Khashayar Gatmiry

Cambridge, MA | gatmiry@mit.edu | gatmiry.github.io | Google Scholar

Education

Massachusetts Institute of Technology, PhD in Computer Science — Sept 2019 – Present

- **Advisors:** Stefanie Jegelka, Jonathan Adam Kelner

Sharif University of Technology, B.S. in Computer Engineering, & Mathematics — Aug 2014, June 2019

- **Mentors:** Manuel Gomez Rodriguez, Abolfazl Motahari

Research interests & Selected Publications

– My research is on designing better machine learning algorithms with provable guarantees, as well as understand and improve the existing ones, through the lens of optimization and sampling.

Selected Publications

- Learning Mixtures of Gaussians Using Diffusion Models
  Khashayar Gatmiry, Jonathan A. Kelner Holden Lee
  ARXIV, 2024
- Sampling Polytopes with Riemannian HMC: Faster Mixing via the Lewis Weights Barrier
  Khashayar Gatmiry, Jonathan A. Kelner Santosh S. Vempala
  COLT, 2024
- The Inductive Bias of Flatness Regularization for Deep Matrix Factorization
  Khashayar Gatmiry, Zhiyuan Li Ching-Yao Chuang Sashank J. Reddi Tengyu Ma Stefanie Jegelka
  NEURIPS, 2024

Industry Experience

Research Intern, Microsoft – Redmond, WA — June 2024 – Aug 2024

- **Mentors:** Ronen Eldan, Adil Salim, Yi Zhang
- Worked on generating paragraph semantic embeddings with small language models (SLMs), based on pytorch

Research intern, Google – New York — Oct 2022 – Jan 2023

- **Mentors:** Sashank J. Reddi, Sobhan Miryoosefi
- Trained language models using novel gradient estimates, authored 4 papers in collaboration with my group

Research Visiting Experience

Visiting Researcher, Yale Computer Science – New Heaven, CT — July 2023 – Aug 2023

- **Mentor:** Andre Wibisono
- Worked on Hamiltonian Monte Carlo and Gibbs sampler

Visiting Researcher, Georgia Tech College of Computing – Atlanta, GA — Oct 2022 – Jan 2023

- **Mentor:** Santosh S. Vempala
- worked on proving isoperimetric inequalities on manifolds, KLS conjecture, and sampling from polytopes Authored 2 publications from projects related to my visit

Research Intern, Max Plank Institute for Software Systems – Kaiserslautern, GER — June 2018 – Aug 2018

- **Mentor:** Manuel Gomez Rodriguez
- worked on the network visibility problem and submodularity

Review Service

Reviewer for Neural Information Processing System (NEURIPS), Conference on Learning Representations (ICLR), Fundamentals of Computer Science (FOCS), Symposium on Discrete Algorithms (SODA), Journal of Machine Learning Research (JMLR), Theoretical Computer Science (TCS)

Teaching Service

Teaching Assistant for Introduction to Probability and Statistics (18.05), MIT	Spring 2024
Teaching Assistant for Computation Theory and Automata, Sharif university	Spring 2018
Teaching Assistant for Artificial Intelligence, Sharif university	Spring 2018
Teaching Assistant for Probability and Statistics, Sharif university	Fall 2017
Teaching Assistant for Discrete Structures, Sharif university	Fall 2016
Teaching Math Olympiad in high schools during my undergraduate studies	

Awards

Multiple NSF fellowship awards including AI Institute TILOS and TRIPODS program	PhD
Fellowship from the Institute of Science and Technology Austria (**IST**)	Summer 2018
Outstanding Student Award from Sharif University of Technology	2018
Recipient of the grant for exceptional talents *National Elites Foundation of Iran*	undergraduate studies
Gold Medal, 27^{th} Iranian National Olmypiad in Mathematics	2013
Ranked 4^{th}, International Tournament of Young Mathematicians (**ITYM**), Germany	2014
Silver Medal, International Mathematics Competition (**IMC**), Taiwan	2012
Bronze Medal, International Mathematics Competition (**IMC**), Indonesia	2011
Gold Medal, Asian Inter-cities Teenagers Mathematics Olympiad (**AITMO**), Nepal	2011

