Professor Animashree Anandkumar
Bren Professor of Computing and Mathematical Sciences
California Institute of Technology
USA

Professor Frances H. Arnold
1200 E. California Blvd.
Pasadena, CA 91125
(626) 395-4162
frances@cheme.caltech.edu
web: http://fhalab.caltech.edu

Caltech Division of Chemistry
and Chemical Engineering MC 210-41

December 5, 2019

Dear Members of the Awards Committee,

I am writing this letter to express my strongest support for the nomination of **Prof. Anima Anandkumar** for the 2020 "Albert Einstein" World Award of Science. Prof. Anandkumar is currently a Bren professor at Caltech in the computing and mathematical sciences department. She also is a director of machine-learning research at NVIDIA.

Prof. Anandkumar's foundational contribution to the field of artificial intelligence is the class of tensor algorithms first proposed in her seminal 2014 paper. Tensors are central for effectively processing multidimensional and multimodal data, and for achieving massive parallelism in large-scale AI applications. The tensor algorithms proposed by Prof. Anandkumar are fundamentally, a new class of AI algorithms. They are the first theoretically guaranteed methods for solving a broad range of problems in unsupervised, supervised, and reinforcement learning. Building on these strong foundations, Prof. Anandkumar has also found great success in making these algorithms practical. She productionized these tensor algorithms at Amazon Web Services, making them the most scalable algorithms for document categorization and probabilistic modeling available on the cloud.

I strongly believe that the machine-learning techniques developed by Prof. Anandkumar will revolutionize protein engineering. Protein engineering is a notoriously challenging task, as the space of protein sequences is too large to be searched exhaustively. While generic deep learning has shown some preliminary gains, I believe that tensor frameworks developed by Prof. Anandkumar will be far more effective in handling this complex high-dimensional domain. We have started an ambitious interdisciplinary collaboration with Prof. Anandkumar and multiple faculty in biology and chemistry to explore these research directions. I expect that this will to lead to transformative advances, resulting in reduction of our experimental costs by 1000x or more, and impacting a broad range of applications from sustainable chemistry and biofuels to biomedicine and biosensors.

In summary, Prof. Anandkumar belongs to a rare breed of researchers whose contributions span a broad spectrum: building novel foundations for AI, transforming the practice of AI, and also promoting fairness and democratization of AI. I consider her a superb candidate for this prestigious award.

Sincerely,

Frances Arnold

Frances Hamilton Arnold, *PhD*
Linus Pauling Professor of Chemical Engineering, Bioengineering and Biochemistry

FHA:cn

UNIVERSITY OF CALIFORNIA, BERKELEY

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO   SANTA BARBARA • SANTA CRUZ

DIVISION OF COMPUTER SCIENCE
DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

387 SODA HALL
BERKELEY, CALIFORNIA 94720-1776
bartlett@cs.berkeley.edu

December 10, 2019

2020 Albert Einstein World Award of Science Committee
Dear Award Committee

I am writing in support of Professor Anima Anandkumar's nomination for the 2020 "Albert Einstein" World Award of Science. Professor Anandkumar has made pioneering contributions in large scale machine learning, contributions that are innovative, have had an enormous impact on the research field, and are likely to have an enormous impact in the near future in the many application areas where machine learning has become a critical technology. Her contributions are in the development of *tensor methods*, which are computationally efficient estimation methods for complex statistical models. Let me give some background to explain the significance of her work. In using data to make effective decisions, there are three critical steps: deciding on a model for the phenomenon of interest, using data to estimate the parameters of that model, and using the estimated parameters to make decisions (whether they be hypothesis tests or other kinds of predictions). In models in which all of the relevant random variables can be observed, parameter estimation is typically very straightforward. Maximum likelihood estimation, for instance, involves the solution of an optimization problem that is often convex. However, when a model involves variables that are not observed, parameter estimation becomes computationally difficult. Models with hidden variables are ubiquitous, because they often allow an accurate but parsimonious representation for natural phenomena. For instance, in density estimation, mixtures of Gaussians (where the variable representing from which mixture component an individual datum was drawn) are a flexible and widely used model that is appropriate for vector data that occurs in multiple clusters. In information retrieval, topic models for documents have become popular; these hypothesize the presence of unobserved variables representing the mix of topics contained in a document. In speech recognition, hidden Markov models use hidden variables to model the state of the vocal tract throughout an utterance; these models have been very widely applied.

