

## References

- [1] M. Hoffman, D. Blei, C. Wang, and J. Paisley. Stochastic variational inference. *JMLR*, 14:1303–1347, 2013.
- [2] R. Ranganath, C. Wang., D. Blei, and E. Xing. An adaptive learning rate for stochastic variational inference. In *ICML*, 2013.
- [3] P. Gopalan, D. M. Mimno, S. Gerrish, M. J. Freedman, and D. M. Blei. Scalable inference of overlapping communities. In *NIPS*, 2012.
- [4] M. Bryant and E. Sudderth. Truly nonparametric online variational inference for hierarchical Dirichlet processes. In *NIPS*, 2012.
- [5] S. Jain and R.M. Neal. A split-merge Markov chain Monte Carlo procedure for the Dirichlet process mixture model. *Journal of Computational and Graphical Statistics*, 13(1):158–182, 2004.
- [6] D. B. Dahl. Sequentially-allocated merge-split sampler for conjugate and nonconjugate Dirichlet process mixture models. *Submitted to Journal of Computational and Graphical Statistics*, 2005.
- [7] D. M. Blei and M. I. Jordan. Variational inference for Dirichlet process mixture models. *Bayesian Analysis*, 1(1):121–144, 2006.
- [8] Y. W. Teh, K. Kurihara, and M. Welling. Collapsed variational inference for HDP. In *NIPS*, 2008.
- [9] K. Kurihara, M. Welling, and N. Vlassis. Accelerated variational Dirichlet process mixtures. In *NIPS*, 2006.
- [10] R. M. Neal and G. E. Hinton. A view of the EM algorithm that justifies incremental, sparse, and other variants. In *Learning in graphical models*, 1999.
- [11] O. Papaspiliopoulos and G. O. Roberts. Retrospective Markov chain Monte Carlo methods for Dirichlet process hierarchical models. *Biometrika*, 95(1):169–186, 2008.
- [12] N. Goodman, V. Mansinghka, D. M. Roy, K. Bonawitz, and J. Tenenbaum. Church: A language for generative models. In *Uncertainty in Artificial Intelligence*, 2008.
- [13] N. Le Roux, M. Schmidt, and F. Bach. A stochastic gradient method with an exponential convergence rate for finite training sets. In *NIPS*, 2012.
- [14] N. Ueda and Z. Ghahramani. Bayesian model search for mixture models based on optimizing variational bounds. *Neural Networks*, 15(1):1223–1241, 2002.
- [15] C. Wang and D. Blei. Truncation-free stochastic variational inference for Bayesian nonparametric models. In *NIPS*, 2012.
- [16] N. Ueda, R. Nakano, Z. Ghahramani, and G. Hinton. SMEM algorithm for mixture models. *Neural Computation*, 12(9):2109–2128, 2000.
- [17] D. Arthur and S. Vassilvitskii. k-means++: The advantages of careful seeding. In *ACM-SIAM Symposium on Discrete Algorithms*, pages 1027–1035, 2007.
- [18] J. Xiao, J. Hays, K. Ehinger, A. Oliva, and A. Torralba. SUN database: Large-scale scene recognition from abbey to zoo. In *CVPR*, 2010.
- [19] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. In *ICCV*, 2011.
- [20] D. Zoran and Y. Weiss. Natural images, Gaussian mixtures and dead leaves. In *NIPS*, 2012.