Folha de Rosto para Projeto de Iniciação Científica — UFABC
Edital 04/2022

Título:	Optimizing the Difference of Submodular Functions in Restricted Settings
Supervisor:	Prof. Dr.
Candidata/o:	
Período:	01/09/2022 a 31/08/2023
Instituição Sede:	Centro de Matemática, Computação e Cognição Universidade Federal do ABC
Palavras Chave:	Submodular Functions, Pseudo-Boolean Functions, Discrete Optimization, Discrete Convex Analysis
Área de Conhecimento:	Ciência da Computação, Matemática

Resumo:

Submodular functions are an important class of objective and constraint functions in discrete optimization problems, with many deep and beautiful results and lots of applications. Recently, the problem of optimizing the difference or the ratio of submodular functions has received some attention in the literature for encompassing a variety of machine learning models. We propose a continuation of that systematic investigation by focusing on some structurally-restricted subclasses of submodular functions, their algorithmic aspects, and their tantalizing consequences in machine learning applications.

Declaração de Interesse por Bolsa

Declaro que a/o candidata/o, nos termos do edital 04/2022, deseja participar do programa de iniciação científica como **bolsista**.

Nota: Este projeto de iniciação científica está inserido no contexto de um projeto de pesquisa que estuda aspectos estruturais e ramificações algorítmicas em espaços dotados de convexidades não euclidianas.

Optimizing the Difference of Submodular Functions in Restricted Settings

Abstract

Submodular functions are an important class of objective and constraint functions in discrete optimization problems, with many deep and beautiful results and lots of applications. Recently, the problem of optimizing the difference or the ratio of submodular functions has received some attention in the literature for encompassing a variety of machine learning models. We propose a continuation of that systematic investigation by focusing on some structurally-restricted subclasses of submodular functions, their algorithmic aspects, and their tantalizing consequences in machine learning applications.

1 Introduction and Definitions

Semimodular functions are real-valued mappings on distributive lattices exhibiting properties of marginal decreasing (if submodular) or increasing (if supermodular) returns – definitions shall be given shortly – and have many applications in areas as diverse as Algebra (vector space and lattice rankings [48]), Combinatorics (matroids and polymatroids [18, 57]), Computer Vision (image segmentation [4]), Economics (markets [55]), Engineering (rigidity and electrical flows [52, 42]), Game Theory (cooperative games [46]), Graph Theory (packing and covering [50]), Machine Learning (document summarization [37, 38], graphical models [33]), Operations Research (inventory [9] and scheduling [39, 2] problems) Statistics (discrete entropy [59]), and Theoretical Computer Science (constraint satisfaction problems [5]).

The current project intends to cover the latest advances in the literature regarding the optimization of semimodular functions in both constrained and unconstrained settings, as well as to advance the field through development of novel structural, algorithmic, and complexity results. In particular, we shall focus our efforts towards optimization problems involving *differences* or *ratios* of semimodular functions. These types of problems were first investigated in [14, 20] and have experienced a rekindled interest in recent years [41, 29, 1, 30] for serving as models in a variety of machine learning applications as sensor placement [24, 34], discriminatively structured graphical models [41], feature selection [33], and probabilistic inference [49].

Despite all the history and some striking recent developments, the whole picture is far from being understood and much remains to be investigated. For instance, when the semimodular functions involved satisfy some syntactic or structural hypotheses as symmetry, bounded algebraic-degree, or self-duality, or when the functions fail semimodular properties, but are close enough under a suitable metric.

In the remainder of this section, we introduce some key definitions that are used in the sequel. The reader familiar with rudiments the field can safely skip it. In Section 2, we state the problems in a more precise manner and present a sample of questions we are interested in and sometimes, discuss how we are currently thinking about attacking them. Section 3 establish formative goals and Section 4 describe the methodology that shall be used. In closing, a scheduled of foreseeing activities in given in Section 5.