Prior to Prof Anandkumar's work, the state of the art for estimation of model parameters in problems involving hidden variables was the *expectation maximization* (EM) algorithm, a heuristic developed in the second half of the twentieth century for finding a local maximum of the likelihood. But like many local optimization procedures for nonconvex objectives, there is no guarantee that EM will lead to an effective estimate. In a series of papers with a variety of collaborators, Prof Anandkumar's work took a different approach: instead of the classical approach of maximizing likelihood in these problems, she developed methods for finding parameters that match the moments (expectations, variances, and higher moments) of the observed random variables. Such methods had been considered since the early twentieth century for simpler statistical problems but have been largely ignored, because their asymptotic variance is worse than that of maximum likelihood. But they have an enormous advantage in large scale problems, because they can often be computed efficiently. This means that an estimate with well-quantified properties can be obtained with a small amount of computation, whereas no such guarantees are known for heuristics like EM. Prof Anandkumar developed tensor methods of this kind for a variety of important statistical models, and introduced efficient algorithms for these methods. This advance has applications across the enormous and diverse range of scientific and technological areas where probabilistic modelling is used.

It is clear from her long list of awards and grants that Prof Anandkumar's work is already widely viewed as pioneering. I give her my strongest recommendation for this award.

Please let me know if you require any more information.

Sincerely

Peter L. Bartlett
Professor, Computer Science Division and Department of Statistics
Associate Director, Simons Institute for the Theory of Computing

Microsoft Research
641 Avenue of the Americas, 7th floor
New York, NY 10011
646-746-6264
schapire@microsoft.com

Robert Schapire
Partner Researcher

December 3, 2019

To Whom It May Concern:

I am pleased to write this strong letter of recommendation for Animashree Anandkumar who is being considered for the Albert Einstein World Award of Science. I have known Anima for many years having seen her regularly but infrequently at conferences, research visits, etc. We have never collaborated together on research, nor ever been colleagues at the same institution.

For many years, Anima's research focused on the study of latent-variable models. These are a broad class of statistical models used in countless applications to understand data of many different varieties, including speech signals, visual images, text data, social networks, and many more. Such models include hidden Markov models, mixtures of Gaussians, latent Dirichlet allocation models, and Bayesian networks, to name just a few. Although used ubiquitously in practice, algorithms for inferring such models typically could not be proved mathematically to be effective in a general sense -- that is, until the work of Anima and her collaborators. Over a period of several years, this team of researchers, with Anima acting as a core leader, developed a new and quickly-growing family of algorithms that they proved theoretically to be effective under general conditions. Their novel approach was based on a deep, elegant, unified and versatile set of methods centered on computing moments and tensor decompositions. Although mathematically sophisticated, their methods are efficient and potentially of long-term practical importance.

I regard this work as a breakthrough, truly inspiring, and very high impact. It also turned out to be an extremely fruitful direction for research, as evidenced by the numerous papers produced in this area by Anima and others (and generally published in the very best and most selective venues in the field). Anima's record in this area has been particularly stunning in terms of her steadfast determination to grow this idea into a well-developed subfield. Over the years, she methodically expanded the range of problems that can be provably handled using this approach, while at the same time remaining focused on practical considerations, including the development of usable software.

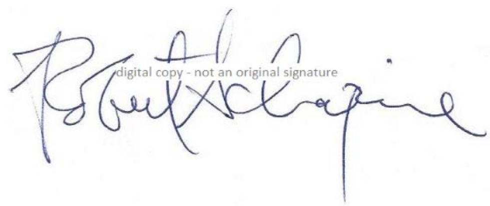
In recent years, Anima's interests have continued to expand in a remarkable way to encompass an amazing array of important machine-learning topics, including neural nets, time series, reinforcement learning, and game-playing. In many cases, she has shown how tensors and related methods, her area of particular expertise and leadership, can be applied, generally in a highly-principled way.

