

# Aniket Das

PRE-DOCTORAL RESEARCHER, GOOGLE RESEARCH INDIA

EDUCATION	<b>Indian Institute of Technology Kanpur</b> <i>BTech in Electrical Engineering and BS in Mathematics (Double Major)</i> <b>GPA : Overall 9.3/10 Mathematics 9.8/10</b>	<i>Aug' 17 - May' 22</i> <i>8 Semesters</i>
	<b>Aalto University, Finland</b> <i>Year-long Academic Exchange in Aalto University School of Science</i> <b>GPA : 4.8/5.0</b>	<i>Jan' 20 - Jan' 21</i> <i>2 Semesters</i>

INTERESTS	Sampling Algorithms, Markov Chains, Applied Probability, High Dimensional Statistics
-----------	--

PUBLICATIONS	<b><i>Provably Fast Finite-Particle Variants of SVGD via Virtual Particle Stochastic Approximation</i></b> <b>Aniket Das, Dheeraj Nagaraj</b> [αβ] <i>Neural Information Processing Systems (NeurIPS) 2023 [Spotlight] [Paper]</i> <b><i>Oral Presentation at Optimal Transport and Machine Learning Workshop @ NeurIPS 2023</i></b>
	<b><i>Utilising the CLT Structure in Stochastic Gradient based Sampling : Improved Analysis and Faster Algorithms</i></b> <b>Aniket Das, Dheeraj Nagaraj, Anant Raj</b> [αβ] <i>Conference On Learning Theory (COLT) 2023 [Paper]</i>
	<b><i>Near-Optimal Heteroscedastic Regression with Symbiotic Learning</i></b> <b>Dheeraj Baby, Aniket Das, Dheeraj Nagaraj, Praneeth Netrapalli</b> [αβ] <i>Conference On Learning Theory (COLT) 2023 [Paper]</i>
	<b><i>Sampling without Replacement Leads to Faster Rates in Finite-Sum Minimax Optimization</i></b> <b>Aniket Das, Bernhard Schölkopf, Michael Muehlebach</b> <i>Neural Information Processing Systems (NeurIPS) 2022 [Paper]</i>
	<b><i>NeurInt - Learning Interpolation by Neural ODEs</i></b> <b>Avinandan Bose*, Aniket Das*, Yatin Dandi, Piyush Rai</b> <i>Workshop on Deep Learning and Differential Equations @ NeurIPS 2021 [Spotlight] [Paper]</i>
	<b><i>Jointly Trained Image and Video Generation using Residual Vectors</i></b> <b>Yatin Dandi, Aniket Das, Soumye Singhal, Vinay P. Namboodiri, Piyush Rai</b> <i>IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2020 [Paper]</i>
	<b><i>TorchGAN: A Flexible Framework for GAN Training and Evaluation</i></b> <b>Avik Pal*, Aniket Das*</b> <i>Journal of Open Source Software (JOSS) [Paper]</i>
	[αβ] : indicates alphabetical ordering      * : indicates equal contribution

EXPERIENCE	<b>Google Research India</b> <i>Pre-Doctoral Researcher, Machine Learning and Optimization (MLO)</i>	<i>July '22 - Present</i>
	<ul style="list-style-type: none"> <li>- Working on sampling, spin systems, high dimensional statistics and stochastic optimization</li> <li>- Published two papers at COLT 2023 on SGLD and minimax optimal heteroscedastic regression, and one spotlight paper at NeurIPS 2023 on finite particle convergence of SVGD.</li> </ul>	

## Max Planck Institute for Intelligent Systems, Tübingen

[Remote]

Advisors : [Michael Muehlebach](#) and [Bernhard Schölkopf](#)

July '21 - Dec '21

- Worked on stochastic minimax optimization and gradient flows for constrained optimization
- Paper on sampling without replacement for minimax optimization published at NeurIPS 2022

## Tata Institute of Fundamental Research, Mumbai

[Remote]

Advisor : [Sandeep Juneja](#)

Apr '21 - Jun '21

- Worked on instance-dependent lower bounds for PAC learning in Markov Decision Processes and structured stochastic bandits

---

## SELECTED PROJECTS

### Rapid Convergence of Finite-Particle SVGD using Virtual Particles

Advisor : [Dheeraj Nagaraj](#), Google Research [\[NeurIPS'23 Spotlight + OTML'23 Oral\]](#)

- Developed computationally efficient variants of **Stein Variational Gradient Descent (SVGD)** with *provably fast convergence in the finite-particle regime*
- Obtained the *first known gradient oracle complexity* for any SVGD-type algorithm that enjoys *polynomial scaling in dimension and error tolerance*
- Demonstrated a *double exponential improvement* over prior state-of-the-art
- Developed a novel and highly general technique for designing stochastic approximations in the space of measures that admit an *exact finite-particle implementation*
- Applied techniques from Optimal Transport, Geometry, Dynamical Systems and Functional Analysis.
- Accepted as a **Spotlight Paper** at *Neural Information Processing Systems (NeurIPS) 2023*.
- Also accepted for an **Oral Presentation** at the *Optimal Transport and Machine Learning (OTML) Workshop, NeurIPS 2023*

