

# 1 References and further reading

## Tutorials, talks and links

- by Jeff Bilmes (at NIPS):  
<http://nips.cc/Conferences/2013/Program/event.php?ID=3688>
- by Andreas Krause and Stefanie Jegelka (at ICML):  
[http://www.cs.berkeley.edu/~stefje/submodularity\\_icml.html](http://www.cs.berkeley.edu/~stefje/submodularity_icml.html)
- by Francis Bach (MLSS):  
[http://www.di.ens.fr/~fbach/submodular\\_fbach\\_mlss2012.pdf](http://www.di.ens.fr/~fbach/submodular_fbach_mlss2012.pdf)
- by Jan Vondrák (Modern Aspects of Submodularity workshop, 2012):  
<http://researcher.watson.ibm.com/researcher/files/us-jvondrak/submod-tutorial-1.pdf> and <http://researcher.watson.ibm.com/researcher/files/us-jvondrak/submod-tutorial-2.pdf>
- talks about various topics related to machine learning and discrete optimization: <http://las.ethz.ch/discml/videos.html>
- links to workshops and tutorials: [submodularity.org](http://submodularity.org)
- list of more submodularity papers than listed here: <http://www.cs.berkeley.edu/~stefje/references.html>

## Early papers on submodularity; books and surveys

- [Bac13] F. Bach. *Learning with Submodular Functions: A Convex Optimization Perspective*. Foundations and Trends in Machine Learning, 2013.
- [Cho54] G. Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295, 1954.
- [Edm70] J. Edmonds. *Combinatorial Structures and Their Applications*, chapter Submodular Functions, Matroids and Certain Polyhedra, pages 69–87. Gordon and Breach, 1970.
- [Fuj05] S. Fujishige. *Submodular functions and optimization*. Number 58 in Annals of Discrete Mathematics. Elsevier Science, 2 edition, 2005.
- [KG14] Andreas Krause and Daniel Golovin. *Tractability: Practical Approaches to Hard Problems*, chapter Submodular Function Maximization. Cambridge University Press, 2014.
- [Lov83] L. Lovász. *Mathematical programming – The State of the Art*, chapter Submodular Functions and Convexity, pages 235–257. Springer, 1983.
- [Sha71] L. S. Shapley. Cores of convex games. *International Journal of Game Theory*, 1(1):11–26, 1971.

## Submodularity and convexity, submodular minimization

- [Bac10] F. Bach. Structured sparsity-inducing norms through submodular functions. In *Advances in Neural Information Processing Systems (NIPS)*, 2010.
- [Bac13] F. Bach. *Learning with Submodular Functions: A Convex Optimization Perspective*. Foundations and Trends in Machine Learning, 2013.
- [CD09] A. Chambolle and J. Darbon. On total variation minimization and surface evolution using parametric maximum flows. *Int. Journal of Computer Vision*, 84(3), 2009.
- [Edm70] J. Edmonds. *Combinatorial Structures and Their Applications*, chapter Submodular Functions, Matroids and Certain Polyhedra, pages 69–87. Gordon and Breach, 1970.

- [FI11] S. Fujishige and S. Isotani. A submodular function minimization algorithm based on the minimum-norm base. *Pacific Journal of Optimization*, 7:3–17, 2011.
- [Fuj05] S. Fujishige. *Submodular functions and optimization*. Number 58 in Annals of Discrete Mathematics. Elsevier Science, 2 edition, 2005.
- [GJ85] F. Granot and A. F. Veinott Jr. Substitutes, complements, and ripples in network flows. *Math. of OR*, 10:471–497, 1985.
- [GLS81] M. Grötschel, L. Lovász, and A. Schrijver. The ellipsoid algorithm and its consequences in combinatorial optimization. *Combinatorica*, 1:499–513, 1981.
- [GLS84] M. Grötschel, L. Lovász, and A. Schrijver. Corrigendum to the paper “the ellipsoid algorithm and its consequences in combinatorial optimization”. *Combinatorica*, 4:291–295, 1984.
- [Hoc01] D.S. Hochbaum. An efficient algorithm for image segmentation, Markov random fields and related problems. *Journal of the ACM*, 48(4):686–701, 2001.
- [IFF01] S. Iwata, L. Fleischer, and S. Fujishige. A combinatorial strongly polynomial algorithm for minimizing submodular functions. *Journal of the ACM*, 48:761–777, 2001.
- [Lov83] L. Lovász. *Mathematical programming – The State of the Art*, chapter Submodular Functions and Convexity, pages 235–257. Springer, 1983.
- [McC05] S. T. McCormick. Submodular function minimization. *Discrete Optimization*, 12:321–391, 2005.
- [Orl09] J. B. Orlin. A faster strongly polynomial time algorithm for submodular function minimization. *Mathematical Programming*, 118(2):237–251, 2009.
- [Que98] W. Queyranne. Minimizing symmetric submodular functions. *Mathematical Programming*, 82:3–12, 1998.
- [Sch00] A. Schrijver. A combinatorial algorithm minimizing submodular functions in strongly polynomial time. *J. Combin. Theory Ser. B*, 80:346–355, 2000.
- [Top78] D.M. Topkis. Minimizing a submodular function on a lattice. *Operations Research*, 26:305–321, 1978.
- [Top98] D.M. Topkis. *Supermodularity and Complementarity*. Princeton University Press, 1998.

