Anima Anandkumar

California Institute of Technology
Computer & Mathematical Sciences,
Pasadena, CA, USA

Email: anima@caltech.edu
Homepage: http://tensorlab.cms.caltech.edu/users/anima/

Current Research Interests

Large-scale machine learning and high-dimensional statistics with a focus on deep learning and probabilistic models, distributed optimization and tensor methods.

Current Appointments

Bren Professor *Since July 2017.*
Computing+Math. Sciences Dept., California Institute of Technology.

Principal Scientist, *Since Oct. 2016.*
Amazon AI, Amazon Web Services (AWS).

Education

Doctor of Philosophy in Electrical Engineering with minor in Applied Mathematics *July 2009*
Electrical & Computer Engr, Cornell University.

Bachelor of Technology in Electrical Engineering with minor in Theoretical Comp. Science *May 2004*
Department of Electrical Engineering, Indian Institute of Technology Madras.

Awards and Honors

1. **Bren Endowed Chair Professorship at Caltech.**
2. **Google Faculty Research Award 2015.**
3. **AFOSR Young Investigator Award (YIP) 2015.**
4. **UCI Early Career Research Award 2015.**
5. **Alfred P. Sloan Research Fellowship 2014.**
6. **Microsoft Faculty Fellowship 2013.**
7. **ARO Young Investigator Award (YIP) 2013.**
8. **NSF CAREER Award 2013.**
9. **ACM SIGMETRICS 2011 Best Paper Award.**
10. **Best Thesis Award 2009** by ACM SIGMETRICS Society.

Previous Appointments

Associate Professor *Aug. 2016-June 2017.*
Electrical Engr & Computer Science, University of California.

Assistant Professor, EECS Dept., University of California, Irvine. *Aug. 2010-2016*

Visiting Researcher at Microsoft Research New England. *April-Dec. 2012*

Post-doctoral Associate at the Laboratory of Information & Decision Systems *July 2009-July 2010*
Massachusetts Institute of Technology.

Selected List of Publications

[1] Peiyun Hu, Zachary C Lipton, Animashree Anandkumar, and Deva Ramanan. Active learning with partial feedback. *arXiv preprint arXiv:1802.07427*, 2018.

[2] Forough Arabshahi, Sameer Singh, and Animashree Anandkumar. Combining Symbolic and Function Evaluation Expressions In Neural Programs. In *Proc. of International Conference on Learning Representation (ICLR)*, 2018.

[3] Ashish Khetan, Zachary C Lipton, and Animashree Anandkumar. Learning From Noisy Singly-labeled Data. In *Proc. of International Conference on Learning Representation (ICLR)*, 2018.

[4] Guneet S Dhillon, Kamyar Azizzadenesheli, Zachary C Lipton, Jeremy Bernstein, Jean Kossaifi, Aran Khanna, and Anima Anandkumar. Stochastic Activation Pruning for Robust Adversarial Defense. In *Proc. of International Conference on Learning Representation (ICLR)*, 2018.

[5] Yanyao Shen, Hyokun Yun, Zachary C Lipton, Yakov Kronrod, and Animashree Anandkumar. Deep active learning for named entity recognition. In *Proc. of International Conference on Learning Representation (ICLR)*, 2018.

[6] Jean Kossaifi, Zachary C Lipton, Aran Khanna, Tommaso Furlanello, and Animashree Anandkumar. Tensor regression networks. In *Proc. of NIPS workshop MLTrain* **Winner of best poster award**, 2017.

[7] Rose Yu, Stephan Zheng, Animashree Anandkumar, and Yisong Yue. Long-term forecasting using tensor-train rnns. In *Proc. of NIPS workshop on timeseries* **Winner of best paper award**, 2017.

[8] Tommaso Furlanello, Zachary C Lipton, AI Amazon, Laurent Itti, and Anima Anandkumar. Born again neural networks. In *NIPS Workshop on Meta Learning*, 2017.

[9] Michael Tschannen, Aran Khanna, and Anima Anandkumar. Strassennets: Deep learning with a multiplication budget. In *Proc. of NIPS Workshop MLTrain*, 2017.

[10] Kamyar Azizzadenesheli, Emma Brunskill, and Animashree Anandkumar. Efficient Exploration through Bayesian Deep Q-Networks. In *Proc. of NIPS workshop on Reinforcement Learning*, 2017.

