Constrained Maximization of Lattice Submodular Functions

Aytunc Sahin [1] **Joachim Buhmann** [1] **Andreas Krause** [1]

Abstract

Submodular optimization over the integer lattice has many applications in machine learning. Although the constrained maximization of submodular functions with coordinate-wise concavity (also called *DR-submodular* functions) is well studied, the maximization of *general* lattice submodular functions is considerably more challenging. In this work, we first show that we can optimize lattice submodular functions subject to a discrete (integer) polymatroid constraint using a recently proposed extension, called the Generalized Multilinear Extension. Then, we establish a bound on the rounding error for the discrete polymatroid constraint, which depends on the "distance" between the lattice submodular function to a DR-submodular function. Lastly, we demonstrate the effectiveness of our algorithm on a Bayesian experimental design problem with repetition and a concave cost.

1. Introduction

Submodular functions, which are defined on the subsets of a ground set V containing n elements, have many applications in machine learning (Tschiatschek et al., 2014; Boykov et al., 1999) and they can be described as a pseudo-Boolean function defined on vertices of the unit hypercube $\{0,1\}^n$. Recently, submodularity has been extended to *integer* (Soma et al., 2014; Soma and Yoshida, 2018) domains and *continuous* domains (Bach, 2019). Submodular functions over the integer lattice arise naturally, e.g., if elements can be selected repeatedly, and have found applications in optimal budget allocation (Soma et al., 2014), sensor placement with different power levels (Soma and Yoshida, 2015) and experimental design for causal structure discovery (Agrawal et al., 2019).

[1]Department of Computer Science, ETH Zürich, Switzerland. Correspondence to: Aytunc Sahin <aytunc.sahin@inf.ethz.ch>.

ICML 2020 workshop on Negative Dependence and Submodularity for ML, Vienna, Austria, PMLR 119, 2020.

On *sets*, submodularity is fully characterized via a certain *diminishing returns* property. However, over integer and continuous domains, this is more subtle (Soma et al., 2014). In particular, integer/continous submodularity makes no restrictions on the functions' variation along any single coordinate. A natural restriction is to require coordinate-wise concavity (diminishing returns, DR) as well. Submodular functions with this additional diminishing returns property are called *DR-submodular* (Soma and Yoshida, 2015; Bian et al., 2017c). Maximization of such DR-submodular functions subject to different constraints has been extensively studied (Soma and Yoshida, 2018; 2017; Bian et al., 2017a; 2019; Chen et al., 2018). Ene and Nguyen (2016) provide a polynomial time reduction technique from the integer lattice to the set function maximization case, thus lifting existing techniques for maximizing submodular set functions to the DR-submodular case. Note that this reduction from the integer to the subset lattice only works when the function is DR-submodular. The maximization of general, non DR-submodular (i.e., lattice submodular) functions is considerably more challenging, and only a few constraint settings have been studied. For example, Niazadeh et al. (2018); Gottschalk and Peis (2015) consider box constraints, Kuhnle et al. (2018); Soma and Yoshida (2018); Qian et al. (2018) consider cardinality constraints, and Soma et al. (2014) consider a single knapsack constraint.

To the best of our knowledge, maximization of lattice submodular functions subject to richer constraints is an open problem. In this work, we address the problem of maximizing lattice submodular functions to a rich family of constraints called *discrete polymatroid constraint*. Our main contributions are summarized as follows: We provide the first method to maximize lattice submodular function subject to discrete polymatroid constraint, using the Generalized Multilinear Extension (GME). Then, we characterize the worst-case rounding error over the polytope associated with the discrete polymatroid using a novel notion of distance to a DR-submodular function.

