3443 S Morgan St, Chicago IL, 60608, US
+1 (630) 297 6261
asorokin@hawk.iit.edu
alegresor.github.io
aleksei-sorokin
alegresor
akk3XSEAAAAJ

Aleksei G. Sorokin

Background

Research Interests Probabilistic Numerics, Quasi-Monte Carlo, Gaussian Processes, Machine Learning

Softwares Docker, AWS, GitHub Actions, Travis CI, LaTeX

Frameworks Python (pandas, PyTorch, GPyTorch), Julia, C, MATLAB, R, SQL, Wolfram

Education

2021 - 2025 **PhD in Applied Math.** Illinois Institute of Technology (IIT). GPA $3.89/4$.

2017 - 2021 **Master of Data Science.** IIT. Summa cum laude. GPA $3.94/4$.

2017 - 2021 **B.S. in Applied Math, Minor in Computer Science.** IIT. Summa cum laude. GPA $3.94/4$.

Experiences

Summer 2023 **Graduate Intern** at **Los Alamos National Laboratory.** I modeled the solution processes of PDEs with random coefficients using efficient and error aware Gaussian processes. Resulted in publication of *Computationally Efficient and Error Aware Surrogate Construction for Numerical Solutions of Subsurface Flow Through Porous Media*.

Summer 2022 **Givens Associate Intern** at **Argonne National Laboratory**. I researched methods to efficiently estimate failure probability using Monte Carlo with non-parametric importance sampling. Resulted in publication of *Credible Intervals for Probability of Failure with Gaussian Processes*.

Summer 2021 **ML Engineer Intern** at **SigOpt, an Intel Company**. I developed novel meta-learning techniques for model-aware hyperparameter tuning via Bayesian optimization. In a six person ML engineering team, I contributed production code and learned key elements of the AWS stack. Resulted in publication of "SigOpt Mulch: An intelligent system for AutoML of gradient boosted trees".

Fall 2021 - Present **Teaching Assistant** at **IIT**. I lead review sessions for PhD qualifying exams in applied math.

Projects

Fast Gaussian Processes with Gradients The cost of Gaussian process regression can be reduced from $\mathcal{O}(n^3)$ to $\mathcal{O}(n \log n)$ when one has control over the design of experiments. This is achieved by pairing quasi-random sampling with matching kernels to induce structure in the kernel matrix. My PhD research studies generalizations for quickly incorporating gradient information into the ML model and using these efficient strategies to solve PDEs.

QMCPy I lead development of the open source project QMCPy, a Quasi-Monte Carlo Python Library. This package provides high quality quasi-random sequence generators, automatic variable transformations, adaptive stopping criteria algorithms, and diverse use cases. Over the past five years, this project has grown to dozens of collaborators and resulted in numerous presentations at a variety of conferences. Publications include "Quasi-Monte Carlo Software", *On Bounding and Approximating Functions of Multiple Expectations using Quasi-Monte Carlo*, *Challenges in Developing Great Quasi-Monte Carlo Software*, and "(Quasi-)Monte Carlo Importance Sampling with QMCPy". See *qmcpy.org* for more information.

Argonne: AI on Supercomputers I studied *AI Driven Science on Supercomputers* during my time at *Argonne National Laboratory*. Key topics included handling large scale data pipelines and parallel training for neural networks. Coursework at *github.com/alegresor/ai-science-training-series*.

Awards

2023 **Outstanding Math Poster**, Los Alamos National Laboratory.

2021 **Best Manuscript**, IIT Undergraduate Research Journal.

2020 **Karl Menger Student Award for Exceptional Scholarship**, IIT.

2017 - Present **Deans List Member**, IIT.