## Minimizing a sum of submodular functions

- [EN15] A. Ene and H. Nguyen. Random coordinate descent methods for minimizing decomposable submodular functions. In *Int. Conference on Machine Learning (ICML)*, 2015.
- [JBS13] S. Jegelka, F. Bach, and S. Sra. Reflection methods for user-friendly submodular optimization. In *Advances in Neural Information Processing Systems (NIPS)*, 2013.
- [Kol12] V. Kolmogorov. Minimizing a sum of submodular functions. *Discrete Applied Mathematics*, 160(15):2246–2258, 2012.
- [KPT11] N. Komodakis, N. Paragios, and G. Tziritas. MRF energy minimization and beyond via dual decomposition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 33(3):531–552, 2011.
- [NJJ14] N. Nishihara, S. Jegelka, and M.I. Jordan. On the linear convergence rate of decomposable submodular function minimization. In *Advances in Neural Information Processing Systems (NIPS)*, 2014.
- [SK10] P. Stobbe and A. Krause. Efficient minimization of decomposable submodular functions. In *Advances in Neural Information Processing Systems (NIPS)*, 2010.
- [SSKS11] B. Savchynskyy, S. Schmidt, J. Kappes, and C. Schnörr. A study of Nesterov’s scheme for Lagrangian decomposition and MAP labeling. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.
- [SSKS12] B. Savchynskyy, S. Schmidt, J.H. Kappes, and C. Schnörr. Efficient MRF energy minimization via adaptive diminishing smoothing. In *Uncertainty in Artificial Intelligence (UAI)*, 2012.

## Frank Wolfe / Conditional gradient

- [Bac13] F. Bach. Conditional gradients everywhere. NIPS workshop talk, 2013. <http://www.cmap.polytechnique.fr/~jaggi/NIPS-workshop-FW-greedy/slides/slides-bach.pdf>.
- [FW56] M. Frank and P. Wolfe. An algorithm for quadratic programming. *Naval Research Logistics Quarterly*, 3:95–100, 1956.
- [Jag13] M. Jaggi. Revisiting Frank-Wolfe: Projection-free sparse convex optimization. In *Int. Conference on Machine Learning (ICML)*, 2013.

## Submodular Maximization

- [BFNS12] N. Buchbinder, M. Feldman, J. Naor, and R. Schwartz. A tight linear time (1/2)-approximation for unconstrained submodular maximization. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, 2012.
- [BV14] A. Badanidiyuru and J. Vondrák. Fast algorithms for maximizing submodular functions. In *SIAM-ACM Symposium on Discrete Algorithms (SODA)*, 2014.
- [CC84] M. Conforti and G. Cornuéjols. Submodular set functions, matroids and the greedy algorithm: tight worst-case bounds and some generalizations of the rado-edmonds theorem. *Discrete Applied Math.*, 7(3):251–274, 1984 1984.
- [CCPV11] G. Calinescu, C. Chekuri, M. Pál, and J. Vondrák. Maximizing a submodular set function subject to a matroid constraint. *SIAM J. Computing*, 40(6), 2011.
- [dPBENW15] Rafael da Ponte Barbosa, Alina Ene, Huy L. Nguyen, and Justin Ward. The power of randomization: Distributed submodular maximization on massive datasets. In *Int. Conference on Machine Learning (ICML)*, 2015.
- [FMV07] U. Feige, V. Mirrokni, and J. Vondrák. Maximizing non-monotone submodular functions. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, 2007.
- [FW14] Y. Filmus and J. Ward. Monotone submodular maximization over a matroid via non-oblivious local search. *SIAM J. Comput.*, 43(2):514–542, 2014.
- [KG14] Andreas Krause and Daniel Golovin. *Tractability: Practical Approaches to Hard Problems*, chapter Submodular Function Maximization. Cambridge University Press, 2014.
- [KMVV13] Ravi Kumar, Benjamin Moseley, Sergei Vassilvitskii, and Andrea Vattani. Fast greedy algorithms in MapReduce and streaming. In *25th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, 2013.
- [LKG<sup>+</sup>07] Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutsos, Jeanne VanBriesen, and Natalie Glance. Cost-effective outbreak detection in networks. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, pages 420–429, August 2007.
- [MBK<sup>+</sup>15] B. Mirzasoleiman, A. Badanidiyuru, A. Karbasi, J. Vondrák, and A. Krause. Lazier than lazy greedy. In *Proc. Conf. on Artificial Intelligence (AAAI)*, 2015.
- [MKS13] B. Mirzasoleiman, A. Karbasi, R. Sarkar, and A. Krause. Distributed submodular maximization: Identifying representative elements in massive data. In *Advances in Neural Information Processing Systems (NIPS)*, 2013.
- [NWF78] G.L. Nemhauser, L.A. Wolsey, and M.L. Fisher. An analysis of approximations for maximizing submodular set functions—I. *Mathematical Programming*, 14(1):265–294, 1978.
- [PJG<sup>+</sup>14] X. Pan, S. Jegelka, J. Gonzalez, J. Bradley, and M.I. Jordan. Parallel double greedy submodular maximization. In *Advances in Neural Information Processing Systems (NIPS)*, 2014.