2. Problem Setting

We consider submodular functions defined on different subsets of $\mathbb{R}^n$. A function f is *submodular* if $\forall \mathbf{x}, \mathbf{y} \in \mathcal{X}$

we have

$$f(\mathbf{x}) + f(\mathbf{y}) \geq f(\mathbf{x} \vee \mathbf{y}) + f(\mathbf{x} \wedge \mathbf{y}) \quad (1)$$

where $\vee$ and $\wedge$ are component-wise maximum and minimum functions respectively and $\mathcal{X} = \prod_{i=1}^n \mathcal{X}_i$, where each $\mathcal{X}_i$ is a compact subset of $\mathbb{R}$ (Bach, 2019). From now on, a submodular function f will be referred as *set submodular* if $\mathcal{X}_i = \{0, 1\}$, *lattice submodular* if $\mathcal{X}_i = \{0, 1, \ldots, k-1\}$ and *continuous submodular* if $\mathcal{X}_i = [a, b]$. For set submodular functions, we define the *ground set* as V with $|V| = n$. We use $\mathbb{Z}^V$ to denote $\{0, 1, \ldots, k-1\}^V$ and denote the ith unit vector as $\mathbf{e}_i$. A function f is called *monotone* if for any $\mathbf{x} \leq \mathbf{y}$, we have $f(\mathbf{x}) \leq f(\mathbf{y})$. Moreover, without loss of generality, we assume that the function is normalized, i.e., $f(\mathbf{0}) = 0$. In particular, we make the distinction between two classes of submodular functions below (Bian et al., 2017c; Soma and Yoshida, 2015). We call a function $f \colon \mathbb{Z}^V \to \mathbb{R}$

Lattice submodular (Weak DR) iff for any $i \in V$, $l \in \mathbb{Z}_+$, $\mathbf{x} \leq \mathbf{y} \in \mathbb{Z}^V$

$$f(\mathbf{x} + l\mathbf{e}_i) - f(\mathbf{x}) \geq f(\mathbf{y} + l\mathbf{e}_i) - f(\mathbf{y}) \quad (2)$$

DR-submodular (Strong DR) iff for any $i \in V$, $\mathbf{x} \leq \mathbf{y} \in \mathbb{Z}^V$

$$f(\mathbf{x} + \mathbf{e}_i) - f(\mathbf{x}) \geq f(\mathbf{y} + \mathbf{e}_i) - f(\mathbf{y}) \quad (3)$$

Thus, for the lattice and continuous case, submodularity is equivalent to the *Weak DR* property and for the set case, it is equivalent to *Strong DR* property (Soma and Yoshida, 2015; Bian et al., 2017c).

Discrete Polymatroids

We now introduce the combinatorial object called *Discrete Polymatroids* (Herzog and Hibi, 2002), which have been used in cryptography (Farras et al., 2007; Farras and Padró, 2012), but not yet in machine learning. We will later use discrete polymatroids as constraints. Discrete polymatroids can be viewed as generalizations of *matroids*, arguably the most common family of constraints for maximizing submodular set functions. Similar to matroids, we can define discrete polymatroids using their rank function, their independent sets or their bases. They are very closely related to submodular functions and polymatroids (Murota, 2003).

A *discrete polymatroid* defined on the ground set V is a nonempty set $\mathcal{M} \subset \mathbb{Z}_+^V$ which satisfies two properties: If $\mathbf{y} \in \mathcal{M}$ and $\mathbf{x} \in \mathbb{Z}_+^V$ with $\mathbf{x} \leq \mathbf{y}$, then we have $\mathbf{x} \in \mathcal{M}$. If $\mathbf{x}, \mathbf{y} \in \mathcal{M}$ with $|\mathbf{x}| < |\mathbf{y}|$, then there is $i \in V$ with $x_i < y_i$ such that $\mathbf{x} + \mathbf{e}_i \in \mathcal{M}$. We call a *base* of $\mathcal{M}$ a vector $\mathbf{x}$ with $\mathbf{x} < \mathbf{y}$ for no $\mathbf{y} \in \mathcal{M}$. The base elements also have same cardinality, $|\mathbf{x}| = |\mathbf{y}|$ where $|\mathbf{x}| = \sum_{i=1}^n x_i$. We call $\mathcal{B} \subset \mathbb{Z}_+^V$ the set of bases of a discrete polymatroid iff they satisfy the following two properties. First, for every $\mathbf{x} \in \mathcal{B}$, their cardinality $|\mathbf{x}|$ is the same. Second, for $\mathbf{x}, \mathbf{y} \in \mathcal{B}$ with $x_i > y_i$, then there is an element j with $x_j < y_j$ such that $\mathbf{x} - \mathbf{e}_i + \mathbf{e}_j \in \mathcal{B}$.