1.1 Semimodular Functions

A **submodular function** over a finite *ground* set U is a mapping $f : 2^U \to \mathbb{R}$ such that

$$f(S) + f(T) \geq f(S \cap T) + f(S \cup T) \qquad \text{for all} \qquad S, T \subseteq U \tag{1}$$

or equivalently, through a simple combinatorial argument, such that for all $x \in U$,

$$f(S \cup \{x\}) - f(S) \geq f(T \cup \{x\}) - f(T) \qquad \text{for all} \qquad S \subseteq T \subseteq U \setminus \{x\}. \tag{2}$$

A function $g : 2^U \to \mathbb{R}$ is **supermodular** if $-g$ is submodular, and it is **modular** if simultaneously super and submodular. A function is termed **semimodular** if it is either submodular or supermodular. A semimodular function f is **increasing** if $f(S) \leq f(T)$ for all $S \subseteq T \subseteq U$, **decreasing** if $-f$ is increasing, **normalized** if $f(\emptyset) = 0$, and **symmetric** if $f(S) = f(U \setminus S)$ for all $S \subseteq U$. It is **integral** if $\mathrm{img}(f) \subseteq \mathbb{Z}$. (Note: we shall focus solely on semimodular functions over the Boolean algebra 2^U, which is clearly a distributive lattice.)

Equation (2) shows that submodular functions exhibit the (marginal) *diminishing-returns property* sometimes available for indivisible goods in economic settings and hence, are related to *concave functions*[1]. That relation was further strengthened recently in [31]. On the other hand, submodular functions also possess many characteristics exhibited by *convex functions*[2] as efficient, exact unconstrained minimization, good approximation guarantees for maximization, *sub-differentials*, *Fenchel's duality*, and are considered, in a sense, their discrete analogue. We proceed with a brief covering of the literature regarding optimization aspects. It serves both as motivation and to illustrate the reasoning underlying some of our proposed directions and ideas.

The unconstrained minimization of submodular functions was shown to be solvable in the value-oracle[3] model in strongly polynomial time in [21, 22, 23] through the use of the *ellipsoid method* (a polynomial time, albeit highly convoluted and impractical, solver for the broad class of *linear programming* problems). Strongly polynomial-time combinatorial algorithms were later, simultaneously introduced in [56] and [28]. A sequence of improvements [60, 16, 27, 26] culminated in a $O(n^6 + \gamma n^5)$ running time [47], where γ means the running time of an oracle call and n the cardinality of the ground set. Under extra assumptions as bounded algebraic-degree, symmetry, or restricted domain, there are more efficient algorithms: $O(n^3)$ in [25, 6, 8, 3, 61] and in [51, 53, 40], and $O(\gamma n^{5/3}/\varepsilon^2)$, with $0 < \varepsilon \ll 1$, in [11], in that order.

Despite being impressive achievements themselves, the n^6 running-time dominant term is too high for modern applications – especially, those inside the *big data* and computer vision realms – and pales in comparison to the $O(n^3)$ cost of flow-based techniques. Furthermore, since they are all value-oracle-based algorithms, they completely neglect the fact that an algebraic expression for the function is / may be available.

The maximization counterpart was shown to be NP-hard in [45, 43], which also introduced the first constant-factor $(1 - 1/e)$ approximation algorithm when the function is increasing and subjected to a cardi-

[1] A function $f : \mathbb{R}^n \to \mathbb{R}$ is **concave** if $f(\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y}) \geq \lambda f(\boldsymbol{x}) + (1 - \lambda) f(\boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathrm{dom}(f)$ and $\lambda \in [0, 1]$.

[2] $f : \mathbb{R}^n \to \mathbb{R}$ is **convex** if $-f$ is concave. Convex functions are know to have global minima that can be determined in polynomial time in n and in the size of the function's description. For details on convexity and concavity, see [54].

[3] A **value-oracle** is any procedure for a function f that, given $\boldsymbol{x} \in \mathrm{dom}(f)$, returns $f(\boldsymbol{x})$ and no extra information about f.

nality constraint. Those results were further extended in [58, 32, 37, 15, 44], some considering also matroid and knapsack constraints. The case when the function is non-increasing was first investigated in [15] and extended in [10, 35, 36].

Perhaps surprising, while unconstrained submodular minimization is in P and both constrained and unconstrained versions of maximization are in NP, but admit constant factor approximations in polynomial time, the constrained minimization is NP-hard and also hard to approximate. Some progress in a restricted sense was made in [30].