- [Svi04] M. Sviridenko. A note on maximizing a submodular set function subject to knapsack constraint. *Operations Research Letters*, 31:41–43, 2004.
- [SVW15] Maxim Sviridenko, Jan Vondrák, and Justin Ward. Optimal approximation for submodular and supermodular optimization with bounded curvature. In *SIAM-ACM Symposium on Discrete Algorithms (SODA)*, 2015.
- [Von10] J. Vondrák. Submodularity and curvature: the optimal algorithm. Technical Report 23, RIMS Kokyuroku Bessatsu, 2010.

## Constrained submodular optimization

- [GHIM09] M.X. Goemans, N.J.A. Harvey, S. Iwata, and V. Mirrokni. Approximating submodular functions everywhere. In *SIAM-ACM Symposium on Discrete Algorithms (SODA)*, 2009.
- [GKTW09] G. Goel, C. Karande, P. Tripathi, and L. Wang. Approximability of combinatorial problems with multi-agent submodular cost functions. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, 2009.
- [IB13] R.K. Iyer and J. Bilmes. Submodular optimization with submodular cover and submodular knapsack constraints. In *Advances in Neural Information Processing Systems (NIPS)*, 2013.
- [IJB13a] R.K. Iyer, S. Jegelka, and J. Bilmes. Curvature and optimal algorithms for learning and minimizing submodular functions. In *Advances in Neural Information Processing Systems (NIPS)*, 2013.
- [IJB13b] R.K. Iyer, S. Jegelka, and J. Bilmes. Fast semidifferential-based submodular function optimization. In *Int. Conference on Machine Learning (ICML)*, 2013.
- [IN09] S. Iwata and K. Nagano. Submodular function minimization under covering constraints. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, 2009.
- [JB11a] S. Jegelka and J. Bilmes. Approximation bounds for inference using cooperative cuts. In *Int. Conference on Machine Learning (ICML)*, 2011.
- [JB11b] S. Jegelka and J. Bilmes. Submodularity beyond submodular energies: coupling edges in graph cuts. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.
- [KOJ13] P. Kohli, A. Osokin, and S. Jegelka. A principled deep random field for image segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.

## Examples of submodular functions and applications

- [Bac11] F. Bach. Shaping level sets with submodular functions. In *Advances in Neural Information Processing Systems (NIPS)*, 2011.
- [BJ01] Y. Boykov and M.-P. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images. In *Int. Conference on Computer Vision (ICCV)*, 2001.
- [BLK12] Olga Barinova, Victor Lempitsky, and Pushmeet Kohli. On detection of multiple object instances using Hough transforms. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 9:1773–1784, 2012.
- [CG07] A. Chechetka and C. Guestrin. Efficient learning of thin junction trees. In *Advances in Neural Information Processing Systems (NIPS)*, 2007.
- [DK14] J. Djolonga and A. Krause. From MAP to marginals: Variational inference in Bayesian submodular models. In *Advances in Neural Information Processing Systems (NIPS)*, 2014.
- [DK15] J. Djolonga and A. Krause. Scalable variational inference in log-supermodular models. In *Int. Conference on Machine Learning (ICML)*, 2015.