Examples: Now we give two examples of discrete polymatroids. First, we can view the cardinality constraint as a discrete polymatroid. The set $\mathcal{M} = \{\mathbf{x} \in \mathbb{Z}_+^V : |\mathbf{x}| \leq K\}$ is a discrete polymatroid with base vectors $\mathcal{B} = \{\mathbf{y} \in \mathbb{Z}_+^V : |\mathbf{y}| = K\}$. Second, we can generalize the partition matroid to the integer vectors. When we divide the ground set into J disjoint parts and have cardinality constraint on these parts, this structure is a discrete polymatroid, namely $\mathcal{M} = \{\mathbf{x} \in \mathbb{Z}_+^V : |\mathbf{x}_{S_j}| \leq b_j \text{ for all } j = 1, \ldots, J\}$ where S_j are disjoint subsets of V.

3. Maximizing Lattice Submodular Functions Over a Discrete Polymatroid

We consider the following problem:

$$\begin{aligned} \underset{\mathbf{x}}{\text{maximize}} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{x} \in \mathcal{M} \end{aligned} \quad (4)$$

where $f(\mathbf{x})$ is a lattice submodular function and $\mathcal{M}$ is a discrete polymatroid. We assume that $\mathcal{M}$ is contained in the natural domain of f, i.e., $\mathcal{M} \subseteq \mathbb{Z}_+^V$. Here is our roadmap to solve this problem: First, we relax the problem into the continuous domain. Second, we solve the relaxed problem with an approximation guarantee. Finally, we round the solution back and bound the rounding error which depends on the "distance" of the function f to a DR-submodular function.

3.1. The Generalized Multilinear Extension (GME) and its Properties

A Continuous Extension: We use the Generalized Multilinear Extension (GME, Sahin et al. (2020), which has several attractive properties. Let f be a lattice submodular function of n variables. The GME is defined probabilistically, as an expectation of f over random integer vectors $R = [x_1, \ldots, x_n]$ where each coordinate x_i is selected independently according to some categorical distribution $\boldsymbol{\rho}_i$. The GME is defined on the product of n simplices, i.e., for $\boldsymbol{\rho} \in \Delta_n^{k-1}$, via

$$F(\boldsymbol{\rho}) = F(\boldsymbol{\rho}_1, \ldots, \boldsymbol{\rho}_n) = \mathbb{E}_{R \sim \boldsymbol{\rho}_1, \ldots, \boldsymbol{\rho}_n}[f(R)] \quad (5)$$

Moreover, the GME has the following properties making it attractive for maximization problems:

Proposition 1. *Let F be the GME of an integer submodular function f. Then we have*

1. *If f is monotone then $\frac{\partial F}{\partial \rho_{ij}} \geq 0$ for all $i \in V$ and $j \in \{1, \ldots, k-1\}$*

2. *F is DR-submodular (even if f is not), i.e., $\frac{\partial^2 F}{\partial \rho_{ij} \partial \rho_{kl}} \leq 0$ for all i, j, k, l.*

Now, we discuss the continuous relaxation of the constraint $\mathcal{M}$. Since $\mathcal{M}$ is a discrete polymatroid, its relaxation $\mathcal{P}(\mathcal{M})$ becomes an integral polymatroid. For every valid point $\boldsymbol{\rho}$, we have its corresponding point $\mathbf{x} \in \mathcal{P}(\mathcal{M})$ which is given by the following relation. We can define a linear function $T : \Delta_n^{k-1} \to \mathbb{R}_+^V$ where $\mathbf{x} = T(\boldsymbol{\rho})$ using $x_i = \sum_{j=1}^{k-1} j\rho_{ij}\ \forall i \in V$.

