Project proposal: Reconstruction in random factor graphs
Ricardo Restrepo
Mentor: Prasad Tetali

Introduction: A factor graph is a graphical representation of a system involving constraints (or interactions) between its variables. It consists in a bipartite graph connecting variable nodes $\{i : i = 1, \ldots, n\}$, with function nodes $\{a : a = 1, \ldots, m\}$. To every function node i we assign a $\chi-$valued variable x_i and to every function node a, of degree k_a, we assign a nonnegative compatibility function $\psi_a : \chi^{k_a} \to \mathbf{R}$. This configuration defines a distribution over χ^n, given by

$$\mu(x_1, \ldots, x_n) \sim \prod_{a=1}^{m} \psi_a(x_{\partial a}),$$

where $x_{\partial a}$ are the variables assigned to the neighbors of the function node a, in certain preassigned order. Examples of factor graph models are graph colorings, $k-$SAT, general CSP's, etc..

The reconstruction problem in a factor graph consists in the estimation of the value of a variable x_i, given 'far away' observations, say the values of the variables in the boundary of certain neighbourhood $B(i,t)$ of the node i. To determine if this estimation is possible, we need to know the 'degree of dependence' of the variables x_i and $x_{\partial B(i,t)}$, which we measure as

$$\left\|\mu_{i,\partial B(i,t)}(\cdot,\cdot) - \mu_i(\cdot)\,\mu_{\partial B(i,t)}(\cdot)\right\|_{TV}. \tag{1}$$

If this quantity is bounded away from zero it means that for some values of $x_{\partial B(i,t)}$, there is a bias on what should be the value of x_i, which should be therefore a good guess (or guesses) for x_i. The reconstruction problem emerges naturally in a variety of contexts like MCMC algorithms, sampling in random CSPs, existence of approximate fixed points in message passing algorithms, philogeny, network tomography and extremality of the Gibbs measure.

Project: Given an increasing sequence of factor graphs $\{G_N\}_{N\geq 1}$ (that may be random), the reconstruction problem is solvable if the quantity in (**??**) stays bounded away from zero when $N \to \infty$ and $t \to \infty$ (in that order). If this sequence of factor graphs is locally tree-like (locally converges to a tree when $N \to \infty$), the solvability of the reconstruction problem is linked with the solvability of the reconstruction problem in the corresponding limiting tree. In fact, Montanari and Gerschenfeld[1] , established a sufficient condition for these two problems to coincide in graphical models (a particular kind of factor graph). This condition allowed them to conclude, for example, that for antiferromagnetic $q-$colorings in a random Poisson graph with parameter γ, where

$$\gamma < (q-1)\log(q-1), \tag{2}$$

the reconstruction problem is solvable if and only if it is solvable for the corresponding limiting tree (A Galton Watson tree with Poisson distributed degrees of mean 2γ.)

[1] *Reconstruction for models on random graphs.* FOCS 2007

The objective of this project is to extend these ideas to the case of locally tree-like random factor graphs and to establish similar reconstruction thresholds for several factor graph models including, but not limited to: graph coloring, hypergraph coloring, $NAE-$SAT and extensions of it (for instance, 'at most k-equal'-SAT). Just as in the case of graphical models, the crux of the problem is not to state the conditions under which factor graph reconstruction will coincide with the corresponding factor tree reconstruction, but to check these conditions and to establish the corresponding thresholds for tree reconstruction. These will be our goals.