Microsoft

Anima is an extremely active and high-energy researcher with a passionate and unusually well-focused research program. She has also been tremendously successful in working with students as well as her many collaborators around the world. Her work has been recognized with prestigious fellowships, multiple best paper awards, etc.

In sum, Anima's research is important, timely, bold, high-impact, broad yet well-focused, and very original. I recommend her strongly.

Sincerely,



Robert Schapire

Resume of Anima Anandkumar
Albert Einstein Prize

Anima Anandkumar is a Bren professor at Caltech in the Computing and Mathematical Sciences (CMS) department. She is the youngest named chair professor at Caltech, the highest honor the university bestows on an individual faculty. She is also the director of machine-learning (ML) research at NVIDIA.

Prof. Anandkumar's primary contribution to the field of artificial intelligence (AI) is the class of tensor algorithms, first proposed in her seminal 2014 paper. Prof. Anandkumar developed an entirely new class of learning algorithms that incorporate tensors. She laid the theoretical foundation for guaranteed solutions to a broad range of problems in unsupervised, supervised, and reinforcement learning using tensor algorithms.

Historical Perspective: Prior to Prof. Anandkumar's seminal work, tensors were not even on the "radar" of ML researchers. They were not considered a viable tool for machine learning. Fast forward today, tensors are ubiquitously associated with machine learning. For instance, TensorFlow is the most popular framework for AI and Tensor Cores are the programmable cores in NVIDIA GPUs, which is the primary platform for AI. This is just the beginning of a tensor revolution; currently these frameworks support only a small set of tensor operations, and there is rapid progress in expanding them.

In 2012, Prof. Anandkumar started focusing on the problem of learning latent variable models. This is a form of unsupervised learning, and it is considered as one of the hardest problems since there is no supervision or labels available at the time of training. The goal is to design algorithms that can automatically extract hidden or latent factors from data.

There are mainly two classes of algorithms for unsupervised learning. One is expectation maximization (EM), an iterative local search approach that alternates between estimating model parameters and configuration of hidden variables. However, in high dimensions, EM is prone to failure and tends to be far from the optimal solution. Hence, EM is expensive and does not guarantee good solutions.

The other approach is the class of matrix methods which use efficient linear-algebraic routines such as spectral decomposition. A popular example is the principal component analysis (PCA), where the principal directions of the data covariance matrix is used to filter out noise. More generally, these matrix methods attempt to find the latent subspaces which can fit the observed data well.

In 2012, Prof. Anandkumar was beginning to understand the limits of such matrix methods. In PCA, only second order moments (covariance) of data are used. Prof. Anandkumar started thinking about using higher order moments for learning. She realized that tensors are a natural approach for expressing them. This led her to pose further questions: how much more information is present in these higher order moments? how can they help us learn about the underlying latent factors? Since the size of these moments grows exponentially, how can we still process them efficiently?

The answers to the above questions are not straightforward. In fact, there have been strong arguments against the use of tensors in learning and computational problems. But Prof. Anandkumar recognized that many things were now on her side: we now have access to enormous amounts of data. Parallel computing, which is at the heart of modern computing revolution, is ideal for scaling up tensor operations. Hence, she realized the untapped potential of tensors for solving AI problems.

Prof. Anandkumar established the first theoretical guarantees for unsupervised learning of a broad class of latent variable models using tensor methods. She showed that these methods enjoy low computational and sample complexity, meaning that they require a limited amount of data and computation for learning. In her theory, she characterized conditions under which these methods succeed in finding the globally optimal solution, and they turn out to be mild and reasonable for most AI problems. Hence, these are the first methods to guarantee unsupervised learning of latent variable models.

Impact on Applications: Prof. Anandkumar helped build these tensor algorithms into Amazon SageMaker and Amazon Comprehend frameworks, the most scalable publicly-available software for document categorization and probabilistic modeling. It offers at least $10x$ speed improvement over state-of-art. These cloud services are used by tens of thousands of customers. She also helped develop an open source framework TensorLy. It integrates tensor-algebraic operations with popular deep learning frameworks. At NVIDIA, the new CuTensor library provides computational primitives for tensor operations on GPUs. They lay the foundation for parallelizing tensor algorithms at scale. Thus, Prof. Anandkumar has had a tremendous impact on AI applications in a short period of time.