1.2 Pseudo-Boolean Functions

A **pseudo-Boolean function** (pBf) in $n \geq 0$ variables is any real-valued mapping of the form $\{0,1\}^n \mapsto \mathbb{R}$. A pBf is **Boolean** if the binary set $\{0,1\}$ is used in place of $\mathbb{R}$. Pseudo-Boolean functions are closely related to **finite-set functions** – mappings of the form $2^S \mapsto \mathbb{R}$, for a finite **ground** sets S – since any subset $T \subseteq S$ can be identified with its *support* binary vector in $\{0,1\}^S$.

It is well known [7, 13] that any pBf f in variables $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ can be uniquely represented by a multi-affine polynomial P_f in the same variables, namely as

$$P_f(\boldsymbol{x}) := \sum_{a \in \{0,1\}^n} f(a) \cdot \prod_{a_i = 1} x_i \cdot \prod_{a_j = 0} (1 - x_j) = \sum_{S \subseteq [n]} a_S \prod_{i \in S} x_i.$$

The **degree** of f is defined as $d(f) := \max\{|S \subseteq [n]| : a_S \neq 0\}$, the degree of its polynomial representation. Henceforth, we shall identify f with P_f and simple write f when talking about the function or its polynomial indiscriminately. The **algebraic-degree** of a semimodular function is then the degree of its multi-affine polynomial representation.

It is well known that a quadratic pBf f is submodular if and only if all quadratic terms of f are non positive [43]. Hence, deciding if a quadratic pBf is submodular can be done in linear time in the number of its terms. A similar, slightly more involved result [3] is known for cubic, submodular pBfs. For degree-4 and higher, no characterization is believed to exist, since the following theorem is known.

Theorem 1.1 (Crama [12], Nemhauser, Wolsey e Fisher [45])**.** *It is* coNP*-hard to determine if a degree-*4 *or higher pBf f is submodular, even if f is given by its multi-affine polynomial representation.*

Despite this negative result, it is known [17, 62] that some subclasses of higher-degree submodular functions admit quadratic representations at the cost of introducing a small number of *auxiliary* variables. Hence, the mentioned algorithms [25, 6, 8, 3, 61] can be used to optimize them.

2 Problems and Questions

Let f and g be submodular functions over a finite ground set U. We shall dedicate most of our efforts to the problem

$$\min \big\{ f(S) - g(S) \, : \, S \subseteq U \big\}, \tag{3}$$

and its close relatives

$$\min\left\{f(S)/g(S)\,:\,S\subseteq U\right\},\tag{4}$$

$$\min\left\{f(S)\,:\,g(S)\geq c, S\subseteq U\right\},\tag{5}$$

$$\max\left\{g(S)\,:\,f(S)\leq b, S\subseteq U\right\},\tag{6}$$

with $b, c \in \mathbb{R}_{\geq 0}$ constants.

The first structural results and algorithms for Problem (3) were given for some special cases of f and g in [14]. It was later shown in [41, 29] that every set function $h : 2^U \to \mathbb{R}$ can be expressed as a difference of submodular functions, thus implying that Problem (3) is NP-hard in general. Furthermore, [29] showed that Problem (3) is inapproximable even when f and g are increasing and even the construction of f and g from h can take exponential time in the worst case. The authors also showed, however, that having a "good" (somewhat far from trivial) lower-bound on h allows one to compute f and g in polynomial time. Clearly, finding such bound is a hard task in itself. Our first questions have a structural flavor:

Question 2.1. *Can we characterize some non-trivial subclasses of set functions that can be expressed as difference of bounded algebraic-degree, symmetric, or covering submodular functions? What if, under a suitable metric, h itself is "close" to being submodular (as* quasi-submodular*,* pseudo-submodular*, or $M^\natural$-*convex*)?*

Affirmative answers span algorithmic ones:

Question 2.2. *Do those characterizations provide polynomial-time algorithms to construct such decompositions? Do they allow us to provide more efficient exact or better approximate algorithms for Problem (3) when restricted to those sub-classes?*

Another way of looking at Question 2.2 would be: can we leverage on the maximum-flow techniques [25, 6, 8, 3, 61] and [51, 53, 40] mentioned in Sections 1.1 and 1.2 or even on submodular-flow techniques (see [19] and references there in) to solve Problem (3) in those restricted settings? Why or why not?

