

# Krishnamurthy (Dj) Dvijotham

## PERSONAL DATA

---

Experienced researcher with expertise in AI, Optimization and Control theory

PHONE: +1 2066179986 EMAIL: [dvijcse@gmail.com](mailto:dvijcse@gmail.com) WEBPAGE: [dvij.github.io](http://dvij.github.io)

## WORK EXPERIENCE

---

*August 7 2017 - Present*

Staff Research Scientist at [DeepMind](#), 6 Pancras Square, London, UK  
*Research in Robust and Verifiable AI, Integrating AI into Google Products*

Lead research effort involving several research scientists, interns and engineers to develop robust and verifiable AI algorithms that can be safely deployed in adversarial and safety-critical environments. Line manager to a team of 4 research scientists. Developed and integrated AI into Google products.

AUG 22 2016-JULY 7 2017

Research Scientist at [PACIFIC NORTHWEST NATIONAL LABORATORY](#)

902 Battelle, Blvd, Richland, WA, USA (on H1B)

*Research in Control & Optimization of Power Grids*

Optimization and control of power grids under high uncertainty arising from fluctuating renewable generation sources like solar and wind power. Wrote successful proposals for internal and external (DoE, NSF) funding and mentored interns and junior researchers.

MARCH 24 2016 - JULY 31 2016

Fellow at [CENTER FOR MATHEMATICS OF INFORMATION, CALTECH](#)

1200 E California Blvd, Pasadena, CA, USA (on H1B)

*Research in Control, Optimization and Game Theory*

Developed algorithms for robust optimization, control and game theoretic analysis of large scale infrastructure networks like power and gas grids. Mentored graduate students and organized group meetings/seminars.

AUGUST 1 2014 - MARCH 23 2016

Fellow at [CENTER FOR MATHEMATICS OF INFORMATION, CALTECH](#)

1200 E California Blvd, Pasadena, CA, USA (on F1 OPT)

*Research in Control, Optimization and Game Theory*

Developed algorithms for robust optimization, control and game theoretic analysis of large scale infrastructure networks like power and gas grids. Mentored graduate students and organized group meetings/seminars.

## EDUCATION

---

MARCH 2014 PhD in COMPUTER SCIENCE AND ENGINEERING, **University of Washington, Seattle**

Thesis: "Automating Stochastic Control"

Advisor: Prof. Emanuel TODOROV, Prof. Maryam FAZEL

GPA: 3.9/4.0

JULY 2010 MS in COMPUTER SCIENCE AND ENGINEERING, **University of Washington, Seattle**

Advisor: Prof. Emanuel TODOROV

GPA: 3.9/4.0

MAY 2008 Bachelor of Technology in COMPUTER SCIENCE AND ENGINEERING, **IIT Bombay**

Advisor: Prof. Soumen CHAKRABARTI

GPA: 8.7/10.0

## RECENT ACHIEVEMENTS

---

- JAN 2020 - PRESENT Line Manager to 4 research scientists in the robust and verified AI team, Mentor to several research scientists and engineers and Co-lead of projects within the robust and verified AI team
- OCT 2020 Promoted to Staff Research Scientist
- NOV 2019 Collaboration with Google Play Store featured on [DeepMind blog](#)
- JAN - DEC 2019 Co-lead for Robust & Verified AI Push at DeepMind, coordinating a team of 10 research scientists, engineers and interns working on robust artificial intelligence  
Research featured on [DeepMind blog](#)
- JULY 2019 Launched algorithm to filter fraudulent banking apps on Google Play App Store  
Doubled precision of previous approach
- DEC. 2018 Best paper award, Neural Information Processing Systems Security Workshop (NeurIPS SecML 2018)  
*Research topic: Verification of AI*
- OCT. 2018 Promoted to Senior Research Scientist at DeepMind with “Superb” performance review  
Rating awarded only to highly select employees at the company
- AUG. 2018 Best paper award,  
Conference on Uncertainty in Artificial Intelligence (UAI 2018)  
*Research topic: Verification of AI*
- JUL. 2018 Applied research at DeepMind featured on [DeepMind blog](#)  
*Research topic: AI for battery management on Android phones*

## AWARDS

---

- JUL. 2016 Best paper award,  
Conference on Principles and Practice of Constraint Programming (CP 2016)  
*Research topic: AI for sustainable electric power grids*
- JUL. 2014 Best student paper award  
Conference on Uncertainty in Artificial Intelligence (UAI 2014)  
*Research topic: Convex optimization formulations of risk-averse machine learning*
- JUL. 2014 Best student paper award, ACM GREENMETRICS workshop  
*Research topic: Analysis of strategic behavior in electricity markets*
- SEP. 2008 Best student paper award  
European Conference on Machine Learning (ECML 2008)  
*Research topic: Approximate inference in probabilistic graphical models*