[11] Jeremy Bernstein, Yu-Xiang Wang, Kamyar Azizzadenesheli, and Anima Anandkumar. signsgd: compressed optimisation for non-convex problems. In *Proc. of NIPS workshop on Optimization*, 2017.

[12] Anima Anandkumar, Yuan Deng, Rong Ge, and Hossein Mobahi. Homotopy analysis for tensor pca. In *Conference on Learning Theory*, pages 79–104, 2017.

[13] Kamyar Azizzadenesheli, Alessandro Lazaric, and Animashree Anandkumar. Reinforcement Learning in Rich-Observation MDPs using Spectral Methods. In *RLDM*, 2017.

[14] Forough Arabshahi and Anima Anandkumar. Spectral methods for correlated topic models. In *Artificial Intelligence and Statistics*, pages 1439–1447, 2017.

[15] Animashree Anandkumar, Rong Ge, and Majid Janzamin. Analyzing tensor power method dynamics in overcomplete regime. *Journal of Machine Learning Research*, 18(22):1–40, 2017.

[16] Yining Wang and Anima Anandkumar. Online and differentially-private tensor decomposition. In *Advances in Neural Information Processing Systems*, pages 3531–3539, 2016.

[17] Yang Shi, UN Niranjan, Animashree Anandkumar, and Cris Cecka. Tensor contractions with extended blas kernels on cpu and gpu. In *High Performance Computing (HiPC), 2016 IEEE 23rd International Conference on*, pages 193–202. IEEE, 2016.

[18] Kamyar Azizzadenesheli, Alessandro Lazaric, and Animashree Anandkumar. Reinforcement learning of pomdps using spectral methods. In *29th Annual Conference on Learning Theory*, pages 193–256, 2016.

[19] Alekh Agarwal, Animashree Anandkumar, Prateek Jain, and Praneeth Netrapalli. Learning sparsely used overcomplete dictionaries via alternating minimization. *SIAM Journal on Optimization*, 26(4):2775–2799, 2016.

[20] Animashree Anandkumar and Rong Ge. Efficient approaches for escaping higher order saddle points in non-convex optimization. In *29th Annual Conference on Learning Theory*, pages 81–102, 2016.

[21] Hanie Sedghi and Anima Anandkumar. Training input-output recurrent neural networks through spectral methods. *arXiv preprint arXiv:1603.00954*, 2016.

[22] Furong Huang, Ioakeim Perros, Robert Chen, Jimeng Sun, and Anima Anandkumar. Scalable latent tree model and its application to health analytics. In *Proc. of NIPS workshop on health analytics*, 2016.

[23] Anima Anandkumar, Prateek Jain, Yang Shi, and Uma Naresh Niranjan. Tensor vs. matrix methods: Robust tensor decomposition under block sparse perturbations. In *Artificial Intelligence and Statistics*, pages 268–276, 2016.

[24] Majid Janzamin, Hanie Sedghi, and Anima Anandkumar. Beating the perils of non-convexity: Guaranteed training of neural networks using tensor methods. *arXiv preprint arXiv:1506.08473*, 2015.

[25] Furong Huang and Animashree Anandkumar. Convolutional dictionary learning through tensor factorization. In *Proc. of NIPS workshop on Feature Extraction: Modern Questions and Challenges*, pages 116–129, 2015.

[26] Majid Janzamin, Hanie Sedghi, UN Niranjan, and Animashree Anandkumar. Feast at play: Feature extraction using score function tensors. In *Feature Extraction: Modern Questions and Challenges*, pages 130–144, 2015.

[27] Hanie Sedghi, Majid Janzamin, and Anima Anandkumar. Provable tensor methods for learning mixtures of generalized linear models. In *Artificial Intelligence and Statistics*, pages 1223–1231, 2016.

[28] Furong Huang, Animashree Anandkumar, Christian Borgs, Jennifer Chayes, Ernest Fraenkel, Michael Hawrylycz, Ed Lein, Alessandro Ingrosso, and Srinivas Turaga. Discovering neuronal cell types and their gene expression profiles using a spatial point process mixture model. In *Proc. of NIPS workshop*, 2016.

[29] Yining Wang, Hsiao-Yu Tung, Alexander J Smola, and Anima Anandkumar. Fast and guaranteed tensor decomposition via sketching. In *Advances in Neural Information Processing Systems*, pages 991–999, 2015.