Alongside with the restricted positive answer for decomposition, the authors provide in [29] some approximation algorithms for Problem (3) (in some settings) based on lower and upper modular bounds and the minimization-maximization method imported from convex optimization. In [1], approximation algorithms for Problem (4) based on fractional optimization methods are given. As Problem (4) can be easily recast into Problem (3), we are left with:

Question 2.3. *How do the inner workings of such algorithms compare on our proposed restricted settings? Do they really belong to different algorithmic-methodology classes and if so, can they ideas be crossed and produce a hybrid?*

While Problem (3) is inapproximable even for increasing submodular functions f and g, Problems (5) and (6) are related to each other and admit constant factor approximation algorithms with f and g under the same conditions. As Problem (3) encompasses Problems (5) and (6), there is clearly a deep question at stake:

Question 2.4. *The Lagrangian relaxation used in recasting Problems* (5) *and* (6) *into Problem* (3) *is introducing an unquantifiable duality-gap. Does this also happen in our restricted settings? Can we adapt the gauge-duality concept from convex functions to submodular ones and benefit from it?*

To the best of the adviser's knowledge, the above questions have not yet been investigated and answers shall not only advance our knowledge in the field, but may have fruitful impacts in machine learning applications.

3 Goals

This project assumes the student has minimal to no research experience whatsoever. Its general goals are then to expose the student to a research environment: the mindset of a researcher; some of the techniques and methodologies commonly used; how to detect roadblocks and how to (try to) bypass them; how to carefully check the validity of one's results; how to read (and possibly write) academic papers; how to research in an ethical manner, to back up one's claims with proofs or evidence, to avoid sheer speculations and plagiarism.

The project will focus on the questions stated in Section 2. The adviser anticipate the following specific goals to be achieved: the strengthening of the student's ability to think in a clear, concise, and formal way; the increase in the student's abilities to navigate through more abstract settings; the acquisition, from the student's perspective, of fundamental results in combinatorics, computational complexity, optimization, and theory of submodular functions; the improvement of the student's coding skills, and the finding of some answers to the questions proposed.

Furthermore, the student will have the opportunity to interact with other advisees, some of which are working on related aspects of submodularity. It is expected for this interaction to be highly beneficial to the student.

4 Methodology

The methodology of this project will be the one compatible with the mathematical nature of the subject:

(1) Adviser and student will have office meetings – at first, every other week; later, on a weekly basis – to discuss the papers or book chapters read by the latter since the last meeting. The adviser will provide clarifications and suggest new material and/or directions to be worked upon.

(2) The student will write about the topics learned in order to develop their writing skills. Focus will be given to precise definitions, correct contextualization, clear and concise proofs of the main results, and the correctness and quality of any computer code that may be written. The final report will be made available under an open-access policy.

(3) At the adviser's discretion, the student may be asked to give a talk of her results in a seminar series run by the center. Any new result will be submitted to publication in conferences and/or international journals.

5 Schedule

The student contacted the adviser for a research experience for undergraduates project at the end of the first academic term of 2022 (Q1). Upon agreement on the topic and under the adviser's guidance, the student already started reading on prerequisite material as maximum-flows/minimum-cuts in graphs and integrality of polyhedra. Besides showing engagement to the project, this shall reduce the knowledge gap in the initial stages and accelerate his entrance in research mode.

Topic	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Ago
Prerequisites	*	*										
Submodularity		*	*	*	*	*	*	*	*	*	*	
Discrete Convexity					*	*	*	*	*	*	*	
Coding and Testing				*	*	*	*	*	*	*	*	
Writing		*	*	*	*	*	*	*	*	*	*	*

References

[1] Bai, W., Iyer, R. K., Wei, K., and Bilmes, J. A. Algorithms for optimizing the ratio of submodular functions. In *Proceedings of the 33nd International Conference on Machine Learning, ICML 2016, New York City, NY, USA, June 19-24, 2016* (2016), M. Balcan and K. Q. Weinberger, Eds., vol. 48 of *JMLR Workshop and Conference Proceedings*, JMLR.org, pp. 2751–2759. → 1 and 4

[2] Baptiste, P., Carlier, J., Kononov, A. V., Queyranne, M., Sevastyanov, S., and Sviridenko, M. Integrality property in preemptive parallel machine scheduling. In *Computer Science - Theory and Applications, Fourth International Computer Science Symposium in Russia, CSR 2009, Novosibirsk, Russia, August 18-23, 2009. Proceedings* (2009), A. E. Frid, A. Morozov, A. Rybalchenko, and K. W. Wagner, Eds., vol. 5675 of *Lecture Notes in Computer Science*, Springer, pp. 38–46. → 1

