Contribution ID: **20**   Type: **not specified**

Near optimal efficient decoding from pooled data

Wednesday 28 September 2022 09:20 (20 minutes)

Consider n items, each of which is characterized by one of d+1 possible features in {0,...,d}. We study the inference task of learning these types by queries on subsets, or pools, of the items that only reveal a form of coarsened information on the features - in our case, the sum of all the features in the pool. Related prominent problems are the quantitative group testing problem, of which it is a generalization, as well as the compressed sensing problem, of which it is a special case.

We are interested in the minimum number of queries needed to efficiently infer the features, in the setting where the feature vector is chosen uniformly while fixing the frequencies, and one of the features, say 0, is dominant in the sense that the number $k = n\,\theta$, $\theta \in (0, 1)$, of non-zero features among the items is much smaller than n. It is known that in this case, all features can be recovered in exponential time using no more than $O(k)$ queries. However, so far, all efficient inference algorithms required at least $\Omega(k \ln)$ queries, and it was unknown whether this gap is artificial or of a fundamental nature. We provide an efficient algorithm that succeeds with high probability and employs no more than $O(k)$ measurements. This also solves a prominent open question for the quantitative group testing problem. This is joint work with Max Hahn-Klimroth.

Presenter: MUELLER, Noela (Eindhoven University of Technology)