[30] Forough Arabshahi, Furong Huang, Animashree Anandkumar, Carter T Butts, and Sean M Fitzhugh. Are you going to the party: Depends, who else is coming?:[learning hidden group dynamics via conditional latent tree models]. In *Data Mining (ICDM), 2015 IEEE International Conference on*, pages 697–702. IEEE, 2015.

[31] Tejaswi Nimmagadda and Anima Anandkumar. Multi-object classification and unsupervised scene understanding using deep learning features and latent tree probabilistic models. In *Proc. of CVPR workshop*, 2015.

[32] Furong Huang, UN Niranjan, Mohammad Umar Hakeem, and Animashree Anandkumar. Online tensor methods for learning latent variable models. *The Journal of Machine Learning Research*, 16(1):2797–2835, 2015.

[33] Animashree Anandkumar, Daniel Hsu, Majid Janzamin, and Sham Kakade. When are overcomplete topic models identifiable? uniqueness of tensor tucker decompositions with structured sparsity. *Journal of Machine Learning Research*, 16:2643–2694, 2015.

[34] Animashree Anandkumar, Rong Ge, and Majid Janzamin. Learning overcomplete latent variable models through tensor methods. In *Conference on Learning Theory*, pages 36–112, 2015.

[35] Praneeth Netrapalli, UN Niranjan, Sujay Sanghavi, Animashree Anandkumar, and Prateek Jain. Non-convex robust pca. In *Advances in Neural Information Processing Systems*, pages 1107–1115, 2014.

[36] Hanie Sedghi, Anima Anandkumar, and Edmond Jonckheere. Multi-step stochastic admm in high dimensions: Applications to sparse optimization and matrix decomposition. In *Advances in neural information processing systems*, pages 2771–2779, 2014.

[37] Furong Huang, Sergiy Matusevych, Anima Anandkumar, Nikos Karampatziakis, and Paul Mineiro. Distributed latent dirichlet allocation via tensor factorization. In *NIPS Optimization Workshop*, volume 1, 2014.

[38] Le Song, Animashree Anandkumar, Bo Dai, and Bo Xie. Nonparametric estimation of multi-view latent variable models. In *International Conference on Machine Learning*, pages 640–648, 2014.

[39] Alekh Agarwal, Animashree Anandkumar, Prateek Jain, Praneeth Netrapalli, and Rashish Tandon. Learning sparsely used overcomplete dictionaries. In *Conference on Learning Theory*, pages 123–137, 2014.

[40] Majid Janzamin and Animashree Anandkumar. High-dimensional covariance decomposition into sparse markov and independence models. *Journal of Machine Learning Research*, 15:1549–1591, 2014.

[41] A. Anandkumar, D. Hsu, M. Janzamin, and S. M. Kakade. When are Overcomplete Topic Models Identifiable? Uniqueness of Tensor Tucker Decompositions with Structured Sparsity. In *Neural Information Processing (NIPS)*, Dec. 2013.

[42] F. Huang and A. Anandkumar. Fast, Concurrent and Distributed Load Balancing under Switching Costs and Imperfect Observations. In *Proc. of IEEE INFOCOM*, Apr. 2013.

[43] A. Anandkumar, R. Ge, D. Hsu, and S. M. Kakade. A Tensor Spectral Approach to Learning Mixed Membership Community Models. In *Conference on Learning Theory (COLT)*, June 2013.

[44] A. Anandkumar, D. Hsu, A. Javanmard, and S. M. Kakade. Learning Bayesian Networks with Latent Variables. In *Proc. of Intl. Conf. on Machine Learning*, June 2013.

[45] A. Anandkumar and R. Valluvan. Learning Loopy Graphical Models with Latent Variables: Efficient Methods and Guarantees. In *Proc. of Neural Information Processing (NIPS)*, Dec. 2012.

[46] A. Anandkumar, D. P. Foster, D. Hsu, S. M. Kakade, and Y. K. Liu. A Spectral Algorithm for Latent Dirichlet Allocation. In *Proc. of Neural Information Processing (NIPS)*, Dec. 2012.

[47] A. Anandkumar, D. Hsu, F. Huang, and S.M. Kakade. Learning Mixtures of Tree Graphical Models. In *Proc. of Neural Information Processing (NIPS)*, Dec. 2012.

[48] M. Janzamin and A. Anandkumar. High-Dimensional Covariance Decomposition into Sparse Markov and Independence Domains. In *Proc. of International Conf. on Machine Learning*, June 2012.

