1,\dots,x_n)$ for certain negatively correlated variables with $\mathbf{E}[X_i] = x_i$.

In the following, we consider ensembles of random variables arising from a random process satisfying certain conditions (which are, in particular, satisfied by randomized swap rounding). We will show that such random variables satisfy the property of increasing expectations for all submodular functions, hence implying the property of negative correlations as well.

Lemma VI.2. *Let $\tau \in \mathbb{N}$ and let $\mathbf{X}_t = (X_{1,t},\dots,X_{n,t})$ for $t \in \{0,\dots,\tau\}$ be a non-negative vector-valued random process with initial distribution given by $\mathbf{X}_0 = x \in \mathbb{R}^n$ with probability 1, and satisfying the following properties:*

1) *For every $t \in \{0,\dots,\tau-1\}$, $\mathbf{E}[\mathbf{X}_{t+1} \mid \mathbf{X}_t] = \mathbf{X}_t$.*
2) *For every $t \in \{0,\dots,\tau-1\}$, with probability 1, $\mathbf{X}_{t+1} - \mathbf{X}_t$ has at most one positive coordinate and at most one negative coordinate; the remaining coordinates are zero.*

Then for any submodular function $f : \{0,1\}^n \to \mathbb{R}$ and its multilinear extension $F : [0,1]^n \to \mathbb{R}$, for any $t \in \{0,\dots,\tau\}$,

$$\mathbf{E}[F(\mathbf{X}_t)] \ge F(x).$$

By Lemma VI.1, this also implies that the variables $X_{1,t},\dots,X_{n,t}$ are negatively correlated for each t. We remark that the conditions do not require that when two coordinates change, their sum remains constant. We only assume that two coordinates do not increase or decrease simultaneously.

Proof: We claim that $\mathbf{E}[F(\mathbf{X}_{t+1}) \mid \mathbf{X}_t] \ge F(\mathbf{X}_t)$ for each t. Since $F(\mathbf{X}_0) = F(x)$, this implies the statement of the lemma.

Consider the difference $F(\mathbf{X}_{t+1}) - F(\mathbf{X}_t)$. We know that with probability 1, $\mathbf{X}_t$ and $\mathbf{X}_{t+1}$ differ in at most two coordinates. Assume for now that these coordinates are $i < j$ (if only one coordinate changes, pick j arbitrarily). We denote by $\mathbf{e}_i = \mathbf{1}_{\{i\}}$ and $\mathbf{e}_j = \mathbf{1}_{\{j\}}$ the canonical basis vectors corresponding to elements i and j. Using the fact that F is linear when only one coordinate varies, we can write the difference $F(\mathbf{X}_{t+1}) - F(\mathbf{X}_t)$ as follows:

$$\begin{aligned}
&F(\mathbf{X}_{t+1}) - F(\mathbf{X}_t)\\
&= F(\mathbf{X}_t + (X_{i,t+1} - X_{i,t})\mathbf{e}_i + (X_{j,t+1} - X_{j,t})\mathbf{e}_j) - F(\mathbf{X}_t)\\
&= (X_{i,t+1} - X_{i,t})\frac{\partial F}{\partial x_i}\Big|_{\mathbf{X}_t} + (X_{j,t+1} - X_{j,t})\frac{\partial F}{\partial x_j}\Big|_{\mathbf{X}_t^+}
\end{aligned}$$

where $\mathbf{X}_t^+ = \mathbf{X}_t + (X_{i,t+1} - X_{i,t})\mathbf{e}_i$. Consider the partial derivative at $\mathbf{X}_t^+$. If $X_{j,t+1} > X_{j,t}$, then $X_{i,t+1} \le X_{i,t}$ by the property that at most one coordinate can increase. Then we are evaluating the partial derivative $\frac{\partial F}{\partial x_j}$ at a point $\mathbf{X}_t^+ \le \mathbf{X}_t$. Here we use the property that $\frac{\partial^2 F}{\partial x_i \partial x_j} \le 0$ for the multilinear extension of any submodular function [33], [7]. This implies that if $\frac{\partial F}{\partial x_j}$ is evaluated at a point dominated by $\mathbf{X}_t$, its value is at least $\frac{\partial F}{\partial x_j}\Big|_{\mathbf{X}_t}$. On the other hand, if $X_{j,t+1} < X_{j,t}$, then $X_{i,t+1} \ge X_{i,t}$ and we are evaluating $\frac{\partial F}{\partial x_j}$ at a point $\mathbf{X}_t^+ \ge \mathbf{X}_t$, which means that the partial derivative is at most $\frac{\partial F}{\partial x_j}\Big|_{\mathbf{X}_t}$. In both cases, we get