3.2. Maximizing and Rounding the GME

Solving the relaxed problem: Now our original problem (Equation (4)) becomes

$$\max\{F(\boldsymbol{\rho}) : \boldsymbol{\rho} \in \Delta_n^{k-1}, T(\boldsymbol{\rho}) \in \mathcal{P}(\mathcal{M})\}. \quad (6)$$

Since our new objective function F is DR-submodular, and the constraints are down-closed and convex, we can use any continuous DR-submodular optimization algorithm for maximization (Mokhtari et al., 2018; Bian et al., 2017c). These methods first linearize the objective function using its gradient and then solve the linearized problem over the constraint set iteratively. Thus, these methods need a separation oracle for $P(\mathcal{M})$, which can be implemented, e.g., via the ellipsoid method. We can thus find, in polynomial time, a vector $\mathbf{v}$ which maximizes $\langle \mathbf{v}, \nabla F(\boldsymbol{\rho})\rangle$ subject to $\{\mathbf{v} \in \Delta_n^{k-1}, T(\mathbf{v}) \in \mathcal{P}(\mathcal{M})\}$. Separation for the constraint $T(\boldsymbol{\rho}) \in \mathcal{P}(\mathcal{M})$ is proved by Ene and Nguyen (2016). In addition, separation for the constraint $\boldsymbol{\rho} \in \Delta_n^{k-1}$ can be done easily by coordinate-wise check and a summation check for n dimensions. Then, we can obtain a $(1 - \frac{1}{e})$ guarantee for monotone f and $\frac{1}{e}$ guarantee (Feldman et al., 2011; Mokhtari et al., 2018) for the optimization (6) over the *continuous* domain.

Rounding Back: Rounding in the polytope of the discrete polymatroid has been considered independently by Soma and Yoshida (2018); Chekuri and Vondrák (2009) and for the DR-submodular case, the integer point can be recovered by a provably lossless rounding scheme (Calinescu et al., 2011; Chekuri et al., 2010). Similarly to Soma and Yoshida (2018); Ene and Nguyen (2016), we can round the fractional solution $\mathbf{x}_{\mathrm{F}}^* = T(\boldsymbol{\rho}^*)$ to an integral solution $\mathbf{x}_{\mathrm{I}}^*$ without any loss in the approximation when the function f is *DR-submodular*, since the GME is always upper bounded by the value of the multilinear extension of the submodular function restricted on the unit cube which $\mathbf{x}$ belongs to. However, for the *lattice submodular* case we incur an error for the rounding which we characterize in Section 3.4. We use randomized pipage rounding (Chekuri and Vondrák, 2009) technique which does not require any function evaluation.

Algorithm 1 Monotone Lattice Submodular Maximization subject to a Discrete Polymatroid

Input: $f : \mathbb{Z}_+^V \to \mathbb{R}, \mathcal{M} \subset \mathbb{Z}_+^V$

Initialize $\boldsymbol{\rho}^0 \in \Delta_n^{k-1}$, $t \leftarrow 1$

while $t \leq T_{\mathrm{F}}$ **do**

  Compute $\nabla F(\boldsymbol{\rho}^{t-1})$

  $\mathbf{v}_t \leftarrow \mathrm{argmax}_{\mathbf{v} \in \Delta_n^{k-1}, T(\mathbf{v}) \in \mathcal{P}(\mathcal{M})} \langle \nabla F(\boldsymbol{\rho}^{t-1}), \mathbf{v}\rangle$