[49] A. Anandkumar, D. Hsu, and S.M. Kakade. A Method of Moments for Mixture Models and Hidden Markov Models. In *Proc. of Conf. on Learning Theory*, June 2012.

[50] Animashree Anandkumar, Rong Ge, Daniel Hsu, Sham M Kakade, and Matus Telgarsky. Tensor decompositions for learning latent variable models. *The Journal of Machine Learning Research*, 15(1):2773–2832, 2014.

[51] A. Anandkumar, D. P. Foster, D. Hsu, S. M. Kakade, and Y. K. Liu. Two SVDs Suffice: Spectral Decompositions for Probabilistic Topic Modeling and Latent Dirichlet Allocation. *Special issue of Algorithmica on New Theoretical Challenges in Machine Learning*, July 2013.

[52] A. Anandkumar and R. Valluvan. Learning Loopy Graphical Models with Latent Variables: Efficient Methods and Guarantees. *Annals of Statistics*, 41(2):401–435, 2013.

[53] A. Anandkumar, V. Y. F. Tan, F. Huang, and A. S. Willsky. High-dimensional structure learning of Ising models: local separation criterion. *The Annals of Statistics*, 40(3):1346–1375, 2012.

[54] A. Anandkumar, V. Y. F. Tan, F. Huang, and A. S. Willsky. High-Dimensional Gaussian Graphical Model Selection: Walk-Summability and Local Separation Criterion. *J. Machine Learning Research*, 13:2293–2337, Aug. 2012.

[55] A. Anandkumar, A. Hassidim, and J. Kelner. Topology discovery of sparse random graphs with few participants. *J. of Random Structures and Algorithms*, 43, June 2013.

[56] Y. Liu, V. Chandrasekaran, A. Anandkumar, and A. Willsky. Feedback Message Passing for Inference in Gaussian Graphical Models. *IEEE Tran. on Signal Processing*, 60(8):4135–4150, Aug. 2012.

[57] A. Anandkumar, V. Y. F. Tan, and A. S. Willsky. High-Dimensional Graphical Model Selection: Tractable Graph Families and Necessary Conditions. In *Proc. of Neural Information Processing (NIPS)*, Dec. 2011. *Oral Presentation, AR 1%.*

[58] A. Anandkumar, K. Chaudhuri, D. Hsu, S.M. Kakade, L. Song, and T. Zhang. Spectral Methods for Learning Multivariate Latent Tree Structure. In *Proc. of Neural Information Processing (NIPS)*, Dec. 2011.

[59] Animashree Anandkumar, Avinatan Hassidim, and Jonathan Kelner. Topology discovery of sparse random graphs with few participants. In *ACM SIGMETRICS* **Winner of Best Paper Award**, volume 39, pages 253–264, 2011.

[60] M. A. Khajehnejad, J. Yoo, A. Anandkumar, and B. Hassibi. Summary Based Structures with Improved Sublinear Recovery for Compressed Sensing. In *Proc. of IEEE ISIT*, July 2011.

[61] M.J. Choi, V.Y.F. Tan, A. Anandkumar, and A. Willsky. Learning latent tree graphical models. *J. of Machine Learning Research*, 12:1771–1812, May 2011.

[62] V.Y.F. Tan, A. Anandkumar, and A. Willsky. Learning Markov forest models: analysis of error rates. *J. of Machine Learning Research*, 12:1617–1653, May 2011.

[63] V.Y.F. Tan, A. Anandkumar, and A. Willsky. A large-deviation analysis for the maximum likelihood learning of tree structures. *IEEE Tran. on Information Theory*, 57(3):1714–1735, March 2011.

[64] V.Y.F. Tan, A. Anandkumar, and A. Willsky. Learning Gaussian tree models: analysis of error exponents and extremal structures. *IEEE Tran. on Signal Processing*, 58(5):2701–2714, May 2010.

PhD Theses Advised

[65] Furong Huang. *Discovery of latent factors in high-dimensional data using tensor methods.* PhD thesis, 2016.

[66] Majid Janzamin. *Non-convex Optimization in Machine Learning: Provable Guarantees Using Tensor Methods*. PhD thesis, 2016.

[67] Hanie Sedghi. *Stochastic Optimization in High Dimension*. PhD thesis, University of Southern California, 2015.