## PEOPLE MANAGER

---

- DeepMind Rudy Bunel (research scientist, DeepMind)  
*Start date: October 21, 2019*
- DeepMind Johannes Welbl (research scientist, DeepMind)  
*Start date: February 21, 2020*
- DeepMind Leonard Berrada (research scientist, DeepMind)  
*Start date: March 2, 2020*
- DeepMind Sumanth Dathathri (research scientist, DeepMind)  
*Start date: March 23, 2020*

## MENTORING

---

DeepMind	Jamie Hayes (research intern, PhD student at University College London) <i>Research topic:</i> Learning robust watermarking of images and audio
DeepMind	Lily Weng (research intern, PhD student at MIT) <i>Research topic:</i> Adversarial evaluation of continuous control policies
DeepMind	Sumanth Dathathri (research intern, PhD student at Caltech) <i>Research topic:</i> Verification of temporal logic specifications for neural networks
DeepMind	Rudy Bunel (research intern, PhD student at Oxford) <i>Research topic:</i> Scalable algorithms for neural network verification
Deepmind	Chenglong Wang (research intern, PhD student at UW) <i>Research topic:</i> Verifying output-length specifications for sequence to sequence models
PNNL	Benjamin Rapone (research intern, PhD student at WSU) <i>Research topic:</i> Computational topology for robust feasibility analysis of power grids
PNNL	Haoxiang Yang (research intern, PhD student at Northwestern) <i>Research topic:</i> Robust optimization for electric power systems
PNNL	Thayagarajan Ramachandran (postdoc) <i>Research topic:</i> Market dynamics in electric power grids

## PEER REVIEW SERVICES

---

Area Chair	International Conference on Learning Representations (ICLR) 2020, 2021
Program Committee	ICML 2019, IJCAI 2019, ICLR 2019, NeurIPS 2018 (top 30% reviewer)
Reviewer	IEEE Controls and Decision Conference, American Controls Conference, IEEE Power System Computation Conference (PSCC), Journal of Machine Learning Research (JMLR), IEEE Transactions on Automatic Control, IEEE Transactions on Control of Network Systems, IEEE Transactions on Power Systems

## INVITED TALKS

---

MARCH 2019	AAAI Spring Symposium 2019, Workshop on Verification of Neural Networks ( <a href="#">VNN-19</a> ), Stanford, CA, USA (Keynote Speaker)
MAY 2018	Workshop on Mathematical and Computational Challenges in real-time decision making, Simons Institute for the Theory of Computing, UC Berkeley (Invited speaker)
APRIL 2018	Workshop on Learning in Verification (LiVE), ETAPS 2018, Thessaloniki, Greece (Keynote Speaker)

## ORGANIZATIONAL ACTIVITIES

---

JAN 2019	DALI symposium on machine learning <a href="#">Workshop on Security and Robustness</a>
MAY 2017	Control of Complex Systems: <a href="#">An Integrated Perspective on Modern Power Grid Control</a> American Controls Conference, Seattle, WA

## SCHOLARSHIPS

---

AUG. 2014	Postdoctoral fellowship, Caltech Center for Mathematics of Information (USD 60,000 p.a)
SEP. 2008	UW CSE Research fellowship, University of Washington (USD 30,000 p.a)

## PROGRAMMING AND TECHNICAL SKILLS

---

Programming Languages: JAVA, C, HTML (basic), C++, JULIA (intermediate), PYTHON  
MATLAB (advanced)  
AI/ML: TENSORFLOW (intermediate)  
Optimization: JUMP, GUROBI, CVX, MOSEK, CVXPY  
Presentations/Writing: POWERPOINT, BEAMER,  $\LaTeX$ , WORD  
Math & CS: Probability, Statistics, Stochastic Calculus  
Convex analysis, Combinatorial and Continuous Optimization  
Machine Learning, Game Theory, Control theory

## INTERESTS AND ACTIVITIES

---

Geopolitics, Global Economy, Neuroscience, Cognitive Psychology  
Mindfulness Meditation (Completed courses at UW and DeepMind), Hiking, Backpacking