  $\boldsymbol{\rho}^t \leftarrow \boldsymbol{\rho}^{t-1} + \frac{1}{T}\mathbf{v}^t$

  $t \leftarrow t + 1$

$\mathbf{x}_{\mathrm{F}} \leftarrow T(\boldsymbol{\rho}^{T_{\mathrm{F}}})$

$\mathbf{x}_{\mathrm{I}} \leftarrow \mathrm{Round}(\mathbf{x}_{\mathrm{F}})$

Output: $\mathbf{x}_{\mathrm{I}} \in \mathcal{M}$

3.3. Distance to a DR-submodular Function

There are several definitions that characterize how close a function is to a submodular function. For the set submodular case, the *submodularity ratio* is introduced by Das and Kempe (2011), and generalized to the lattice case by Kuhnle et al. (2018). Bian et al. (2017b) define the *generalized curvature* of a set function However, we are primarily interested in how close a given lattice submodular function is to a DR-submodular function. The *DR-ratio*, introduced by Lehmann et al. (2006); Qian et al. (2018); Kuhnle et al. (2018) characterize how far a lattice submodular function is to a DR-submodular function. The DR-ratio is defined by

$$\beta_f = \min_{\mathbf{x} \leq \mathbf{y}, i \in V} \frac{f(\mathbf{x} + \mathbf{e}_i) - f(\mathbf{x})}{f(\mathbf{y} + \mathbf{e}_i) - f(\mathbf{y})} \quad (7)$$

For monotone DR-submodular functions, β_f is always 1. For monotone lattice submodular functions, we have $0 \leq \beta_f < 1$. In order to bound the rounding error, we will use an *additive* version of the DR-ratio which was also considered in McMeel and Parpas (2019):

$$\lambda = \max_{\mathbf{x} \in \mathbb{Z}_+^V, i \in V} f(\mathbf{x} + 2\mathbf{e}_i) - 2f(\mathbf{x} + \mathbf{e}_i) + f(\mathbf{x}) \quad (8)$$

As stated in McMeel and Parpas (2019), we can have the following decomposition. Let f be a lattice submodular function. Then f can be represented as the sum of a modular function and a DR-submodular function.

Proposition 2. *$f_{\mathrm{DR}}(\mathbf{x}) = f(\mathbf{x}) - \lambda m(\mathbf{x})$ is DR-submodular with $m(\mathbf{x}) = \sum_{i=1}^n x_i^2$*

Note that, DR-ratio is defined for *monotone* submodular functions whereas λ and this decomposition holds for both *monotone* and *non-monotone* functions.

3.4. Bounding the Rounding Loss

So far, we have seen how to maximize a lattice submodular function subject to a discrete polymatroid constraint. Now, using the decomposition above, we will bound the rounding error, which depends on the λ.

Table 1. Bayesian Experimental Design with Repetition and Concave Cost. TFW performs well in the continuous domain and its rounded values are better than greedy.

		$K = 3, p = 2$		$K = 3, p = 4$	
	type	$\alpha = 0.5$ $\gamma = 0.2$	$\alpha = 0.5$ $\gamma = 0.3$	$\alpha = 0.5$ $\gamma = 0.2$	$\alpha = 0.5$ $\gamma = 0.3$
f_{GR}	integer	5.56	4.96	7.01	5.95
F_{TFW}	fractional	5.52	4.92	7.02	5.91
F_{ShFW}	fractional	4.84	4.30	6.35	5.28
f_{TFW}	integer	5.62	5.02	7.11	5.97
f_{ShFW}	integer	5.34	4.57	6.77	5.70