### CLT Analysis of Stochastic-Gradient Based Sampling

Advisor : [Dheeraj Nagaraj](#), Google Research

[\[COLT'23 Paper\]](#)

- Developed novel non-asymptotic Central Limit Theorems to analyze the interaction between the stochastic approximation noise and diffusion noise in stochastic-gradient based sampling algorithms
- Obtained state of the art convergence rates for **Stochastic Gradient Langevin Dynamics (SGLD)** under *minimal isoperimetric assumptions*
- Analyzed the **Random Batch Method (RBM)** for simulating Interacting Particle Dynamics and derived state of the art trajectory-level guarantees under minimal assumptions
- Derived the first known convergence rates for the Covariance Correction heuristic and proved that it enables faster convergence of SGLD and RBM without added computational complexity
- Applied techniques from Stochastic Calculus, Markov Chains and Optimal Transport
- Work is published at *Conference on Learning Theory (COLT), 2023*

### Minimax Optimal Heteroscedastic Linear Regression

Advisors : [Dheeraj Nagaraj](#) and [Praneeth Netrapalli](#), Google Research

[\[COLT'23 Paper\]](#)

- Developed a computationally efficient **Alternating Minimization** algorithm for heteroscedastic linear regression that exhibits *minimax optimal sample complexity* (modulo log factors).
- Designed provably efficient algorithms for *linear regression and phase retrieval under the multiplicative noise (or noisy covariate) model*
- Utilised the spectral properties of rank-deficient Wishart matrices to design a *novel adaptation of LeCam's method which is robust to infinite mutual information quantities*
- Applied techniques from High Dimensional Statistics, Random Matrices and Information Theory.
- Work is published at *Conference on Learning Theory (COLT), 2023*

## Sampling without Replacement for Finite-Sum Minimax Optimization

Advisors : *Michael Muehlebach and Bernhard Schölkopf, MPI-IS* [NeurIPS'22 Paper]

- Analyzed stochastic gradient minimax optimization algorithms that sample the data points *without replacement* and demonstrated that they lead to faster convergence than uniform sampling.
- Derived *near-optimal rates for GDA and PPM with Random Reshuffling, Single Shuffling and Incremental Gradient* for solving finite-sum strongly monotone variational inequalities
- Used *alternating updates* and without-replacement sampling to design an algorithm that outperforms sampling with replacement for *nonconvex-nonconcave problems satisfying a two-sided PL inequality*
- Provably demonstrated the effectiveness of *data ordering attacks* on finite-sum minimization and minimax optimization by deriving *near-optimal rates under the Adversarial Shuffling model*.
- Utilised techniques from Game Theory, Variational Inequalities and Stochastic Optimization
- Work is published at *Neural Information Processing Systems (NeurIPS), 2022*

## Near-Optimal Streaming Heavy Tailed Stochastic Optimization

Advisors : *Dheeraj Nagaraj and Arun Sai Suggala, Google Research*

- Rigorously analyzed the popular clipped SGD heuristic for *Heavy-tailed Stochastic Convex Optimization (HT-SCO)* in the streaming setting
- Proved that clipped SGD nearly achieves the *optimal sub-Gaussian statistical rate* for HT-SCO under smoothness and strong convexity
- Applied techniques from Stochastic Optimization and Heavy-tailed Statistics
- Currently working on extending our analysis to problems without strong convexity

---

### RELEVANT COURSEWORK

<b>Computer Science</b>	Introduction to Programming, Data Structures and Algorithms, Advanced Algorithms, Toolkit for Theoretical Computer Science <sup>II</sup> ,
<b>Probability Statistics &amp; ML</b>	Optimization in ML, Kernel Methods and Learning Theory, Advanced Probability Theory <sup>II</sup> , Markov Chains and Mixing Times <sup>II</sup> , Probabilistic ML, State Space Models
<b>Mathematics</b>	Real Analysis, Complex Analysis, Functional Analysis, Topology, Measure Theory, Differential Geometry, Dynamical Systems, Ordinary Differential Equations, Partial Differential Equations, Linear Algebra, Abstract Algebra, Numerical Methods

<sup>II</sup> : Audited Remotely at Tata Institute of Fundamental Research (TIFR), Mumbai

---

### TALKS

#### Sampling Through the Lens of Optimization : Recent Advances and Insights

*MSR-IISc Theory Seminar 2023* [Upcoming]

#### Utilising the CLT Structure in Stochastic Gradient-Based Sampling

*Conference on Learning Theory 2023*

#### Near-Optimal Heteroscedastic Regression with Symbiotic Learning

*Conference on Learning Theory 2023*

---

### SERVICE

<b>Reviewer</b>	JMLR, NeurIPS 2023, AISTATS 2022, NeurIPS 2021 DLDE Workshop
<b>Co-ordinator</b>	Special Interest Group in Machine Learning, IIT Kanpur
<b>Project Mentor</b>	Programming Club, IITK and Stamatics (Math Club), IITK