List of PhD Projects 2025

Division of Computational Mathematics
Department of Mathematics
Stockholm University

— Project 1 —

Computational methods for complex SV detection using sequencing data

Main supervisor: Kristoffer Sahlin (ksahlin@math.su.se)
Co-supervisor: Adam Ameur (adam.ameur@igp.uu.se)

In the Department of mathematics at Stockholm University, we are announcing the position as DDLS PhD student in **data driven cell and molecular biology**. Data driven cell and molecular biology covers research that fundamentally transforms our knowledge about how cells function by peering into their molecular components in time and space, from single molecules to native tissue environments.

The project aims to design data-driven algorithms to detect complex structural variants in humans using long DNA sequencing reads.

A *structural variant* (SV) is a large-scale alteration in the genome that involves rearranged, deleted, or inserted DNA segments with respect to the reference genome. A combination of SV events in a genomic region characterizes a complex SV. Complex SVs play an important role in many diseases. The data used to detect and examine SVs are whole genome *sequencing reads* (or simply *reads*). The reads mapping over and around variation sites can be used to classify the SV. Previous sequencing technologies often produced relatively short reads, making it impossible for reads to span over large and complex SVs. However, recent advancements in long-read sequencing technologies now offer the potential of a cheap and powerful method for understanding complex genomic alterations. While sequencing technologies have made a leap forward, work is still needed on the computational side to fully use the technology.

This is an interdisciplinary project. The project concerns algorithm design, implementations of algorithms, and simulated and biological data analysis. The student is expected to learn a bit of relevant molecular biology to efficiently interact in the interdisciplinary project. We seek candidates with a strong computer science, mathematics, statistics, or bioinformatics background and strong programming skills. Some previous knowledge or experience working with biological sequencing data is beneficial but not necessary.

Feel free to email me if you have any question when preparing your application.

About the DDLS PhD student program
Data-driven life science (DDLS) uses data, computational methods and artificial intelligence to study biological systems and processes at all levels, from molecular structures and cellular processes to human health and global ecosystems. The SciLifeLab and Wallenberg National Program for Data-Driven Life Science (DDLS) aims to recruit and train the next generation of data-driven life scientists and to create globally leading computational and data science capabilities in Sweden.

The program is funded with a total of 3.7 billion SEK (about 290 MUSD) over 12 years from the Knut and Alice Wallenberg (KAW) Foundation.

In 2025 the DDLS Research School will be expanded with the recruitment of 19 academic and 7 industrial PhD students. During the course of the DDLS program more than 260 PhD students and 200 postdocs will be part of the Research School. The DDLS program has four strategic research areas: cell and molecular biology, evolution and biodiversity, precision medicine and diagnostics, epidemiology and biology of infection. For more information, please see https://www.scilifelab.se/data-driven/ddls-research-school/

The future of life science is data-driven. Will you be part of that change? Then join us in this unique program!

Keywords: Bioinformatics, computational biology, algorithms, data structures, sequencing data

— Project 2 —

Explainable AI: Characterizing fidelity and creativity of deep generative models

Main supervisor: Chun-Biu Li (cbli@math.su.se)

Deep diffusion models (DDMs) have been one of the main drivers of AI developments in the last few years. They are widely used in generating new realistic data such as in large language models (e.g. ChatGPT), text-to-image (e.g. Midjourney, DALL-E), text-to-video (e.g, Sora, Runway), 3D protein structure prediction (e.g., AlphaFold3), material designs (e.g. MatterGen), etc. However, the underlying generative mechanisms of these models remain elusive.

This project aims to establish deep understanding of the generative processes in DDMs in terms of advances in explainable AI, such as variable explainability, representation learning and interpretation, high dimensional flow dynamics, counterfactual analyses, etc. Focuses will be put on elucidating the tradeoff mechanism between fidelity (how realistic is the generated data) and creativity (how novel is the generative data) in different types of DDMs. The outcomes are expected to provide intuitive guidelines to control, improve and generalize contemporary generative AI models.