$$\begin{aligned}
&F(\mathbf{X}_{t+1}) - F(\mathbf{X}_t)\\
&\quad\ge (X_{i,t+1} - X_{i,t})\frac{\partial F}{\partial x_i}\Big|_{\mathbf{X}_t} + (X_{j,t+1} - X_{j,t})\frac{\partial F}{\partial x_j}\Big|_{\mathbf{X}_t}.
\end{aligned}$$

Recall that i, j are in fact random. However, whatever the values of i, j, it holds that

$$F(\mathbf{X}_{t+1}) - F(\mathbf{X}_t) \ge \sum_{i=1}^{n} (X_{i,t+1} - X_{i,t})\frac{\partial F}{\partial x_i}\Big|_{\mathbf{X}_t}$$

since the terms for coordinates that do not change are zero. Now we can take expectation conditioned on $\mathbf{X}_t$, and use the fact that $\mathbf{E}[X_{i,t+1} - X_{i,t} \mid \mathbf{X}_t] = 0$:

$$\begin{aligned}
&\mathbf{E}[F(\mathbf{X}_{t+1}) - F(\mathbf{X}_t) \mid \mathbf{X}_t]\\
&\qquad\ge \sum_{i=1}^{n} \mathbf{E}[X_{i,t+1} - X_{i,t} \mid \mathbf{X}_t]\frac{\partial F}{\partial x_i}\Big|_{\mathbf{X}_t} = 0.
\end{aligned}$$

∎

B. Conditions for negative correlation in matroids

In this section we prove that randomized swap rounding satisfies the conditions of Lemma VI.2. We denote by an *elementary operation* of the swap rounding algorithm one iteration of the while loop in the **MergeBases** procedure, which is repeatedly called in **SwapRound**. Hence, an elementary operation changes two components in one of the bases used in the convex representation of the current point. For example, if the first elementary operation transforms the base B_1 into B_1', then this can be interpreted on the matroid base polytope as transforming the point $x = \sum_{\ell=1}^{n} \beta_\ell \mathbf{1}_{B_\ell}$ into $\beta_1 \mathbf{1}_{B_1'} + \sum_{\ell=2}^{n} \beta_\ell \mathbf{1}_{B_\ell}$. Hence, the **SwapRound** algorithm can be seen as a sequence of dn elementary operations leading to a random sequence $\mathbf{X}_0,\dots,\mathbf{X}_\tau$ where $\mathbf{X}_t$ denotes the convex combination after t elementary operations.

Lemma VI.3. *Randomized swap rounding for matroids satisfies the conditions of Lemma VI.2.*

Proof: Let $\mathbf{X}_t$ denote the fractional solution after t steps, let $X_{i,t}$ denote the i-th component of $\mathbf{X}_t$. To prove the first condition of Lemma VI.2 we condition on a particular vector $\mathbf{X}_t$ at time t of the process and on its convex representation $\mathbf{X}_t = \sum_{\ell=1}^{k} \beta_\ell \mathbf{1}_{B_\ell}$. The vector $\mathbf{X}_{t+1}$ is obtained from $\mathbf{X}_t$ by an elementary operation. Without loss of generality we assume that the elementary operation performs a swap between the bases B_1 and B_2 involving the elements $i \in B_1 \setminus B_2$ and

