

Hierarchical clustering algorithms are common tools for simplifying, exploring and analyzing datasets in many areas of research. For flow cytometry, a specific variant of agglomerative clustering has been proposed, that uses cluster linkage based on Mahalanobis distance to produce results better suited for the domain. Applicability of this clustering algorithm is currently limited by its relatively high computational complexity, which does not allow it to scale to common cytometry datasets. This thesis describes a specialized, GPU-accelerated version of the Mahalanobis-average linked hierarchical clustering, which improves the algorithm performance by several orders of magnitude, thus allowing it to scale to much larger datasets. The thesis provides an overview of current hierarchical clustering algorithms, and details the construction of the variant used on GPU. The result is benchmarked on publicly available high-dimensional data from mass cytometry.