In addition to theoretical and computational studies, the project could also be carried out in collaboration with other biostatistics groups in the Karolinska Institutes who will provide us with real biomedical training samples. Furthermore, the Mathematics Department of Stockholm University provides a full range of graduate courses in machine learning/AI that enable admitted PhD students to build up firm statistical and computational backgrounds of the field.

Keywords: Deep generative model, explainable AI, deep neural networks, fidelity and creativity

— Project 3 —

Efficient algorithms for modeling ice sheets

Main supervisor: Josefin Ahlkrona (ahlkrona@math.su.se)

In this PhD project, you will develop mathematical and computational tools to simulate ice sheets more efficiently.

Understanding the effects of climate change — such as future sea level rise and shifts in ocean circulation — requires improved models of how ice on Greenland and Antarctica deforms. At the core of an ice sheet model is a solver for a nonlinear version of the Stokes equations, coupled with an advection equation that describes changes in ice sheet shape. Current models are computationally inefficient, limiting their resolution and accuracy. This project aims to develop more efficient algorithms based on the finite element method.

We seek candidates with a strong background in scientific computing, numerical methods, fluid dynamics, and programming. The work will involve algorithm development, programming, and theoretical analysis.

— Project 4 —

Discrete mathematics & algorithm design

Main supervisor: Marc Hellmuth (marc.hellmuth@math.su.se)

The research group of Marc Hellmuth (https://marc-hellmuth.github.io) at the Department of Mathematics (Division of Computational Mathematics), Stockholm University, is inviting applications for a PhD position in Discrete Mathematics and Theoretical Computer Science.

This PhD project aims to provide a deeper understanding of discrete structures in life sciences, particularly those observed in biology and chemistry. The focus will be on the design of efficient algorithms for recognizing such structures and analyzing the computational complexity of modification problems related to them. This, in turn, will provide the foundation for establishing exact algorithms or efficient heuristics to tackle underlying optimization problems.

We seek candidates with strong background in
• Discrete Mathematics, in particular, Graph Theory and Discrete Optimization and
• Theoretical Computer Science, in particular, in Algorithm Design and Computational Complexity

(Optional) Knowledge in Bioinformatics is beneficial but not required.
For any further inquire, please contact Marc Hellmuth at marc.hellmuth@math.su.se.

We look forward to receiving your application!

— Project 5 —

Programming, logic and constructive mathematics

Main supervisor: Anders Mörtberg (anders.mortberg@math.su.se)

My research is in the intersection of programming, logic and constructive mathematics. Specific topics that interest me, with some keywords in parentheses, are:

- logic (dependent type theory, proof automation, proof verification, proof assistants),
- constructive mathematics (commutative algebra, computer algebra, algebraic geometry and topology),
- category theory (semantics of type theories, higher category theory),
- functional programming (connections to category theory and logic, implementation of proof assistants),
- computer formalization of mathematics and computer science (using proof assistants like Agda, Coq, Lean), and
- homotopy type theory (HoTT, univalent foundations, higher inductive types, cubical type theory, Cubical Agda, synthetic homotopy theory).

Interest in one or more of the above topics can be a starting point for a possible PhD thesis topic. My webpage contains recent papers and other relevant information about my research: https://staff.math.su.se/anders.mortberg/

Some more specific potential thesis topics with me are:

- Formalize mathematics and computer science in a proof assistant based on type theory, e.g. in Cubical Agda. Possible topics to formalize include synthetic homotopy theory, synthetic algebraic geometry, algebra or data structures using the structure identity principle of HoTT, topological data analysis in HoTT, etc.
- Develop new type theories and study their properties. This includes both implementation and proving meta-theoretic properties (canonicity, normalization, etc.) as well as general model theoretic considerations.
- Develop models of type theory in higher categories and/or using methods from categorical logic.

Feel free to email me if you want to discuss some potential topics or have any question when preparing your application.