Publications

Learning mixtures of gaussians using diffusion models
Khashayar Gatmiry, Jonathan A. Kelner, Holden Lee,
PREPRINT, 2024

Can Looped Transformers Learn to Implement Multi-step Gradient Descent for In-context Learning?
Khashayar Gatmiry, Nikunj Saunshi, Sashank J. Reddi, Stefanie Jegelka, Sanjiv Kumar,
ICML, 2024

Simplicity Bias via Global Convergence of Sharpness Minimization
Khashayar Gatmiry, Zhiyuan Li, Sashank J. Reddi, Stefanie Jegelka
ICML, 2024

What does guidance do? A fine-grained analysis in a simple setting
Muthu Chidambaram, ***Khashayar Gatmiry***, Sitan Chen, Holden Lee, Jianfeng Lu,
NEURIPS, 2024

Adversarial Online Learning with Temporal Feedback Graphs
Khashayar Gatmiry, Jon Schneider, Stefanie Jegelka,
COLT, 2024

Sampling Polytopes with Riemannian HMC: Faster Mixing via the Lewis Weights Barrier
Khashayar Gatmiry, Jonathan A. Kelner, Santosh Vempala,
COLT, 2024

EM for Mixture of Linear Regression with Clustered Data
Amir Reisizadeh, ***Khashayar Gatmiry***, Asuman Ozdaglar,
AISTATS, 2024

Bandit Algorithms for Prophet Inequality and Pandora's Box
Khashayar Gatmiry, Thomas Kesselheim, Sahil Singla, Yifan Wang

SODA, 2024

Rethinking Invariance in In-context Learning
Lizhe Fang, Yifei Wang, ***Khashayar Gatmiry***, Yisen Wang
ICML WORKSHOP, 2024

Projection-free online convex optimization via efficient newton iterations
Khashayar Gatmiry, Zak Mhammedi
NEURIPS, 2023

The Inductive Bias of Flatness Regularization for Deep Matrix Factorization
Khashayar Gatmiry, Zhiyuan Li, Tengyu Ma, Sashank J. Reddi, Stefanie Jegelka, Ching-Yao Chuang,
NEURIPS, 2023

Quasi-newton steps for efficient online exp-concave optimization
Khashayar Gatmiry, Zak Mhammedi
COLT, 2023

Quasi-newton steps for efficient online exp-concave optimization
Khashayar Gatmiry, Zak Mhammedi
COLT, 2023

When does Metropolized Hamiltonian Monte Carlo provably outperform Metropolis-adjusted Langevin algorithm?
Yuansi Chen, ***Khashayar Gatmiry***
PREPRINT, 2023

A unified approach to controlling implicit regularization via mirror descent
Haoyuan Sun, ***Khashayar Gatmiry***, Kwangjun Ahn, Navid Azizan
JMLR, 2023

Optimal algorithms for group distributionally robust optimization and beyond
Tasuku Soma, ***Khashayar Gatmiry***, Stefanie Jegelka
PREPRINT, 2022

On the generalization of learning algorithms that do not converge
Nisha Chandramoorthy, Andreas Loukas, ***Khashayar Gatmiry***, Stefanie Jegelka
NEURIPS, 2022

Adaptive Generalization and Optimization of Three-Layer Neural Networks
Khashayar Gatmiry, Stefanie Jegelka, Jonathan A. Kelner
ICLR, 2022

Convergence of the riemannian langevin algorithm
Khashayar Gatmiry, Santosh S. Vempala
PREPRINT

Testing determinantal point processes
Khashayar Gatmiry, Maryam Aliakbarpour, Stefanie Jegelka
NEURIPS, 2020

The network visibility problem
Khashayar Gatmiry, Manuel Gomez-Rodriguez
ACM TOIS, 2021

Information Theoretic Bounds on Optimal Worst-case Error in Binary Mixture Identification
Khashayar Gatmiry, Seyed Abolfazl Motahari
PREPRINT