## PUBLICATIONS

---

- [1] Rudy Bunel, Alessandro De Palma, Alban Desmaison, Krishnamurthy Dvijotham, Pushmeet Kohli, Philip Torr, and M. Pawan Kumar. Lagrangian decomposition for neural network verification. volume 124 of *Proceedings of Machine Learning Research*, pages 370–379, Virtual, 03–06 Aug 2020. PMLR.
- [2] Taylan Cemgil, Sumedh Ghaisas, Krishnamurthy (Dj) Dvijotham, and Pushmeet Kohli. Adversarially robust representations with smooth encoders. In *International Conference on Learning Representations*, 2020.
- [3] Krishnamurthy (Dj) Dvijotham\*, Jamie Hayes, Borja Balle, Zico Kolter, Chongli Qin, Andras Gyorgy, Kai Xiao, Sven Gowal, and Pushmeet Kohli. A framework for robustness certification of smoothed classifiers using f-divergences. In *International Conference on Learning Representations*, 2020.
- [4] Krishnamurthy (Dj) Dvijotham\*, Aditi Raghunathan, Jonathan Uesato, Sumanth Dathathri, Alexey Kurakin, Ian Goodfellow, Pushmeet Kohli, Jacob Steinhardt, and Percy Liang. Enabling certification of verification-agnostic networks via memory-efficient semidefinite programming. In *Advances in Neural Information Processing Systems*, pages –, 2020.
- [5] Sumedh Ghaisas, Taylan Cemgil, Krishnamurthy (Dj) Dvijotham, Sven Gowal, and Pushmeet Kohli. The autoencoding variational autoencoder. In *Advances in Neural Information Processing Systems*, pages –, 2020.
- [6] Olliver Hinder, Rudy Bunel, Shrinadh Bhojanapalli, and Krishnamurthy (Dj) Dvijotham. An efficient nonconvex reformulation of stagewise convex optimization problems. In *Advances in Neural Information Processing Systems*, pages –, 2020.
- [7] Johannes Welbl, Po-Sen Huang, Robert Stanforth, Sven Gowal, Krishnamurthy (Dj) Dvijotham, Martin Szummer, and Pushmeet Kohli. Towards verified robustness under text deletion interventions. In *International Conference on Learning Representations*, 2020.
- [8] Tsui-Wei Weng, Krishnamurthy (Dj) Dvijotham, Jonathan Uesato, Kai Xiao, Sven Gowal, Robert Stanforth, and Pushmeet Kohli. Toward evaluating robustness of deep reinforcement learning with continuous control. In *International Conference on Learning Representations*, 2020.
- [9] Anton Zhernov, Krishnamurthy Dj Dvijotham, Ivan Lobov, Dan A. Calian, Michelle Gong, Natarajan Chandrashekar, and Timothy A. Mann. The nodehopper: Enabling low latency ranking with constraints via a fast dual solver. In *Proceedings of the 26th ACM SIGKDD*

*International Conference on Knowledge Discovery & Data Mining*, KDD '20, page 1285–1294, New York, NY, USA, 2020. Association for Computing Machinery.