- [DVOB12] A. DeLong, O. Veksler, A. Osokin, and Y. Boykov. Minimizing sparse high-order energies by submodular vertex-cover. In *Advances in Neural Information Processing Systems (NIPS)*, 2012.
- [GGG15] Michael Gygli, Helmut Grabner, and Luc Van Gool. Video summarization by learning submodular mixtures of objectives. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [GK10] R. Gomes and A. Krause. Budgeted nonparametric learning from data streams. In *ICML*, 2010.
- [GK11] D. Golovin and A. Krause. Adaptive submodularity: Theory and applications in active learning and stochastic optimization. *Journal of Artificial Intelligence Research*, 42:427–486, 2011.
- [GKT12] J. Gillenwater, A. Kulesza, and B. Taskar. Near-optimal map inference for determinantal point processes. In *NIPS*, 2012.
- [GRLK12] M. Gomez-Rodriguez, J. Leskovec, and A. Krause. Inferring networks of diffusion and influence. *ACM Transactions on Knowledge Discovery from Data*, 5(4), 2012.
- [JB11] S. Jegelka and J. Bilmes. Submodularity beyond submodular energies: coupling edges in graph cuts. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.
- [KG11] A. Krause and C. Guestrin. Submodularity and its applications in optimized information gathering. *ACM Trans. on Int. Systems and Technology*, 2(4), 2011.
- [KKT03] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 2003.
- [KLT09] P. Kohli, L. Ladický, and P.H.S. Torr. Robust higher order potentials for enforcing label consistency. *Int. Journal of Computer Vision*, 82(3):302–324, 2009.
- [KSX14] Gunhee Kim, Leonid Sigal, and Eric P. Xing. Joint summarization of large-scale collections of web images and videos for storyline reconstruction. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
- [KT10] A. Kulesza and B. Taskar. Structured determinantal point processes. In *Advances in Neural Information Processing Systems (NIPS)*, 2010.
- [KXFFK11] G. Kim, E.P. Xing, L. Fei-Fei, and T. Kanade. Distributed cosegmentation via submodular optimization on anisotropic diffusion. In *Int. Conference on Computer Vision (ICCV)*, 2011.
- [KZ04] V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26(2):147–159, 2004.
- [LB11] H. Lin and J. Bilmes. A class of submodular functions for document summarization. In *The 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL/HLT-2011)*, 2011.
- [LB12] H. Lin and J. Bilmes. Learning mixtures of submodular shells with application to document summarization. In *Uncertainty in Artificial Intelligence (UAI)*, 2012.
- [LKG<sup>+</sup>07] Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutsos, Jeanne VanBriesen, and Natalie Glance. Cost-effective outbreak detection in networks. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, pages 420–429, August 2007.
- [LWK<sup>+</sup>13] Yuzong Liu, Kai Wei, Katrin Kirchhoff, Yisong Song, and Jeff Bilmes. Submodular feature selection for high-dimensional acoustic score spaces. In *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing*, 2013.
- [MR07] E. Mossel and S. Roch. On the submodularity of influence in social networks. In *Symposium on Theory of Computing (STOC)*, 2007.
- [NJB05] M. Narasimhan, N. Jojic, and J. Bilmes. Q-clustering. In *Advances in Neural Information Processing Systems (NIPS)*, 2005.

- [PJB14] A. Prasad, S. Jegelka, and D. Batra. Submodular meets structured: Finding diverse subsets in exponentially-large structured item sets. In *Advances in Neural Information Processing Systems (NIPS)*, 2014.
- [RG13] C. Reed and Z. Ghahramani. Scaling the indian buffet process via submodular maximization. In *ICML*, 2013.
- [SGJ<sup>+</sup>14] H.O. Song, R. Girshick, S. Jegelka, J. Mairal, Z. Harchaoui, and T. Darrell. On learning to localize objects with minimal supervision. In *Int. Conference on Machine Learning (ICML)*, 2014.
- [SLJD14] H. Song, Y.J. Lee, S. Jegelka, and T. Darrell. Weakly-supervised discovery of visual pattern configurations. In *Advances in Neural Information Processing Systems (NIPS)*, 2014.
- [SSGK14] Nathan Silberman, Lior Shapira, Ran Gal, and Pushmeet Kohli. A contour completion model for augmenting surface reconstructions. In *Europ. Conference on Computer Vision (ECCV)*, 2014.
- [TIWB14] S. Tschitschek, R. Iyer, H. Wei, and J. Bilmes. Learning mixtures of submodular functions for image collection summarization. In *Advances in Neural Information Processing Systems (NIPS)*, 2014.
- [XML<sup>+</sup>15] Jia Xu, Lopamudra Mukherjee, Yin Li, Jamieson Warner, James M. Rehg, and Vikas Singh. Gaze-enabled egocentric video summarization via constrained submodular maximization. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [YG11] Yisong Yue and Carlos Guestrin. Linear submodular bandits and their application to diversified retrieval. In *NIPS*, 2011.