[3] Billionnet, A., and Minoux, M. Maximizing a supermodular pseudoboolean function: A polynomial algorithm for supermodular cubic functions. *Discrete Applied Mathematics 12*, 1 (1985), 1 – 11. → 2, 3, and 4

[4] Blake, A., Kohli, P., Rother, C., Boykov, Y., Kolmogorov, V., Veksler, O., Zabih, R., Ishikawa, H., Yuille, A., Weiss, Y., et al. *Markov Random Fields for Vision and Image Processing*. MIT Press. MIT Press, 2011. → 1

[5] Bodirsky, M., Mamino, M., and Viola, C. Submodular functions and valued constraint satisfaction problems over infinite domains. *CoRR abs/1804.01710* (2018). → 1

[6] Boros, E., and Hammer, P. L. A max-flow approach to improved roof-duality in quadratic $0-1$ minimization. RUTCOR Research Report RRR 15-89, Rutgers University, New Brunswick, NJ, USA, 1989. → 2, 3, and 4

[7] Boros, E., and Hammer, P. L. Pseudo-boolean optimization. *Discrete Applied Mathematics 123*, 1-3 (2002), 155–225. → 3

[8] Boros, E., Hammer, P. L., Sun, R., and Tavares, G. A max-flow approach to improved lower bounds for quadratic unconstrained binary optimization (QUBO). *Discrete Optimization 5*, 2 (2008), 501–529. → 2, 3, and 4

[9] Bramel, J., and Simchi-Levi, D. *The logic of logistics : theory, algorithms, and applications for logistics management*. Springer, New York, 1997. → 1

[10] Buchbinder, N., Feldman, M., Naor, J., and Schwartz, R. A tight linear time (1/2)-approximation for unconstrained submodular maximization. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012* (2012), IEEE Computer Society, pp. 649–658. → 3

[11] Chakrabarty, D., Lee, Y. T., Sidford, A., and Wong, S. C. Subquadratic submodular function minimization. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017* (2017), H. Hatami, P. McKenzie, and V. King, Eds., ACM, pp. 1220–1231. → 2

[12] Crama, Y. Recognition problems for special classes of polynomials in $0-1$ variables. *Mathematical Programming 44* (1989), 139–155. 10.1007/BF01587085. → 3

[13] Crama, Y., and Hammer, P. L. *Boolean functions: Theory, algorithms, and applications*, vol. 142 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2011. → 3

[14] Cunningham, W. H. Decompostiton of submodular functions. *Comb. 3*, 1 (1983), 53–68. → 1 and 4

[15] Feige, U. A threshold of ln n for approximating set cover. *J. ACM 45*, 4 (1998), 634–652. → 3

[16] Fleischer, L., and Iwata, S. A push-relabel framework for submodular function minimization and applications to parametric optimization. *Discrete Appl. Math. 131*, 2 (2003), 311–322. Submodularity. → 2

[17] Freedman, D., and Drineas, P. Energy minimization via graph cuts: Settling what is possible. In *CVPR (2)* (2005), IEEE Computer Society, pp. 939–946. → 3

[18] Fujishige, S. *Submodular functions and optimization*, second ed., vol. 58 of *Annals of Discrete Mathematics*. Elsevier B. V., Amsterdam, 2005. → 1

[19] Fujishige, S. *Submodular functions and optimization*. Elsevier, 2005. → 4

[20] Gondran, M., and Minoux, M. *Graphs and algorithms*. Wiley-Interscience Series in Discrete Mathematics. John Wiley & Sons Ltd., Chichester, 1984. Translated from the French by Steven Vajda, A Wiley-Interscience Publication. → 1

[21] Grötschel, M., Lovász, L., and Schrijver, A. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica 1*, 2 (1981), 169–197. → 2

[22] Grötschel, M., Lovász, L., and Schrijver, A. Corrigendum to our paper: "The ellipsoid method and its consequences in combinatorial optimization" [Combinatorica **1** (1981), no. 2, 169–197; MR0625550 (84a:90044)]. *Combinatorica 4*, 4 (1984), 291–295. → 2

[23] Grötschel, M., Lovász, L., and Schrijver, A. *Geometric algorithms and combinatorial optimization*, vol. 2 of *Algorithms and Combinatorics: Study and Research Texts*. Springer-Verlag, Berlin, 1988. → 2

