Lea Popovic (10-10:45 Monday)

Multi-scale Markov process for interacting particles diffusing in spatially heterogeneous system

Measure-valued Markov processes are used to model the dynamics of spatial systems of interacting species (for example branching processes, population genetics, interacting particle systems, SPDEs). Typically its values are either discrete or continuous measures on a spatial domain. Our focus is on models with mixed (multi-scale) values, some coordinates discrete and some continuous, coupled in their dynamics.
We derive such multi-scale process as a limit of particles in a spatially heterogeneous system interacting through reaction dynamics and undergoing independent movement in space. The multi-scale nature is due to the differing abundances between species types. The resulting limiting dynamics is a measure-valued piecewise Markov process in which the discrete mass of the process evolves stochastically, and the continuous mass evolves in either a deterministic way between consecutive jump times of the discrete part, or as a diffusion between the jumps of the discrete part. Long-term behaviour of such processes will be explored.

Tai Melcher (11:15-12 Monday)

Infinite-dimensional diffusions under Hormander's condition

A collection of vector fields on a manifold M is said to satisfy the Hormander condition if they generate the span of the tangent space at all points of M. Such a collection has a naturally associated diffusion with generator the sum of squares of the vector fields, and often induces a natural geometry on M via the horizontal distance. In finite dimensions, the associated diffusions are known to enjoy various regularity properties, and it's known that the topology induced by the horizontal distance is equivalent to the intrinsic (locally Euclidean) one. The situation in infinite dimensions is more complicated and less understood.

Finite-dimensional Heisenberg groups are model spaces for studying properties of distances and diffusions associated to Hormander collections of vector fields. We'll discuss some things we've learned from studying infinite-dimensional Heisenberg groups. This includes an analytic condition which is equivalent to the Hormander condition in this setting. Under this condition, we may prove regularity properties of the associated diffusion but can show that the horizontal topology may no longer be equivalent. This construction also allows us to discuss the distribution of the central element of the diffusion, which is an infinite-dimensional stochastic Levy area.

Ivana Bozic (2-2:45 Monday)

Quantifying the evolutionary dynamics of cancer.

https://www.ivanabozic.com/publications

Samantha Petti (3-3:45 Monday)

Learning functions in biological sequence space

A fundamental goal of genetics is to understand how variation in biological sequences gives rise to differences in measurable characteristics called phenotypes. The mapping from genotype (DNA, RNA or protein sequence) to phenotype can be difficult to predict and interpret because combinations of mutations interact in complex ways. We explore the use of Gaussian process regression to learn genotype-phenotype maps, which to a mathematician are real-valued functions over a discrete space of sequences with a fixed length and alphabet. We propose families of prior distributions that are both tractable and expressive. We demonstrate that Gaussian process regression with our priors yields better predictive performance on several datasets and show how the learned prior can be interpreted to provide clues into how phenotype arises from genotype biologically. Finally, we discuss connections between regularized regression on sequence space and MAP estimates under particular priors. This is joint work with Carlos Marti-Gomez, David McCandlish, and Juannan Zhou.

Jasmine Foo (4:15-5 Monday)

Spatial Models of Cancer Evolution

https://www-users.cse.umn.edu/~jyfoo/publications.html

Tamara Broderick (10-10:45 Tuesday)

Double trouble: Predicting new variant counts across two heterogeneous populations

Collecting genomics data across multiple heterogeneous populations (e.g., across different cancer types) has the potential to improve our understanding of disease. Despite sequencing advances, though, resources often remain a constraint when gathering data. So it would be useful for experimental design if experimenters with access to a pilot study could predict the number of new genetic variants they might expect to find in a follow-up study: both the number of new variants shared between the populations and the total across the populations. While many authors have developed prediction methods for the single-population case, we show that these predictions can fare poorly across multiple populations that are heterogeneous. A state-of-the-art single-population predictor built on the beta-Bernoulli process offers a natural extension to multiple populations. But we prove that this extension fails for fundamental reasons. We use a similar Poisson point process framework to provide the first predictor for the number of new shared variants and new total variants that can handle heterogeneity in multiple populations. We show that our proposed method works well empirically using real cancer and population genetics data.

Dana Randall (11:15-12 Tuesday)

Programmable Matter and Emergent Computation

Programmable matter explores how collections of computationally limited agents acting locally and asynchronously can achieve some useful coordinated behavior. We take a stochastic approach using techniques from randomized algorithms and statistical physics to develop distributed algorithms for emergent collective behaviors that give guarantees and are robust to failures.