



The Hong Kong University of Science and Technology

Department of Mathematics

PhD THESIS EXAMINATION

**Deep Generative Model Learning and 3D Construction of
Spatial Transcriptomics with Deep Neural Networks**

By

Mr. Gefei WANG

ABSTRACT

Deep generative models have shown extraordinary performance in a diverse range of applications over recent years. However, despite the empirical success of deep generative models, there still exist challenges in learning complex data distributions and providing theoretical justification of these methods. Besides deep generative learning, another important task in data science is to use machine learning methods to model and analyze data in cross-discipline fields. One prominent example is spatial transcriptomics (ST) data, which characterize the transcriptomic landscape on biological tissue samples. Although many efforts have been made to analyze ST data, unique challenges exist in building effective models to leverage rich information shared across a collection of sample slices and additional information such as single-cell transcriptomics data to decipher biological structures.

In this thesis, we propose two deep learning approaches to address challenges in generative learning and spatial transcriptomics data analysis, respectively. To address challenges in generative learning of complex distributions, we propose to learn a generative model with a Schrödinger Bridge. We characterize the generative learning task as stochastic differential equations with time-varying drift terms, and derive our Schrödinger Bridge algorithm by plugging the drift term estimated by a deep neural networks. Under some mild assumptions, we establish the consistency of our approach. Experimental results indicate that the generative model via Schrödinger Bridge is comparable with state-of-the-art generative models, suggesting a new formulation of generative learning. To overcome the difficulties in ST data analysis, we present STitch3D, a unified deep learning-based framework that integrates multiple 2D tissue slices to reconstruct 3D cellular structures from the tissue level to the whole organism level. By jointly modeling multiple 2D tissue slices and integrating them with cell-type-specific expression profiles derived from single-cell RNA-sequencing data, STitch3D simultaneously identifies 3D spatial regions with coherent gene expression levels and reveals 3D distributions of cell types. Through comprehensive experiments using diverse datasets, we demonstrate the performance of STitch3D in building comprehensive 3D tissue architectures.

Date : 19 July 2023, Wednesday

Time : 3:30 p.m.

Venue : Room 4472 (Lift 25/26)

Zoom ID: 934 4970 6287 (passcode: 501425) ~ EE opted via online mode.

<https://hkust.zoom.us/j/93449706287>

Thesis Examination Committee:

Chairman : Prof. Ho Yi MAK, LIFS/HKUST

Thesis Supervisor : Prof. Can YANG, MATH/HKUST

Thesis Supervisor : Prof. Yang WANG, MATH/HKUST

Member : Prof. Dong XIA, MATH/HKUST

Member : Prof. Xinzhou GUO, MATH/HKUST

Member : Prof. Angela Ruohao WU, LIFS/HKUST

External Examiner : Prof. Shuqin ZHANG,
School of Mathematical Sciences/ Fudan University

(Open to all faculty and students)

The student's thesis is now being displayed on the reception counter in the General Administration Office (Room 3461).