1145-00-2242 **Mariano Tepper*** (marianotepper@gmail.com), **Victor Minden** and **Anirvan M Sengupta**.
From clusters to manifolds with semidefinite and completely-positive approximations. Preliminary report.

In solving hard computational problems, convex relaxations often play an important role as they come with a guarantee of optimality. Here, we focus on a popular semidefinite relaxation of K-means clustering. In previous work, we reported an unexpected finding: when data contains (multiple) manifolds, the solution captures such geometrical structures.

Ideally, completely positive (CP) formulations provide a tighter relaxation to the original problem than semidefinite programs (SDP). Whereas SDPs are relatively easy to solve, CP programs are very hard, hindering their applicability. An important empirical question then becomes: are these semidefinite output matrices close to be CP? We answer this question in the positive, showing that they are indeed closely approximated by a sub-family of CP matrices, the RBF kernel matrix, which can be easily obtained with a convex program. Remarkably, these special CP matrices give rise to the natural low-dimensional embeddings of the input manifolds.

For each sub-problem, we present new, convex and efficient, algorithms based on the conditional gradient method. Our results render this SDP-CP hybrid algorithm a versatile, understandable, and powerful tool for manifold learning. (Received September 25, 2018)