Theorem 1. *Let f be a monotone lattice submodular function with the decomposition $f(\mathbf{x}) = f_{\mathrm{DR}}(\mathbf{x}) + \lambda \|\mathbf{x}\|^2$. Then we have*

$$f(\mathbf{x}_{\mathrm{I}}^*) \geq (1 - \tfrac{1}{e}) \max_{\mathbf{x} \in \mathcal{M}} f(\mathbf{x}) - \lambda \left(\mathbb{E}_{\boldsymbol{\rho}*} \|\mathbf{x}\|^2 - \|\mathbf{x}_{\mathrm{I}}^*\|^2 \right) \quad (9)$$

where $\mathbb{E}_{\boldsymbol{\rho}*} \|\mathbf{x}\|^2 - \|\mathbf{x}_{\mathrm{I}}^*\|^2 = \left(\mathbb{E}_{\boldsymbol{\rho}*} \|\mathbf{x}\|^2 - \|x_{\mathrm{F}}^*\|^2 \right) + \left(\|x_{\mathrm{F}}^*\|^2 - \|\mathbf{x}_{\mathrm{I}}^*\|^2 \right) \leq \frac{nk^2}{4} + n$

For proof, please refer to Appendix A.

Corollary 1.1. *Let f be a general (non-monotone) lattice submodular function with the decomposition $f(\mathbf{x}) = f_{\mathrm{DR}}(\mathbf{x}) + \lambda \|\mathbf{x}\|^2$. Then we have*

$$f(\mathbf{x}_{\mathrm{I}}^*) \geq \tfrac{1}{e} \max_{\mathbf{x} \in \mathcal{M}} f(\mathbf{x}) - \lambda \left(\mathbb{E}_{\boldsymbol{\rho}*} \|\mathbf{x}\|^2 - \|\mathbf{x}_{\mathrm{I}}^*\|^2 \right) \quad (10)$$

Note that Algorithm 1 does *not* require an estimate of λ to be run. λ is only required to get an upper bound for the rounding error.

4. Experiments

Bayesian Experimental Design with Diminishing Cost

We consider a Bayesian Experimental Design (Chaloner and Verdinelli, 1995) task. Instead of selecting a *set* of experiments, we select *multisets* by allowing repetitions. Following the setting in Srinivas et al. (2010), we consider $g(\mathbf{x}) = \frac{1}{2} \log \det \left(\mathbf{I} + \sigma^{-2} \mathbf{K}_{\mathbf{x}} \right)$. Note that since we allow repetitions, $\mathbf{K}_{\mathbf{x}}$ is a positive definite kernel matrix of size $|\mathbf{x}| \times |\mathbf{x}|$. This function g measures the informativeness of the selected multiset and is known as the *information gain*, of primary interest in Bayesian D-optimal design. Notably, g is monotone DR-submodular. In many experimental design settings, repeating the same experiment multiple times is cheaper than performing multiple different experiments, as there are shared fixed costs when setting up an experiment. Motivated by this, we also consider the cost of selecting an experiment multiple times. We model the shared costs using a separable concave function $c_{\gamma,\alpha}(\mathbf{x}) = \gamma \sum_{i=1}^n x_i^\alpha$, where $\alpha \in (0, 1)$. Combined, we have the following utility function: $f(\mathbf{x}) = g(\mathbf{x}) - c_{\gamma,\alpha}(\mathbf{x})$. Note that since this cost c is separable, f still remains submodular but no longer DR-submodular. Furthermore, as γ increases, f becomes non-monotone.

We test our function on the temperature data collected from 46 sensors deployed at Intel Research Berkeley. Following the experimental setup in (Srinivas et al., 2010), we use $\sigma^2 = 0.5$. We use an integer lattice of size $k = 5$, concavity parameter of the cost $\alpha = 0.5$, and the fixed cost parameter $\gamma = 0.2$. To optimize the non-monotone GME, we use the Two-Phase FW (TFW) and Shrunken FW (ShFW) from Bian et al. (2017a). As a baseline, we implemented a basic greedy method (GR) which adds one element with the largest marginal gain at each iteration. If the marginal gain becomes negative, the algorithm terminates. Otherwise, it terminates when the constraint can no longer be satisfied. For non-monotone lattice submodular functions, we are not aware of an approximation guarantee of this basic algorithm. We use the oblivious step size $\frac{2}{t+2}$ for TFW with a total number of 20 steps. For rounding, we use the randomized pipage rounding procedure for non-monotone submodular functions (Vondrák, 2011) 10 times and report the maximum value. As for the constraints, we use discrete polymatroids with varying size of partition (p) and cardinality (K) constraints.

