Rupanshu Soi

rsoi@stanford.edu

Education

2022—Now **Stanford University.** PhD in Computer Science.
Advisor: Alex Aiken.
Qualcomm Innovation Fellow 2024. Link.

2018-22 **BITS Pilani.** B.E. in Computer Science (with Distinction).
Thesis: Scaling Implicit Parallelism with Index Launches. Link.
Advisor: Elliott Slaughter, SLAC National Accelerator Laboratory.

Experience

2025—Now **NVIDIA and Stanford University.** Optimal compiler algorithms for Tensor Core GPUs.

2025 **Stanford University.** An incremental algorithm for equality saturation.

2024 **Stanford University.** Search-based optimization of ML models with equality saturation.

2023 **Stanford University.** Distributed fault resilience in the Legion programming model.

2022 **NVIDIA.** Automatic generation of error-checking code in the `ptxas` parser.

2021 **SLAC National Accelerator Laboratory.** Program analysis and optimizations in the Regent compiler.

2021 **UQAM, Canada.** Mapping LLVM passes using a microservice reverse-engineering approach.

2020 **BITS Pilani.** An implicitly-parallel meshfree solver in Regent.

Publications

Optimal Software Pipelining and Warp Specialization for Tensor Core GPUs.
RS*, Rohan Yadav*, Fredrik Kjolstad, Alex Aiken, Maryam Mehri Dehnavi,
Michael Garland and Michael Bauer. (* indicates equal contribution.)
Under submission.

Relight: Automatic, Distributed Checkpointing with Fast-Forward Replay.
Elliott Slaughter, RS, Michael Bauer, and Alex Aiken.
Under submission.

2025 **Incremental Equality Saturation.** Extended abstract. Talk.
RS, Benjamin Driscoll, Ke Wang, and Alex Aiken.
EGRAPHS '25.

2021 **Index Launches: Scalable, Flexible Representation of Parallel Task Groups.** Paper. Slides. Talk.
RS, Michael Bauer, Sean Treichler, Manolis Papadakis, Wonchan Lee, Patrick McCormick,
Alex Aiken, and Elliott Slaughter.
SC 2021.

2020 **An Implicitly Parallel Meshfree Solver in Regent.** Paper.
RS, Nischay Mamidi, Elliott Slaughter, Kumar Prasun, Anil Nemili, and SM Deshpande.
PAW-ATM 2020.

Selected Talks

2025 **Optimal Software Pipelining and Warp Specialization for Tensor Core GPUs.**
Stanford Software Lunch.
Tile IR Group, NVIDIA.
End-of-internship talk, NVIDIA.

2024 **Scalable Fault Resilience with Task Replay and Restart.** Slides.
NVIDIA.

Awards

2024 **Qualcomm Innovation Fellowship.** Link.

2021 **MITACS Globalink Research Internship.**

2019 **BITS Pilani Merit Scholarship.** Spring, Fall 2019.

2019 **Sir CV Raman Prize.** Department of Physics, BITS Pilani.