- [10] D. Lee, H. D. Nguyen, K. Dvijotham, and K. Turitsyn. Convex restriction of power flow feasibility sets. *IEEE Transactions on Control of Network Systems*, 6(3):1235–1245, Sep. 2019.
- [11] Chenglong Wang, Rudy Bunel, Krishnamurthy Dvijotham, Po-Sen Huang, Edward Grefenstette, and Pushmeet Kohli. Knowing when to stop: Evaluation and verification of conformity to output-size specs. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.
- [12] Krishnamurthy (Dj) Dvijotham\*, Robert Stanforth, Sven Gowal, Chongli Qin, Soham De, and Pushmeet Kohli. Efficient neural network verification with exactness characterization. In *Proceedings of the Thirty-Fifth Conference on Uncertainty in Artificial Intelligence, UAI 2019, Tel Aviv, Israel, July 22-25, 2019*, page 164, 2019.
- [13] Sven Gowal, Krishnamurthy Dvijotham\*, Robert Stanforth, Rudy Bunel, Chongli Qin, Jonathan Uesato, Relja Arandjelovic, Timothy A. Mann, and Pushmeet Kohli. On the effectiveness of interval bound propagation for training verifiably robust models. In *ICCV*, volume 9, pages –, 2019.
- [14] Po-Sen Huang, Robert Stanforth, Johannes Welbl, Chris Dyer, Dani Yogatama, Sven Gowal, Krishnamurthy Dvijotham, and Pushmeet Kohli. Achieving verified robustness to symbol substitutions via interval bound propagation. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing, EMNLP 2019, Hong Kong, China, November 5-7, 2019*, pages –, 2019.
- [15] Chongli Qin, Krishnamurthy (Dj) Dvijotham\*, Brendan O’Donoghue, Rudy Bunel, Robert Stanforth, Sven Gowal, Jonathan Uesato, Grzegorz Swirszcz, and Pushmeet Kohli. Verification of non-linear specifications for neural networks. In *International Conference on Learning Representations*, 2019.
- [16] Chongli Qin, James Martens, Sven Gowal, Dilip Krishnan, Krishnamurthy Dvijotham, Alhussein Fawzi, Soham De, Robert Stanforth, and Pushmeet Kohli. Adversarial robustness through local linearization. In *Advances in Neural Information Processing Systems*, pages –, 2019.
- [17] Jonathan Uesato, Ananya Kumar, Csaba Szepesvari, Tom Erez, Avraham Ruderman, Keith Anderson, Krishnamurthy (Dj) Dvijotham, Nicolas Heess, and Pushmeet Kohli. Rigorous agent evaluation: An adversarial approach to uncover catastrophic failures. In *International Conference on Learning Representations*, 2019.
- [18] S. Misra, D. K. Molzahn, and Krishnamurthy Dvijotham. Optimal adaptive linearizations of the ac power flow equations. In *2018 Power Systems Computation Conference (PSCC)*, pages 1–7, June 2018.
- [19] Krishnamurthy Dvijotham\*, Marta Garnelo, Alhussein Fawzi, and Pushmeet Kohli. Verification of deep probabilistic models. In *Advances in Neural Information Processing Systems, SecML workshop*, 2018.
- [20] Krishnamurthy Dvijotham\*, Robert Stanforth, Sven Gowal, Timothy Mann, and Pushmeet Kohli. A dual approach to scalable verification of deep networks. In *Proceedings of the Thirty-Fourth Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-18)*, Corvallis, Oregon, 2018. AUAI Press.
- [21] Vinod Nair, Krishnamurthy Dvijotham\*, Iain Dunning, and Oriyol Vinyals. Learning fast optimizers for contextual stochastic integer programs. In *Proceedings of the Thirty-Fourth Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-18)*, Corvallis, Oregon, 2018. AUAI Press.

- [22] H. D. Nguyen, Krishnamurthy Dvijotham\*, and K. Turitsyn. Constructing convex inner approximations of steady-state security regions. *IEEE Transactions on Power Systems*, pages 1-1, 2018.
- [23] H. D. Nguyen, Krishnamurthy Dvijotham\*, S. Yu, and K. Turitsyn. A framework for robust long-term voltage stability of distribution systems. *IEEE Transactions on Smart Grid*, pages 1-1, 2018.
- [24] Krishnamurthy Dvijotham\*, E. Mallada, and J. W. Simpson-Porco. High-voltage solution in radial power networks: Existence, properties, and equivalent algorithms. *IEEE Control Systems Letters*, 1(2):322-327, Oct 2017.
- [25] Krishnamurthy Dvijotham\*, H. Nguyen, and K. Turitsyn. Solvability regions of affinely parameterized quadratic equations. *IEEE Control Systems Letters*, PP(99):1-1, 2017.
- [26] Krishnamurthy Dvijotham\*, Yuval Rabani, and Leonard Schulman. Convergence of incentive-driven dynamics in fisher markets. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, January 10-12, 2017*, pages 2039-2052, 2017.
- [27] N. Azizan Ruhi, Krishnamurthy Dvijotham\*, N. Chen, and A. Wierman. Opportunities for price manipulation by aggregators in electricity markets. *IEEE Transactions on Smart Grid*, PP(99):1-1, 2017.
- [28] Y. Tang, Krishnamurthy Dvijotham, and S. Low. Real-time optimal power flow. *IEEE Transactions on Smart Grid*, PP(99):1-1, 2017.
- [29] D. Wu, D. K. Molzahn, B. C. Lesieutre, and Krishnamurthy Dvijotham. A deterministic method to identify multiple local extrema for the ac optimal power flow problem. *IEEE Transactions on Power Systems*, PP(99):1-1, 2017.
- [30] Krishnamurthy Dvijotham\* and D. Molzahn. Error bounds on the dc power flow approximation: A convex relaxation approach. In *2016 55th IEEE Conference on Decision and Control (CDC)*, pages 23-30, Dec 2016.
- [31] Krishnamurthy Dvijotham\*, Michael Chertkov, Pascal Van Hentenryck, Marc Vuffray, and Sidhant Misra. Graphical models for optimal power flow. *Constraints*, pages 1-26, 2016.
- [32] Krishnamurthy Dvijotham\*, M. Chertkov, and S. Low. A differential analysis of the power flow equations. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 23-30, Dec 2015.
- [33] Krishnamurthy Dvijotham\* and M. Chertkov. Convexity of structure preserving energy functions in power transmission: Novel results and applications. In *2015 American Control Conference (ACC)*, pages 5035-5042, July 2015.
- [34] Krishnamurthy Dvijotham\*. Systems of quadratic equations: Efficient solution algorithms and conditions for solvability. In *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1027-1031. IEEE, 2015.
- [35] Krishnamurthy Dvijotham\*, Emanuel Todorov, and Maryam Fazel. Convex structured controller design in finite horizon. *IEEE Transactions on Control of Network Systems*, 2(1):1-10, 2015.
- [36] Krishnamurthy Dvijotham\*. *Automating Stochastic Optimal Control*. PhD thesis, 2014.
- [37] Krishnamurthy Dvijotham\*, Misha Chertkov, and Scott Backhaus. Storage sizing and placement through operational and uncertainty-aware simulations. In *2014 47th Hawaii International Conference on System Sciences*, pages 2408-2416. IEEE, 2014.
- [38] Krishnamurthy Dvijotham\*, Maryam Fazel, and Emanuel Todorov. Convex risk averse control design. In *53rd IEEE Conference on Decision and Control*, pages 4020-4025. IEEE, 2014.

