Tuesday, May 11, 2021
Qualifying Exam for Ali Heydari

Title:

A Tale of Two Scales: Deep Learning and Mathematical Frameworks for Complex Biological Systems

Time: 11 am PDT

Zoom Link: https://ucmerced.zoom.us/j/8324815313

Abstract: In my qualifying examination, I will present my on-going work in two mathematical modeling frameworks for studying complex biological systems at the single-cell scale. First, I will present an Eulerian conservative finite volume formulation for diffusion-reaction systems defined over arbitrarily deforming geometries. Biological systems, such as cells, often exhibit moving and deforming geometries while maintaining the conservation of mass within the system. In this project I applied this approach to modeling the reaction and diffusion of protein aggregates within a 3D actively dividing yeast cell. The critical idea of this novel approach is to construct the variational formulation by integrating over the spatio-temporal volumes instead of integrating the time-discretized equations in space (as it is common practice). I will demonstrate the convergence of this method with both theoretical analysis and computational exploration. I will also highlight how the traditional approach can lead to slow or non-converging solutions.

In my second project, I developed a novel deep learning framework for generating realistic synthetic single-cell RNA sequencing (scRNAseq) data. ScRNAseq technologies allow for measurements of gene expression at a single-cell resolution. These technologies provide researchers with a tremendous advantage for detecting heterogeneity, delineating cellular maps, or identifying rare subpopulations. However, a critical complication to the use of scRNAseq data is the low number of single-cell observations due to limitations by the rarity of a subpopulation, tissue degradation, or cost. I will present our state-of-the-art deep generative model for generating realistic in-silico single-cell RNA-seq data, called ACTIVA (Automated Cell-Type-informed Introspective Variational Autoencoder). Within a single framework, ACTIVA can generate data representative of the entire population, or specific subpopulations on demand, as opposed to two separate models (such as scGAN and cscGAN). Data generation and augmentation with ACTIVA can enhance scRNAseq pipelines and analysis, such as benchmarking new algorithms, studying the accuracy of classifiers, and detecting marker genes. ACTIVA will facilitate analysis of smaller datasets, potentially reducing the number of patients and animals necessary in initial studies.

PhD Committee:
Suzanne S. Sindi (Chair) ssindi@ucmerced.edu,
Erica Rutter,
Maxime Theillard