A data set is often given as a point cloud, i.e. a non-empty finite metric space. An important problem is to detect the topological shape of data — for example, to approximate a point cloud by a low-dimensional non-linear subspace such as a graph or a simplicial complex. Classical clustering methods and principal component analysis work very well when data points split into well-separated groups or lie near linear subspaces.

Methods from topological data analysis detect more complicated patterns such as holes and voids that persist for a long time in a 1-parameter family of shapes associated to a point cloud. These features were recently visualized in the form of a 1-dimensional homologically persistent skeleton, which optimally extends a minimal spanning tree of a point cloud to a graph with cycles. I will talk about a generalization of this 1-skeleton to higher dimensions and optimality results that we proved. (Received July 21, 2017)