



## **Inference for hierarchical clustering of variables**

Max G'Sell\*

Carnegie Mellon University, United States [-mgsell@cmu.edu](mailto:mgsell@cmu.edu)

Hierarchical clustering is often applied to detect dependence structure among the variables in large data sets. While this gives some scientific intuition, it has traditionally been difficult to obtain statistical error guarantees for the resulting clusters. In this problem, we present approaches to obtaining inferential guarantees about the groups of variables identified by these methods. We present closed-form asymptotic results in an idealized setting, along with more general permutation-based approaches. We also introduce connections to the graphical lasso and to recent work in sequential false discovery rate control.