In the Table 1, we observe two main findings. First, TFW performs better than ShFW in the continuous domain, as also observed by (Bian et al., 2017a). Although the approximation guarantee for TFW is $\frac{1}{4}$ and for ShFW is $\frac{1}{e}$, TFW performs better. When we compare the rounded function values with the greedy algorithm, we see that the rounded value of TFW is better. Here, the parameters α and γ both affect the non-DR submodularity of the function and although Corollary 1.1 tells us that in the worst case, the rounded function value may be much lower than the fractional function value, we do not observe this behaviour on this real dataset.

5. Conclusion

In this work, we present the first approach for lattice submodular maximization subject to a discrete polymatroid constraint. Our approach exploits the Generalized Multilinear Extension, which is a DR-submodular extension even if the function is only lattice submodular. After maximizing the extension, we show that we can bound the rounding error. To do this, we decompose the lattice submodular function into two parts, where one part is DR-submodular and the remaining part is modular. The modular part depends on the distance between the lattice submodular function to a DR-submodular function and we characterize the relationship between rounding error and the decomposition.

References

Agrawal, R., Squires, C., Yang, K., Shanmugam, K., and Uhler, C. (2019). Abcd-strategy: Budgeted experimental design for targeted causal structure discovery. *arXiv preprint arXiv:1902.10347*.

Bach, F. (2019). Submodular functions: from discrete to continuous domains. *Mathematical Programming*, 175(1-2):419–459.

Bian, A., Levy, K., Krause, A., and Buhmann, J. M. (2017a). Continuous dr-submodular maximization: Structure and algorithms. In *Advances in Neural Information Processing Systems*, pages 486–496.

Bian, A. A., Buhmann, J. M., Krause, A., and Tschiatschek, S. (2017b). Guarantees for greedy maximization of non-submodular functions with applications. *CoRR*, abs/1703.02100.

Bian, A. A., Mirzasoleiman, B., Buhmann, J., and Krause, A. (2017c). Guaranteed Non-convex Optimization: Submodular Maximization over Continuous Domains. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54 of *Proceedings of Machine Learning Research*, pages 111–120, Fort Lauderdale, FL, USA.

Bian, Y., Buhmann, J., and Krause, A. (2019). Optimal continuous dr-submodular maximization and applications to provable mean field inference. In *International Conference on Machine Learning*, pages 644–653.

Boykov, Y., Veksler, O., and Zabih, R. (1999). Fast approximate energy minimization via graph cuts. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*, volume 1, pages 377–384. IEEE.

Calinescu, G., Chekuri, C., Pál, M., and Vondrák, J. (2011). Maximizing a monotone submodular function subject to a matroid constraint. *SIAM J. Comput.*, 40(6):1740–1766.

Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: A review. *Statist. Sci.*, 10(3):273–304.

Chekuri, C. and Vondrák, J. (2009). Randomized pipage rounding for matroid polytopes and applications. *CoRR*, abs/0909.4348.

Chekuri, C., Vondrak, J., and Zenklusen, R. (2010). Dependent randomized rounding via exchange properties of combinatorial structures. In *Proceedings of the 2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, FOCS '10, page 575–584, USA. IEEE Computer Society.

Chen, L., Hassani, H., and Karbasi, A. (2018). Online continuous submodular maximization. *arXiv preprint arXiv:1802.06052*.

