

Inferring single-cell trajectories from RNA velocity fields with varifold distances

Elodie Maignant
Technical University of Denmark (DTU)
`elodie.maignant@orange.fr`

In this talk, we address the critical problem in single-cell transcriptomics of reconstructing dynamic cellular processes from sequencing data. In particular, given a population of cells sampled from such a process, we are interested in the problem of ordering the cells according to their progression in the process. This is known as trajectory inference. If the process is differentiation, this amounts to reconstructing the corresponding differentiation tree. One way of doing this in practice is to estimate the path distance between nodes – labeled by cells – based on cell similarities observed in the sequencing data. Recent sequencing techniques make it possible to measure two types of data: gene expression levels, and RNA velocity, a vector that predicts changes in gene expression. The data then consist of a discrete vector field on a Euclidean space of dimension equal to the number of genes under consideration. By integrating this velocity field backward, we trace the evolution of gene expression levels in each single cell from some initial state to its current state. Using varifold distances between the individual trajectories thus constructed, we introduce a new dissimilarity measure for inferring the differentiation tree. Under the assumption that the sequencing data describe a faithful embedding of such a tree, we prove that the dissimilarity measure we propose approximates the shortest-path distance in a tree that is isomorphic to the target tree. We apply the newly introduced dissimilarity measure to several simulated examples from the **dyngen** library and we show that we are able to recover the corresponding differentiation tree with high accuracy.

Joint work with: Tim Conrad, Christoph von Tycowicz.