Seyedmohammadhadi Daneshmand

Email: hdanesh@mit.edu, hadi.daneshmand@gmail.com.
Phone: +18577990358,
Google Scholar

Professional Experiences

- **Postdoctoral Research Associate** at MIT Laboratory for Information & Decision Systems, USA, since December 2022.
- **Postdoctoral Research Fellow** at Princeton University, USA 2022.
- **Postdoctoral Researcher** at French Institute for Research in Computer Science and Automation, France 2020-2022.
- **Research and Teaching Assistant** at ETH Zurich, Switzerland 2014-2020.
- **Research Intern** at Vector Institute, Canada 2019.
- **Research Intern** at Max Planck Institute for Intelligent Systems, Germany 2013.

Education

PhD. in *Computer Science*, ETH-Zurich 2014-2020.
M.S. in *Computer Engineering – Artificial Intelligence*, Sharif university of technology, Tehran, GPA: 18.54/20
B.S. in *Computer Engineering*, Sharif university of technology, Tehran, GPA: 17.97/20

Interested in

optimization, theory of representation learning, deep neural networks, statistical learning

Preprints

Kwangjun Ahn*, Xiang Cheng*, Hadi Daneshmand*, Suvrit Sra Transformers learn to implement preconditioned gradient descent for in-context learning, 2013.

Amir Joudaki, Hadi Daneshmand, Francis Bach. On the impact of activation and normalization in obtaining isometric embeddings at initialization, 2013.

Publications

Hadi Daneshmand, Jason D Lee, Chi Jin, Efficient Displacement Convex Optimization with Particle Gradient Descent, International Conference on Machine Learning, 2023.

Amir Joudaki, Hadi Daneshmand, Francis Bach, On Bridging the Gap between Mean Field and Finite Width in Deep Random Neural Networks with Batch Normalization, International Conference on Machine Learning, 2023.

Hadi Daneshmand, Amir Joudaki, Francis Bach, Batch Normalization Orthogonalizes Representations in Deep Random Networks, Conference on Neural Information Processing Systems 2021. **Spotlight presentation (among the top 3 % submissions)**.

Peiyuan Zhang*, Antonio Orvieto*, Hadi Daneshmand Rethinking the Variational Interpretation of Nesterov's Accelerated Method, Conference on Neural Information Processing Systems 2021.

Peiyuan Zhang, Antonio Orvieto, Hadi Daneshmand, Thomas Hofmann, Roy S. Smith, Revisiting the Role of Euler Numerical Integration on Acceleration and Stability in Convex Optimization, International Conference on Artificial Intelligence and Statistics 2021.

Hadi Daneshmand*[1], Jonas Kohler*, Francis Bach, Thomas Hofmann, Aurelien Lucchi. Batch Normalization Provably Avoids Rank Collapse for Randomly Initialised Deep Networks, Conference on Neural Information Processing Systems 2020.

Hadi Daneshmand. Optimization for Neural Networks: Quest for Theoretical Understandings. Ph.D. thesis, ETH Zurich 2020.

Jonas Kohler* , Hadi Daneshmand* , Aurelien Lucchi, Ming Zhou , Klaus Neymeyr, Thomas Hofmann. Exponential convergence rates for Batch Normalization: The power of length-direction decoupling in non-convex optimization, International Conference on Artificial Intelligence and Statistics 2019.

Leonard Adolphs, Hadi Daneshmand, Aurelien Lucchi, Thomas Hofmann Local Saddle Point Optimization: A Curvature Exploitation Approach, International Conference on Artificial Intelligence and Statistics 2019.

Hadi Daneshmand*, Jonas Kohler*, Aurelien Lucchi, Thomas Hofmann. Escaping Saddles with Stochastic Gradients, International Conference on Machine Learning 2018. **Long presentation (among the top %8 submissions)**.

Aryan Mokhtari*, Hadi Daneshmand*, Aurelien Lucchi, Thomas Hofmann, Alejandro Ribeiro. Adaptive Newton method for empirical risk minimization to statistical accuracy. Conference on Neural Information Processing Systems 2016.

Hadi Daneshmand, Aurelien Lucchi, Thomas Hofmann, Starting Small ‐ Learning with Adaptive Sample Sizes, International Conference on Machine Learning (ICML) 2016.

Steven M Hill, Laura M Heiser, et al. Inferring causal molecular networks: empirical assessment through a community-based effort, Nature Methods 2016.

Manuel Gomez Rodriguez, Le Song, Hadi Daneshmand, and Bernhard Scholkopf; Estimating Diffusion Network Structure: Recovery Conditions, Sample Complexity, and a Soft-thresholding algorithm, Journal of Machine Learning Researches (JMLR) 2016.

Hadi Daneshmand, Manuel Gomez Rodriguez, Le Song, and Bernhard Scholkopf; Estimating Diffusion Network Structures: Recovery Conditions, Sample Complexity & Soft-thresholding Algorithm, International Conference on Machine Learning (ICML) 2014. **Recommended to JMLR fast track (18 out of 1260+)**

Hadi Daneshmand, Amin Javari, Seyed Ebrahim Abtahi, and Mahdi Jalili; A Time-aware Recommender System based on Dependency Network of Items, Oxford computer journal 2014.

Awards

- Early Postdoctoral Mobility Grant of Swiss National Science Foundation 2020.
- Outstanding reviewer award, ICML 2022 and 2019.
- Tenure-track (W2) offer from Saarland University 2022 (declined).

Teaching Assistant

ETH Zurich, Computational Intelligence Lab (2019, 2016, and 2015)

ETH Zurich, Deep Learning (2018, and 2017)

ETH Zurich, Machine Learning (2018, and 2016)

[1] * marks equal contributions.

Mentorships

- Alexandru Meterez, Master's thesis on dynamical isometry in deep neural networks, ETHZ 2023-current.
- Flowers Alec Massimo, Master's thesis on data representation in deep neural networks, ETHZ 2023-current.
- Kwangjun Ahn, for his PhD project on in-context learning.
- Bense Alexandre's Master's Thesis on Batch Normalization for non-linear deep neural networks, ETH Zurich.
- Leonard Adolphs's Master's Thesis on saddle point optimization published at AISTATS19, ETH Zurich.
- Peiyuan Zhang's Master's Thesis on Accelerated Optimization published at AISTATS21, and NeurIPS21 conferences, ETH Zurich.
- Amir Joudaki for his PhD research on optimization for deep neural networks published at ICML23 and NeurIPS21.
- Antonio Orvieto for his PhD research on accelerated optimization published at AISTATS21, and NeurIPS21.

Talks

ML tea talk at MIT, 2023.

Machine Learning Seminar Series at Princeton University, USA 2022.

Winter Seminar Series at Sharif University of Technology, Iran 2017, 2022.

INRIA Junior Seminar, France 2021.

Reviewer for

Neurocomputing Journal, IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Transactions on Signal and Information Processing over Networks, Elsevier Journal on Online Social Networks and Media, Conference on Neural Information Processing Systems, International Conference on Machine Learning, Data Mining and Knowledge Discovery, International Conference on Artificial Intelligence and Statistics, and International Conference on Learning Representations.