$j \in B_2 \setminus B_1$. Let B_1' and B_2' be the bases after the swap. Hence, with probability $\beta_1/(\beta_1+\beta_2)$, $B_1' = B_1$ and $B_2' = B_2 - j + i$, and with probability $\beta_2/(\beta_1+\beta_2)$, $B_1' = B_1 - i + j$ and $B_2' = B_2$. Thus,

$$\begin{aligned}
&\mathbf{E}[\beta_1 \mathbf{1}_{B_1'} + \beta_2 \mathbf{1}_{B_2'}] \\
&\quad = \frac{\beta_1}{\beta_1+\beta_2}(\beta_1 \mathbf{1}_{B_1} + \beta_2(\mathbf{1}_{B_2} - \mathbf{e}_j + \mathbf{e}_i)) \\
&\qquad + \frac{\beta_2}{\beta_1+\beta_2}(\beta_1(\mathbf{1}_{B_1} - \mathbf{e}_i + \mathbf{e}_j) + \beta_2 \mathbf{1}_{B_2}) \\
&\quad = \beta_1 \mathbf{1}_{B_1} + \beta_2 \mathbf{1}_{B_2},
\end{aligned}$$

where $\mathbf{e}_i = \mathbf{1}_{\{i\}}$ and $\mathbf{e}_j = \mathbf{1}_{\{j\}}$ denote the canonical basis vectors corresponding to element i and j, respectively. Since the vector $\mathbf{X}_{t+1}$ is given by $\mathbf{X}_{t+1} = \beta_1 \mathbf{1}_{B_1'} + \beta_2 \mathbf{1}_{B_2'} + \sum_{\ell=3}^{k} \beta_\ell \mathbf{1}_{B_\ell}$, we obtain $\mathbf{E}[\mathbf{X}_{t+1} \mid \mathbf{X}_t] = \mathbf{X}_t$. The second condition of Lemma VI.2 is clearly satisfied since only two coordinates are involved in each elementary operation, and exactly one of them increases and one decreases. ∎

ACKNOWLEDGMENTS

CC and JV thank Anupam Gupta for asking about the approximability of maximizing a monotone submodular set function subject to a matroid constraint and a constant number of knapsack constraints; this started their work on this project. JV thanks Ilias Diakonikolas for fruitful discussions concerning multiobjective optimization that inspired one of the applications. RZ is grateful to Michel Goemans for introducing him to a version of Shannon's switching game that inspired the randomized swap rounding algorithm. We thank Nick Harvey for pointing out [31], Mohammad Taghi Hajiaghayi for pointing out relevant work on concentration bounds for independent rounding, Mohit Singh for pointing out work on the crossing spanning tree problem, and Aravind Srinivasan for comments and clarifications on dependent rounding schemes and his work on that topic.

REFERENCES

[1] S. Arora, A. Frieze and H. Kaplan. A new rounding procedure for the assignment problem with applications to dense graph assignment problems. *Math. Programming Ser A*, 92: 1–36, 2002.

[2] A. Asadpour, M. Goemans, A. Madry, S. Oveis Gharan, and A. Saberi. An $O(\log n/\log\log n)$-approximation algorithm for the assymetric traveling salesman problem. *Proc. of ACM-SIAM SODA*, 2010.

[3] A. Asadpour and A. Saberi. An approximation algorithm for max-min fair allocation of indivisible goods. *Proc. of ACM STOC*, 2007.

[4] A. Ageev and M. Sviridenko. Pipage rounding: a new method of constructing algorithms with proven performance guarantee. *J. of Combinatorial Optimization*, 8:307–328, 2004.

[5] J. Byrka, A. Srinivasan, and C. Swamy. Fault-tolerant facility location: a randomized dependent LP-rounding algorithm. *Proc. of IPCO*, 2010.

[6] G. Calinescu, C. Chekuri, M. Pál and J. Vondrák. Maximizing a submodular set function subject to a matroid constraint. *Proc. of IPCO*, 182–196, 2007.

[7] G. Calinescu, C. Chekuri, M. Pál and J. Vondrák. Maximizing a submodular set function subject to a matroid constraint. To appear in *SIAM Journal on Computing, special issue for STOC 2008*.

[8] D. Chakrabarty, J. Chuzhoy and S. Khanna. On allocating goods to maximize fairness. *Proc. of IEEE FOCS*, 107–116, 2009.

[9] C. Chekuri, S. Im and B. Moseley. Manuscript, 2009.

[10] C. Chekuri, J. Vondrák, R. Zenklusen. Dependent randomized rounding for matroid polytopes and applications. Preprint, arXiv:0909.4348, November 2009.

