Statistical tools for TDA via Optimal Transport

Theo Lacombe
Datashape, Inria
theo.lacombe@inria.fr

Persistence diagrams (PDs) are now routinely used to provide succinct descriptors encoding the underlying topology of sophisticated data encountered in challenging learning problems. Despite several appealing properties, integrating PDs in statistical and learning pipelines raises difficulties because their base geometry is not Hilbertian. In particular, algorithms to average a family of PDs have only been considered recently and are known to be computationally prohibitive. We present in this talk a tractable framework which links standard PD distances and optimal transport (OT). It allows us to leverage recent advances in OT, in particular entropic regularization and its convolutional formulation on regular grids. It provides efficient, (GPU) scalable, and differentiable computations, making PDs more suited to machine-learning and statistical analysis than before.

Joint work with: Marco Cuturi (CREST, Ensae), Steve Oudot (Datashape, Inria).