## Learning submodular functions, sketching and learning to optimize

- [BCIW12] M.-F. Balcan, F. Constantin, S. Iwata, and L. Wang. Learning valuation functions. In *Conference on Learning Theory (COLT)*, 2012.
- [BDF<sup>+</sup>12] A. Badanidiyuru, S. Dobzinski, H. Fu, R. Kleinberg, N. Nisan, and T. Roughgarden. Sketching valuation functions. In *SIAM-ACM Symposium on Discrete Algorithms (SODA)*, 2012.
- [BH11] M.-F. Balcan and N. Harvey. Learning submodular functions. In *STOC*, 2011.
- [BH12] N. Balcan and N. Harvey. Submodular functions: Learnability, structure, and optimization. *arXiv:0486478*, 2012.
- [CKKL12] M. Cheraghchi, A. Klivans, P. Kothari, and H. Lee. Submodular functions are noise stable. In *SIAM-ACM Symposium on Discrete Algorithms (SODA)*, 2012.
- [DLBS14a] N. Du, Y. Liang, M.-F. Balcan, and L. Song. Influence function learning in information diffusion networks. In *Int. Conference on Machine Learning (ICML)*, 2014.
- [DLBS14b] N. Du, Y. Liang, M.-F. Balcan, and L. Song. Learning time-varying coverage functions. In *Advances in Neural Information Processing Systems (NIPS)*, 2014.
- [FJPZ13] A. Fix, T. Joachims, S. Min Park, and R. Zabih. Structured learning of sum-of-submodular higher order energy functions. In *iccv*, 2013.
- [FKV13] Vitaly Feldman, Pravesh Kothari, and Jan Vondrak. Representation, approximation and learning of submodular functions using low-rank decision trees. In *Conference on Learning Theory (COLT)*, 2013.
- [FV13] Vitaly Feldman and Jan Vondrák. Optimal bounds on approximation of submodular and xos functions by juntas. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, 2013.
- [GB11] A. Guillory and J. Bilmes. Simultaneous learning and covering with adversarial noise. In *Int. Conference on Machine Learning (ICML)*, 2011.
- [GHIM09] M.X. Goemans, N.J.A. Harvey, S. Iwata, and V. Mirrokni. Approximating submodular functions everywhere. In *SIAM-ACM Symposium on Discrete Algorithms (SODA)*, 2009.

- [GHRU11] A. Gupta, M. Hardt, A. Roth, and J. Ullman. Privately releasing conjunctions and the statistical query barrier. In *Symposium on Theory of Computing (STOC)*, 2011.
- [GS08] D. Golovin and M. Streeter. Online algorithms for maximizing submodular set functions. In *Advances in Neural Information Processing Systems (NIPS)*, 2008.
- [HK12] E. Hazan and S. Kale. Online submodular minimization. *Journal of Machine Learning Research*, 13:2903–2922, 2012.
- [JB11] S. Jegelka and J. Bilmes. Online submodular minimization with combinatorial constraints. In *Int. Conference on Machine Learning (ICML)*, 2011.
- [LB12] H. Lin and J. Bilmes. Learning mixtures of submodular shells with application to document summarization. In *Uncertainty in Artificial Intelligence (UAI)*, 2012.
- [RY13] S. Raskhodnikova and G. Yaroslavtsev. Learning pseudo-boolean k-DNF and submodular functions. In *SIAM-ACM Symposium on Discrete Algorithms (SODA)*, 2013.
- [RZY<sup>+</sup>13] S. Ross, J. Zhou, Y. Yue, D. Dey, and D. Bagnell. Learning policies for contextual submodular prediction. In *ICML*, 2013.
- [SGK09] M. Streeter, D. Golovin, and A. Krause. Online learning of assignments. In *Advances in Neural Information Processing Systems (NIPS)*, 2009.
- [SK12] P. Stobbe and A. Krause. Learning Fourier sparse set functions. In *Proc. Int. Conference on Artificial Intelligence and Statistics (AISTATS)*, 2012.
- [TIWB14] S. Tschiatschek, R. Iyer, H. Wei, and J. Bilmes. Learning mixtures of submodular functions for image collection summarization. In *Advances in Neural Information Processing Systems (NIPS)*, 2014.
- [YG11] Yisong Yue and Carlos Guestrin. Linear submodular bandits and their application to diversified retrieval. In *NIPS*, 2011.