Das, A. and Kempe, D. (2011). Submodular meets spectral: Greedy algorithms for subset selection, sparse approximation and dictionary selection. In *Proceedings of the 28th International Conference on International Conference on Machine Learning*, ICML'11, page 1057–1064, Madison, WI, USA. Omnipress.

Ene, A. and Nguyen, H. L. (2016). A reduction for optimizing lattice submodular functions with diminishing returns. *arXiv preprint arXiv:1606.08362*.

Farras, O., Martí-Farré, J., and Padró, C. (2007). Ideal multipartite secret sharing schemes. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 448–465. Springer.

Farras, O. and Padró, C. (2012). Ideal hierarchical secret sharing schemes. *IEEE transactions on information theory*, 58(5):3273–3286.

Feldman, M., Naor, J., and Schwartz, R. (2011). A unified continuous greedy algorithm for submodular maximization. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pages 570–579. IEEE.

Gottschalk, C. and Peis, B. (2015). Submodular function maximization on the bounded integer lattice. In *International Workshop on Approximation and Online Algorithms*, pages 133–144. Springer.

Herzog, J. and Hibi, T. (2002). Discrete polymatroids. *J. Algebraic Comb.*, 16(3):239–268.

Kuhnle, A., Smith, J. D., Crawford, V. G., and Thai, M. T. (2018). Fast maximization of non-submodular, monotonic functions on the integer lattice. *arXiv preprint arXiv:1805.06990*.

Lehmann, B., Lehmann, D., and Nisan, N. (2006). Combinatorial auctions with decreasing marginal utilities. *Games and Economic Behavior*, 55(2):270 – 296.

McMeel, C. and Parpas, P. (2019). Majorisation-minimisation algorithms for minimising the difference between lattice submodular functions. *CoRR*, abs/1905.13492.

Mokhtari, A., Hassani, H., and Karbasi, A. (2018). Stochastic conditional gradient methods: From convex minimization to submodular maximization.

Murota, K. (2003). *Discrete Convex Analysis: Monographs on Discrete Mathematics and Applications 10*. Society for Industrial and Applied Mathematics, USA.

Niazadeh, R., Roughgarden, T., and Wang, J. R. (2018). Optimal algorithms for continuous non-monotone submodular and dr-submodular maximization. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 9617–9627, Red Hook, NY, USA. Curran Associates Inc.

Qian, C., Zhang, Y., Tang, K., and Yao, X. (2018). On multiset selection with size constraints. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

Sahin, A., Bian, Y., Buhmann, J. M., and Krause, A. (2020). From sets to multisets: Provable variational inference for probabilistic integer submodular models.

Soma, T., Kakimura, N., Inaba, K., and Kawarabayashi, K.-i. (2014). Optimal budget allocation: Theoretical guarantee and efficient algorithm. In *International Conference on Machine Learning*, pages 351–359.

Soma, T. and Yoshida, Y. (2015). A generalization of submodular cover via the diminishing return property on the integer lattice. In *Advances in Neural Information Processing Systems*, pages 847–855.

Soma, T. and Yoshida, Y. (2017). Non-monotone dr-submodular function maximization. In *Thirty-First AAAI conference on artificial intelligence*.

Soma, T. and Yoshida, Y. (2018). Maximizing monotone submodular functions over the integer lattice. *Mathematical Programming*, 172(1-2):539–563.

Srinivas, N., Krause, A., Kakade, S., and Seeger, M. (2010). Gaussian process optimization in the bandit setting: No regret and experimental design. In *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ICML'10, page 1015–1022, Madison, WI, USA. Omnipress.

Tschiatschek, S., Iyer, R. K., Wei, H., and Bilmes, J. A. (2014). Learning mixtures of submodular functions for image collection summarization. In *Advances in neural information processing systems*, pages 1413–1421.

Vondrák, J. (2011). Symmetry and approximability of submodular maximization problems. *CoRR*, abs/1110.4860.