Other Contributions: Prof. Anandkumar has invented a broad array of techniques for establishing theoretical guarantees for learning problems involving non-convex optimization. For example, she has shown that it is possible to have “free lunch”: algorithms that require significantly less communication for distribution AI training, but have strong theoretical guarantees for non-convex optimization. Recently, she has developed a novel guaranteed algorithm for competitive optimization, which involves multiple players who want to maximize their own objectives. The next leaps in machine learning will be critically dependent on efficient algorithms for competitive optimization. Thus, she is one of the rare researchers who is able to span between theoretical foundation and practical impact.

The Future: AI is a trinity of data, algorithms, and compute infrastructure. Tensors will continue to play a central role in strengthening these ties among the three facets. As we collect richer multi-dimensional and multi-modal data, tensors are the natural data structures to store and manipulate them. Tensor algorithms that manipulate such data tensors will harness the most useful information effectively. On the computational front, with the end of Moore’s law, single threaded computing is no longer seeing rapid improvements in performance. This means having a strong reliance on parallel computing to speed up and scale up our algorithms. Tensor computations will enable higher parallelism due to their multi-dimensional nature. Thus, tensors will continue play a pivotal role in the ongoing AI revolution.

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

Artificial intelligence and machine learning with a focus on deep learning, probabilistic models, optimization, and tensor methods.

Current Appointments

Director of Machine Learning Research, NVIDIA, Santa Clara, CA. *Since 2018*

Bren Professor, CMS Dept., California Institute of Technology, Pasadena, CA. *Since 2017*

Previous Appointments

Principal Scientist, Amazon AI, Amazon Web Services (AWS), Palo Alto, CA. *2016 - 2018*

Associate Professor, ICS Dept., University of California, Irvine, CA. *2016 - 2017*

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

Visiting Researcher at Microsoft Research New England, Cambridge, MA. *2012 - 2012*

Post-doctoral Associate at Massachusetts Institute of Technology, Cambridge, MA. *2009 - 2010*

Education

Doctor of Philosophy in Electrical and Computer Engineering, Cornell University. *2009*

Bachelor of Technology in Electrical Engineering, Indian Institute of Technology Madras. *2004*

Awards and Honors

1. **NYTimes GoodTech Award**
2. **Bren Named Chair Professorship at Caltech, 2017.**
3. **Expert network of World Economic Forum, 2017.**
4. **Google Faculty Research Award 2015.**
5. **AFOSR Young Investigator Award (YIP) 2015.**
6. **Alfred P. Sloan Research Fellowship 2014.**
7. **Microsoft Faculty Fellowship 2013.**
8. **ARO Young Investigator Award (YIP) 2013.**
9. **NSF CAREER Award 2013.**
10. **ACM SIGMETRICS 2011 Best Paper Award.**
11. **Best Thesis Award 2009** by ACM SIGMETRICS Society.

Teaching

Foundations of Machine Learning (2018-), 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).

Scientific Leadership

Expert network of the World Economic Forum.

Chaired the committee on mapping AI progress at Global Governance of AI Roundtable (GGAR).

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

Co-director of Decision, Optimization and Learning (Dolcit), Caltech.

Co-director of AI4Science initiative, Caltech.

Advisory Council for NORC, University of Chicago and ECE Department, Cornell University.

Judge for MIT Technology Review 35 under 35 and Forbes AI50.

PC for ICML 2012-19, NIPS 2014-18, AISTATS 2016, UAI 2013-14, SIGMETRICS 2014-16.

Action Editor for Journal of Machine Learning Research. Assoc. Editor for Harvard Data Science Review, Assoc. Editor for IEEE Tran. on Signal Processing (2012-2014).

Workshop Chair for ICML 2017. Organizer of several workshops at ICML, NIPS, Fields institute, Dagstuhl.

Democratizing AI through NVIDIA inception program, cloud credit program at AWS and through sponsorships of ML conferences, hackathons and student-run tech events.