[24] Guestrin, C., Krause, A., and Singh, A. P. Near-optimal sensor placements in gaussian processes. In *Machine Learning, Proceedings of the Twenty-Second International Conference (ICML 2005), Bonn, Germany, August 7-11, 2005* (2005), L. D. Raedt and S. Wrobel, Eds., vol. 119 of *ACM International Conference Proceeding Series*, ACM, pp. 265–272. → 1

[25] Hammer, P., and Rudeanu, S. *Boolean Methods in Operations Research and Related Areas*. Springer Verlag, 1968. → 2, 3, and 4

[26] Iwata, S. A fully combinatorial algorithm for submodular function minimization. *J. Combin. Theory Ser. B 84*, 2 (2002), 203–212. → 2

[27] Iwata, S. A faster scaling algorithm for minimizing submodular functions. *SIAM J. Comput. 32*, 4 (2003), 833–840 (electronic). → 2

[28] Iwata, S., Fleischer, L., and Fujishige, S. A combinatorial strongly polynomial algorithm for minimizing submodular functions. *J. ACM 48*, 4 (2001), 761–777 (electronic). → 2

[29] Iyer, R. K., and Bilmes, J. A. Algorithms for approximate minimization of the difference between submodular functions, with applications. *CoRR abs/1207.0560* (2012). → 1 and 4

[30] Iyer, R. K., and Bilmes, J. A. Submodular optimization with submodular cover and submodular knapsack constraints. In *Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013, Lake Tahoe, Nevada, United States* (2013), C. J. C. Burges, L. Bottou, Z. Ghahramani, and K. Q. Weinberger, Eds., pp. 2436–2444. → 1 and 3

[31] Iyer, R. K., and Bilmes, J. A. Concave aspects of submodular functions. In *IEEE International Symposium on Information Theory, ISIT 2020, Los Angeles, CA, USA, June 21-26, 2020* (2020), IEEE, pp. 72–77. → 2

[32] Krause, A., and Golovin, D. Submodular function maximization. In *Tractability: Practical Approaches to Hard Problems*, L. Bordeaux, Y. Hamadi, and P. Kohli, Eds. Cambridge University Press, 2014, pp. 71–104. → 3

[33] Krause, A., and Guestrin, C. Near-optimal nonmyopic value of information in graphical models. In *UAI '05, Proceedings of the 21st Conference in Uncertainty in Artificial Intelligence, Edinburgh, Scotland, July 26-29, 2005* (2005), AUAI Press, pp. 324–331. → 1

[34] Krause, A., Singh, A. P., and Guestrin, C. Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. *J. Mach. Learn. Res. 9* (2008), 235–284. → 1

[35] Lee, J., Mirrokni, V. S., Nagarajan, V., and Sviridenko, M. Non-monotone submodular maximization under matroid and knapsack constraints. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009* (2009), M. Mitzenmacher, Ed., ACM, pp. 323–332. → 3

[36] Lee, J., Sviridenko, M., and Vondrák, J. Submodular maximization over multiple matroids via generalized exchange properties. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 12th International Workshop, APPROX 2009, and 13th International Workshop, RANDOM 2009, Berkeley, CA, USA, August 21-23, 2009. Proceedings* (2009), I. Dinur, K. Jansen, J. Naor, and J. D. P. Rolim, Eds., vol. 5687 of *Lecture Notes in Computer Science*, Springer, pp. 244–257. → 3

[37] Lin, H., and Bilmes, J. A. Multi-document summarization via budgeted maximization of submodular functions. In *Human Language Technologies: Conference of the North American Chapter of the Association of Computational Linguistics, Proceedings, June 2-4, 2010, Los Angeles, California, USA* (2010), The Association for Computational Linguistics, pp. 912–920. → 1 and 3

[38] Lin, H., and Bilmes, J. A. A class of submodular functions for document summarization. In *The 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference, 19-24 June, 2011, Portland, Oregon, USA* (2011), D. Lin, Y. Matsumoto, and R. Mihalcea, Eds., The Association for Computer Linguistics, pp. 510–520. → 1

[39] Margot, F., Queyranne, M., and Wang, Y. Decompositions, network flows, and a precedence constrained single-machine scheduling problem. *Oper. Res. 51*, 6 (2003), 981–992. → 1