Selected Publications on PhD Research

[68] Paul Balister, Béla Bollobás, Animashree Anandkumar, and Alan Willsky. Energy-latency tradeoff for in-network function computation in random networks. In *INFOCOM, 2011 Proceedings IEEE*, pages 1575–1583. IEEE, 2011.

[69] A. Anandkumar, N. Michael, A.K. Tang, and A. Swami. Distributed algorithms for learning and cognitive medium access with logarithmic regret. *Selected Areas in Communications, IEEE Journal on*, 29(4):731–745, 2011. **Best Readings on Cognitive Radio by IEEE Comsoc society**.

[70] A. Anandkumar, J.E. Yukich, L. Tong, and A. Swami. Energy Scaling Laws for Distributed Inference in Random Networks. *IEEE J. Selec. Area Comm.*, 27(7):1203–1217, Sept. 2009.

[71] A. Anandkumar, L. Tong, and A. Swami. Detection of Gauss-Markov Random Fields with Nearest-neighbor Dependency. *IEEE Tran. Information Theory*, 55(2):816–827, Feb. 2009.

[72] A. Anandkumar, L. Tong, and A. Swami. Distributed Estimation Via Random Access. *IEEE Tran. Information Theory*, 54(7):3175–3181, July 2008.

[73] A. Anandkumar and L. Tong. Type-Based Random Access for Distributed Detection over Multiaccess Fading Channels. *IEEE Trans. Signal Proc.*, 55(10):5032–5043, Oct. 2007 **Winner of Young Author Best Paper Award**.

[74] A. Anandkumar and L. Tong. A Large Deviation Analysis of Detection over Multi-Access Channels with Random Number of Sensors. In *Proc. of ICASSP'06*, volume IV, pages 1097–1101, Toulouse, France, May 2006 **Winner of Student Paper Award**.

Teaching

Special topics in ML (2013,2015), Signals & Systems (2012-15), Large-scale ML (2014), Stat. Learning Theory (2014), Estimation Theory (2011-15), Random Processes (2010-11).

Service and Community Outreach

Democratizing AI through cloud credit program at AWS and through sponsorships of ML conferences, hackathons and student-run tech events. Spearheaded the partnership between Amazon Web Services and Caltech through $2 million cloud credit program.

Board of directors at "Behind her eyes", a non-profit focusing on the impact of immersive VR on underrepresented minorities.

Scientific advisory committee for the Center for Autonomous Systems and Technologies (CAST) at Caltech.

Judge for MIT Technology Review 35 under 35.

PC for ICML 2012-18, NIPS 2014-18, AISTATS 2016, UAI 2013-14, SIGMETRICS 2014-16. Action Editor for Journal of Machine Learning Research. Assoc. Editor for IEEE Tran. on Signal Processing (2012-2014).

In the News

Story of Anima Anandkumar, the machine learning guru powering Amazon AI. Yourstory. (Link)

In the Research Spotlight: Anima Anandkumar. Amazon AWS AI Blog. (Link)

Teaching Machines How to Learn: An Interview with Animashree Anandkumar, Caltech, 2017. (Link)

Flying ambulances, space robots and the ethics of artificial intelligence. KPCC. (Link)

Global Governance of AI Roundtable - World Government Summit 2018. The Future Society. (Link)

Robots Get Human-like Brains With Machine Learning and A.I. PBS reporter David Nazar. (Link)

China and the US are bracing for an AI showdownin the cloud. MIT Technology Review. (Link)

The Amazon Empire, an interview in Chinese by MIT Technology Review (Link)

Future customer needs in focus at Shaastra 2018. Indian Express.(Link)

Article on non-convex methods in Huffington post and Forbes. (Link 1), (Link 2).

Quora session on machine learning. (Link)

Interview with Amplify partner David Beyer. (Link)

Announcement of AFOSR YIP, Feb. 2015. (Link)

Announcement of the Sloan fellowship, Feb. 2014. (Link)

Announcement of Microsoft faculty award, July 2013. (Link 1), (Link 2). (Interview).

Talk Videos and Podcasts

Deep learning demystified. Experian 2018. (Link)
Machine learning at scale: deep, distributed and multi-dimensional. MLConf, Nov 2017. (Link)
Tensors for Large-scale Topic Modeling and Deep Learning. AWS re:Invent 2017. (Link)
Deep learning that's easy to implement and easy to scale. O'Reilly podcast (Link)
Tensors for large-scale ML. Strata 2015. (Link)
O'Reilly Data Show Podcast: tensor decomposition techniques for machine learning. (Link)