Board of directors at GoBeyondResumes, a non-profit focusing on the goal to help create a true meritocracy amongst the 50,000 students that will be graduating from US colleges with Computer Science degrees in 2019 and to help companies recruit based on skills, not resume keywords

Research Support (Current and Past)

1. Bren named professorship at Caltech.
2. DARPA Physics of AI, DARPA Purposeful Learning (Purple).
3. Research support from Raytheon and BMW.
4. Microsoft, Google and Adobe faculty fellowships.
5. Alfred. P. Sloan fellowship.
6. AFOSR and ARO Young Investigator Awards
7. NSF Career, NSF BigData, NSF CCF.

Invited Talks, Podcasts and Media

Keynotes and Named Lectures

Top 50 Innovators, Royal Society, London, 2019.

UW Boeing Distinguished Lecture, 2019.

SIAM CSE plenary talk, 2019.

Michigan Institute of Data Science (MIDAS) Distinguished lecture, 2019.

Techfest, IIT Bombay, 2019.

Simons Institute Open Lecture, UC Berkeley, 2018.

TEDx, Indiana University, 2018.

Geekpark Rebuild, Chengdu, 2018.

Digital Innovation Forum, Taipei, 2018.

QS Caltech Innovator Series, NYC, 2018.

ACM India Joint Intl. Conf. on Data Science and Management of Data (CoDS-COMAD), 2018.

Global Mulan Forum and China Business Mulan Annual Meeting, Beijing, 2018.

EmTech China, MIT Technology Review, Beijing, 2018.

Data Science Annual Conference (DSCO), UCSF, 2017.

Information Theory and Applications, San Deigo, 2017.

Indaba Deep Learning, South Africa, 2017.

Podcasts

AI Podcast: Tensor Operations for Machine Learning with Anima Anandkumar. (Link)

Practical AI: Growing up to become a world-class AI expert. (Link)

Deep learning demystified Podcast. Experian 2018. (Link)

Deep learning that's easy to implement and easy to scale. O'Reilly podcast (Link)

O'Reilly Data Show Podcast: tensor decomposition techniques for machine learning. (Link)

In the News

Caltech Celebrates Newest Cohort of Named Professors. (Link)

NVIDIA Opening Core AI and ML Research Lab in Santa Clara - NVIDIA Developer News Center. (Link)

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

List of 10 Significant Publications

Anima Anandkumar

Albert Einstein Award

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] Florian Schäfer and Anima Anandkumar. Competitive gradient descent. In *Proc. of NeurIPS*, 2019.

[6] Jeremy Bernstein, Jiawei Zhao, Kamyar Azizzadenesheli, and Anima Anandkumar. signSGD with Majority Vote is Communication Efficient And Byzantine Fault Tolerant. In *Proc. of ICLR*, 2019.

[7] 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.

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

[9] 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.

[10] 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.

Anima Anandkumar

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

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

Selected List of Publications

[1] Florian Schäfer and Anima Anandkumar. Competitive gradient descent. In *Proc. of NeurIPS*, 2019.

[2] Furong Huang, Ioakeim Perros, Robert Chen, Jimeng Sun, and Anima Anandkumar. Guaranteed scalable latent tree models. In *Proc. of uncertainty in AI (UAI)*, 2019.

[3] Chris Swierczewski, Sravan Bodapati, Anurag Beniwal, David Leen, and Animashree Anandkumar. Large scale cloud deployment of spectral topic modeling. In *KDD 2019, ParLearning Workshop*, 2019.

[4] Milan Cvitkovic, Badal Singh, and Anima Anandkumar. Open vocabulary learning on source code with a graph-structured cache. In *Proc. of ICML*, 2019.

[5] Yang Shi and Anima Anandkumar. Multi-dimensional tensor sketch. In *Proc. of KDD workshop*, 2019.

[6] Guanya Shi, Xichen Shi, Michael O'Connell, Rose Yu, Kamyar Azizzadenesheli, Animashree Anandkumar, Yisong Yue, and Soon-Jo Chung. Neural lander: Stable drone landing control using learned dynamics. In *Proc. of ICRA*, 2019.

[7] Jeremy Bernstein, Jiawei Zhao, Kamyar Azizzadenesheli, and Anima Anandkumar. signSGD with Majority Vote is Communication Efficient And Byzantine Fault Tolerant. In *Proc. of ICLR*, 2019.

