Inside the clustering threshold for random linear equations

Pu Gao

University of Toronto

We study a random system of cn linear equations over n variables in $\mathrm{GF}(2)$, where each equation contains exactly r variables. Previous results determined the clustering threshold c_r^*: if $c > c_r^*$, the solution space is partitioned into well-separated clusters whereas if $c < c_r^*$, all solutions lie inside one well-connected cluster. This clustering phenomenon is hypothesized in many random constraint satisfaction problems and explains why it is difficult to design algorithms for general CSP's. In this talk, we discuss the solution space of the linear system when $c = c_r^* + n^{-\delta}, \delta > 0$. This is joint work with Mike Molloy.