- [39] Krishnamurthy Dvijotham\*, Maryam Fazel, and Emanuel Todorov. Universal convexification via risk-aversion. In *Proceedings of the Thirtieth Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-14)*, pages 162–171, Corvallis, Oregon, 2014. AUAI Press.
- [40] Krishnamurthy Dvijotham\* and R. Sharma. Battery life estimation in a real-time energy management system. In *2013 IEEE Power Energy Society General Meeting*, pages 1–5, July 2013.
- [41] Krishnamurthy Dvijotham\*, Evangelos Theodorou, Emanuel Todorov, and Maryam Fazel. Convexity of optimal linear controller design. In *52nd IEEE Conference on Decision and Control*. IEEE, 2013.
- [42] Krishnamurthy Dvijotham\* and E. Todorov. *Linearly Solvable Optimal Control*, pages 119–141. John Wiley & Sons, Inc., 2013.
- [43] Krishnamurthy Dvijotham\*, Emanuel Todorov, and Maryam Fazel. Convex control design via covariance minimization. In *Annual Allerton Conference on Communication, Control, and Computing*. IEEE, 2013.
- [44] Evangelos Theodorou, Krishnamurthy Dvijotham\*, and Emo Todorov. From information theoretic dualities to path integral and kullback-leibler control: Continuous and discrete time formulations. In *The Sixteenth Yale Workshop on Adaptive and Learning Systems*, 2013.
- [45] Evangelos Theodorou, Krishnamurthy Dvijotham\*, and Emo Todorov. Time varying non-linear policy gradients. In *CDC*, pages 7765–7770, 2013.
- [46] Krishnamurthy Dvijotham\*, Scott Backhaus, and Michael Chertkov. Distributed control of generation in a transmission grid with a high penetration of renewables. In *Smart Grid Communications (SmartGridComm), 2012 IEEE Third International Conference on*, pages 635–640. IEEE, 2012.
- [47] Krishnamurthy Dvijotham\* and Emo Todorov. Linearly solvable markov games. In *2012 American Control Conference (ACC)*, pages 1845–1850. IEEE, 2012.
- [48] Krishnamurthy Dvijotham\* and Emanuel Todorov. A unifying framework for linearly solvable control. In *Proceedings of the Twenty-Seventh Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-11)*, pages 179–186, Corvallis, Oregon, 2011. AUAI Press.
- [49] Krishnamurthy Dvijotham\* and Maryam Fazel. A nullspace analysis of the nuclear norm heuristic for rank minimization. In *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3586–3589. IEEE, 2010.
- [50] Krishnamurthy Dvijotham\* and Emanuel Todorov. Inverse optimal control with linearly-solvable mdps. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pages 335–342, 2010.
- [51] Krishnamurthy Dvijotham\*, Soumen Chakrabarti, and Subhasis Chaudhuri. New closed-form bounds on the partition function. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 8–8. Springer Berlin Heidelberg, 2008.

\* indicates I was a primary contributor