[8] Kamyar Azizzadenesheli, Anqi Liu, Fanny Yang, and Animashree Anandkumar. Regularized learning for domain adaptation under label shifts. In *Proc. of ICLR*, 2019.

[9] Peiyun Hu, Zachary C Lipton, Animashree Anandkumar, and Deva Ramanan. Active learning with partial feedback. In *Proc. of ICLR*, 2019.

[10] Yang Shi, Tommaso Furlanello, Sheng Zha, and Animashree Anandkumar. Question type guided attention in visual question answering. In *Proc. of ECCV*, 2018.

[11] Jean Kossaifi, Yannis Panagakis, Anima Anandkumar, and Maja Pantic. Tensorly: Tensor learning in python. *Journal of Machine Learning Research*, 20(26):1–6, 2019.

[12] Milan Cvitkovic, Badal Singh, and Anima Anandkumar. Deep learning on code with an unbounded vocabulary. In *Proc. of FLoC 2018, Machine Learning for Programming Workshop*, 2018.

[13] Ben Athiwaratkun, Andrew Gordon Wilson, and Anima Anandkumar. Probabilistic fasttext for multi-sense word embeddings. In *Proc. of ACL*, 2018.

[14] Michael Tschannen, Aran Khanna, and Anima Anandkumar. Strassennets: Deep learning with a multiplication budget. In *Proc. of ICML*, 2018.

[15] Tommaso Furlanello, Zachary C Lipton, AI Amazon, Laurent Itti, and Anima Anandkumar. Born again neural networks. In *Proc. of ICML*, 2018.

[16] signsgd: compressed optimisation for non-convex problems.

[17] 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.

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

[19] 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.

[20] 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.

[21] 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.

[22] 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.

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

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

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

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

[27] 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.

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

[29] 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.

[30] 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.

[31] 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.

[32] 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.

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

[34] 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.

[35] 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.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] 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.

[41] 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.

[42] 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.

[43] 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.

[44] 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.

[45] 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.

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

[47] 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.

[48] 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.

[49] 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.

[50] 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.

[51] 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.

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

[53] 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.

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

[55] 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.

[56] 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.

[57] 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.

[58] 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.

[59] 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.

[60] 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.

[61] 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.

[62] 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.

[63] 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.

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

[65] 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.

[66] 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.

[67] 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.

[68] 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.

[69] 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%.*

[70] 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.

[71] 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.

[72] 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.

[73] 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.

[74] 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.

[75] 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.

[76] 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.

[77] A. Anandkumar, J.E. Yukich, and A. Willsky. Scaling Laws for Random Spatial Graphical Models. In *Proc. of IEEE ISIT*, Austin, USA, June 2010.

[78] Y. Liu, V. Chandrasekaran, A. Anandkumar, and A. Willsky. Feedback Message Passing for Inference in Gaussian Graphical Models. In *Proc. of IEEE ISIT*, Austin, USA, June 2010.

[79] V. Tan, A. Anandkumar, and A. Willsky. Error Exponents for Composite Hypothesis Testing of Markov Forest Distributions. In *Proc. of IEEE ISIT*, Austin, USA, June 2010.

PhD Theses Advised

[80] Kamyar Azzizadenesheli. *Reinforcement Learning in Partially Observed and Structured Environments.* PhD thesis, 2019.

[81] Yang Shi. *Efficient Tensor Operations via Sketching and Parallel Computation.* PhD thesis, 2019.

[82] Forough Arabshahi. *Learning Latent Hierarchical Structures via Probabilistic Models and Deep Learning.* PhD thesis, 2018.

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

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

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

Selected Publications from PhD Research

[86] 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.

[87] 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**.

[88] 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.

[89] 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.

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

[91] 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**.

[92] 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**.

[93] A. Anandkumar, C. Bisdikian, and D. Agrawal. Tracking in a Spaghetti Bowl: Monitoring Transactions Using Footprints. In *Proc. of ACM SIGMETRICS*, Annapolis, Maryland, USA, June 2008.

[94] A. Anandkumar, L. Tong, A. Swami, and A. Ephremides. Minimum Cost Data Aggregation with Localized Processing for Statistical Inference. In *Proc. of INFOCOM*, pages 780–788, Phoenix, USA, April 2008.