[40] Nagamochi, H., and Ibaraki, T. A note on minimizing submodular functions. *Inform. Process. Lett. 67*, 5 (1998), 239–244. → 2 and 4

[41] Narasimhan, M., and Bilmes, J. A. A submodular-supermodular procedure with applications to discriminative structure learning. In *UAI '05, Proceedings of the 21st Conference in Uncertainty in Artificial Intelligence, Edinburgh, Scotland, July 26-29, 2005* (2005), AUAI Press, pp. 404–412. → 1 and 4

[42] Narayanan, H. *Submodular Functions and Electrical Networks.* Annals of discrete mathematics. Elsevier, 1997. → 1

[43] Nemhauser, G., and Wolsey, L. Maximizing submodular set functions: formulations and analysis of algorithms. *In Studies of Graphs and Discrete Programming* (1981), 279–301. → 2 and 3

[44] Nemhauser, G. L., and Wolsey, L. A. Best algorithms for approximating the maximum of a submodular set function. *Math. Oper. Res. 3*, 3 (1978), 177–188. → 3

[45] Nemhauser, G. L., Wolsey, L. A., and Fisher, M. L. An analysis of approximations for maximizing submodular set functions. I. *Math. Programming 14*, 3 (1978), 265–294. → 2 and 3

[46] Nisan, N., Roughgarden, T., Tardos, E., and Vazirani, V. V. *Algorithmic Game Theory.* Cambridge University Press, New York, NY, USA, 2007. → 1

[47] Orlin, J. B. A faster strongly polynomial time algorithm for submodular function minimization. *Math. Program. 118*, 2 (2009), 237–251. → 2

[48] Oxley, J. G. *Matroid theory.* Oxford University Press, 1992. → 1

[49] Pearl, J. *Probabilistic reasoning in intelligent systems - networks of plausible inference.* Morgan Kaufmann series in representation and reasoning. Morgan Kaufmann, 1989. → 1

[50] Plummer, M., and Lovász, L. *Matching Theory.* North-Holland Mathematics Studies. Elsevier Science, 1986. → 1

[51] Queyranne, M. Minimizing symmetric submodular functions. *Math. Programming 82*, 1-2, Ser. B (1998), 3–12. Networks and matroids; Sequencing and scheduling. → 2 and 4

[52] Recski, A. *Matroid Theory and its Applications in Electric Network Theory and in Statics.* Algorithms and Combinatorics. Springer Berlin Heidelberg, 2013. → 1

[53] Rizzi, R. On minimizing symmetric set functions. *Combinatorica 20*, 3 (2000), 445–450. → 2 and 4

[54] Rockafellar, R. T. *Convex Analysis.* Princeton Landmarks in Mathematics and Physics. Princeton University Press, 1970. → 2

[55] Schipper, B. C. Submodularity and the evolution of Walrasian behavior. *International Journal of Game Theory 32*, 4 (August 2004), 471–477. → 1

[56] Schrijver, A. A combinatorial algorithm minimizing submodular functions in strongly polynomial time. *J. Combin. Theory Ser. B 80*, 2 (2000), 346–355. → 2

[57] Schrijver, A. *Combinatorial optimization. Polyhedra and efficiency. Vol. B*, vol. 24 of *Algorithms and Combinatorics.* Springer-Verlag, Berlin, 2003. Matroids, trees, stable sets, Chapters 39–69. → 1

[58] Sviridenko, M. A note on maximizing a submodular set function subject to a knapsack constraint. *Oper. Res. Lett. 32*, 1 (2004), 41–43. → 3

[59] Topkis, D. *Supermodularity and Complementarity.* Frontiers of Economic Research. Princeton University Press, 2011. → 1

[60] Vygen, J. A note on Schrijver's submodular function minimization algorithm. *J. Combin. Theory Ser. B 88*, 2 (2003), 399–402. → 2

[61] Wang, D., and Kleinberg, R. Analyzing quadratic unconstrained binary optimization problems via multicommodity flows. *Discrete Appl. Math. 157*, 18 (2009), 3746–3753. → 2, 3, and 4

[62] Živný, S., and Jeavons, P. G. Classes of submodular constraints expressible by graph cuts. *Constraints 15*, 3 (2010), 430–452. → 3