[11] C. Chekuri, J. Vondrák, R. Zenklusen. Multi-budgeted matchings and matroid intersection via dependent rounding. Manuscript, July 2010.

[12] W. H. Cunningham. Testing membership in matroid polyhedra. *Journal of Combinatorial Theory, Series B*, 36(2):161–188, April 1984.

[13] J. Edmonds. Matroids, submodular functions and certain polyhedra. *Combinatorial Structures and Their Applications*, Gordon and Breach, New York, 1970, 69–87.

[14] B. Doerr. Generating randomized roundings with cardinality constraints and derandomizations. *Proc. of STACS*, 571–583, 2006.

[15] B. Doerr. Randomly rounding rationals with cardinality constraints and derandomizations. *Proc. of STACS*, 441–452, 2007.

[16] J. Edmonds. Matroid intersection. *Discrete optimization I*, North Holland, Amsterdam, 1979, 39–49.

[17] U. Feige. A threshold of $\ln n$ for approximating set cover. *Journal of the ACM*, 45(4):634–652, 1998.

[18] M. L. Fisher, G. L. Nemhauser and L. A. Wolsey. An analysis of approximations for maximizing submodular set functions - II. *Math. Prog. Study*, 8:73–87, 1978.

[19] R. Gandhi, S. Khuller, S. Parthasarathy and A. Srinivasan. Dependent rounding and its applications to approximation algorithms. *Journal of the ACM* 53:324-360, 2006.

[20] M.X. Goemans, N. Harvey, S. Iwata and V. Mirrokni. Approximating submodular functions everywhere. *Proc. of ACM-SIAM SODA*, 535–544, 2009.

[21] S. Khot and A.K. Ponnuswami. Approximation algorithms for the max-min allocation problem. In *APPROX-RANDOM 2007*, 204–217.

[22] A. Kulik, H. Shachnai and T. Tamir. Maximizing submodular set functions subject to multiple linear constraints. *Proc. of ACM-SIAM SODA*, 545–554, 2009.

[23] V. S. A. Kumar, M. V. Marathe, S. Parthasarathy and A. Srinivasan. A unified approach to scheduling on unrelated parallel machines. *JACM*, Vol. 56, 2009.

[24] J. Lee, V. Mirrokni, V. Nagarajan and M. Sviridenko. Maximizing non-monotone submodular functions under matroid and knapsack constraints. *Proc. of ACM STOC 2009*, 323–332.

[25] J. Lee, M. Sviridenko, and J. Vondrák. Submodular maximization over multiple matroids via generalized exchange properties. *Proc. of APPROX*, Springer LNCS, 244–257, 2009.

[26] V. Mirrokni, M. Schapira and J. Vondrák. Tight information-theoretic lower bounds for welfare maximization in combinatorial auctions. *Proc. of ACM EC, 2008*.

[27] A. Panconesi and A. Srinivasan. Randomized distributed edge coloring via an extension of the Chernoff-Hoeffding bounds. *SIAM Journal on Computing* 26:350-368, 1997.

[28] P. Raghavan and C. D. Thompson. Randomized rounding: a technique for provably good algorithms and algorithmic proofs. *Combinatorica* 7(4):365–374, 1987.

[29] B. Saha and A. Srinivasan. A new approximation technique for resource-allocation problems. *Proc. of ICS*, 2010.

[30] A. Schrijver. *Combinatorial optimization - polyhedra and efficiency.* Springer, 2003.

[31] P.D. Seymour and D.J. Welsh. Combinatorial applications of an inequality from statistical mechanics. *Mathematical proceedings of the Cambridge Philosophical Society* 77, 485–495, 1975.

[32] A. Srinivasan. Distributions on level-sets with applications to approximation algorithms, *Proc. of IEEE FOCS*, 588–597, 2001.

[33] J. Vondrák. Optimal approximation for the submodular welfare problem in the value oracle model. *Proc. of ACM STOC*, 67–74, 2008.

[34] J. Vondrák. Symmetry and approximability of submodular maximization problems. In *Proc. of IEEE FOCS*, 251–270, 2009.

[35] G. J. Woeginger. When does a dynamic programming formulation guarantee the existence of a fully polynomial time approximation scheme (FPTAS)? *INFORMS Journal on Computing*, 12:5775, 2000.