A NEW CHAPTER:

PROBABILISTIC ANALYSIS

Until now we looked at graphs with bounded degree.

In many applications it is interesting to consider RANDOM GRAPHS.

→ Two motivating examples:

1. RANDOM k-SATISFIABILITY: (k-sat)
(Franco et al., Selman et al.,...)

\[ x_i \in \{0,1\} = \{\text{FALSE, TRUE}\} \quad \mathbf{x} = (x_1, x_n) \]

k-SAT FORMULA (here k=3)
\[ F = (x_4 \lor x_5 \lor \overline{x}_3) \land (x_5 \lor x_3 \lor \overline{x}_2) \land \ldots \land (x_8 \lor x_9 \lor \overline{x}_{12}) \]

\[ \text{logic OR} \quad \text{logic AND} \quad \text{negative} \]

\[ F \text{ is satisfiable if there exists an assignment such that...} \]
**RANDOM k-SAT**: Pick a formula \( \Phi \) with \( m \) clauses and \( n \) variables uniformly at random. ["Clause density, \( \alpha = \frac{m}{n} \)]

**Equivalently**: For each \( q=1, \ldots, m \), construct the \( q \)-th clause by drawing \( k \) variable indices \( i_1, \ldots, i_k \) uniformly at random in \( \{1, \ldots, n\} \) and including \( x_{i_q} \) or \( \overline{x_{i_q}} \) independently with prob \( \frac{1}{2} \).

To a formula \( F \) we can associate a factor graph \( G = (V, F, E) \)

\[
(x_1 \lor \overline{x_2} \lor \overline{x_2}) \land (x_2 \lor x_5 \lor x_6) \land (x_3 \lor x_5 \lor x_4)
\]

Is this formula SAT?
A "natural" measure associated to $F$ is the uniform over solutions

$$
\mu_g(x) = \frac{1}{Z_g} \prod_{\text{all } a \text{ satisfies clause } a} 
$$

\[ \rightarrow \text{ FACTOR GRAPH } G \]

If $F$ is random, $G$ (and $\mu(.)$) is random as well.

Luby, Mitzenmacher, Shokrollahi, Spielman, Sudan (1997) introduced the following ensemble of irregular LDPC codes

- Ensemble parameters
  - $n \rightarrow \# \text{ variable nodes}$
  - $m \rightarrow \# \text{ function nodes}$
  - $\Lambda = \{ \lambda_e : e \geq 0 \}$ \[ \rightarrow \text{ distribution of } r \text{-nodes degrees} \]
  - $P = \{ P_k : k \geq 0 \}$ \[ \rightarrow \text{ distr. of } f \text{-nodes degree} \]

(rho)
\[ n, m \text{ integers} \]
\[ \Lambda \geq 0 \sum_{e} \Lambda e = 1 \quad \text{and} \quad P_k \geq 0 \sum_{k} P_k = 1 \]
\[ n \sum_{e} \Lambda e = m \sum_{k} P_k \quad (*) \]

→ Draw \( n \) variable nodes and \( m \) function nodes

\[ 1 \quad 2 \quad 3 \quad \cdots \quad n \]

\[ a \quad b \quad m \]

→ Associate \( a \) half-edges to the first \( n \Lambda \) \( v \)-nodes

1. half-edge to the following \( n \Lambda_1 \) \( v \)-nodes

2. half edges

\[ a \] half-edge to the first \( n P_1 \) \( f \)-nodes

2. half edges to the following \( m P_2 \) \( f \)-nodes
\[ \mu(x) = \frac{1}{Z} \prod_{i \in F} \mathbb{I}(x_{i|a}) \prod_{i \in V} \mathbb{I}(x_{i|a}) = 0 \] 

\[ \mathbb{P}\{X=x|Y=y\} \]
Two ensembles of factor graphs

\[ C_n(k, \alpha) \leftarrow \text{as in the k-SAT example} \]

\[ D_n(n, p) \leftarrow \text{as in the coding example} \]

[Always assume a bounded \( \ell_{\text{max}}, k_{\text{max}} \) bounded \( n \to \infty \).

Basic Properties [in Random Graph Theory]

1. **Degree Distribution**

   \[ C_n(k, \alpha) \rightarrow \text{regular of degree } k \text{ at } f\text{-nodes} \]

   \[ \rightarrow \text{Poisson } (k\alpha) \text{ degree at } f\text{-nodes} \]

   \[ D_n(n, p) \rightarrow \text{Degree distribution \{Lef at } \]

   \[ v\text{-nodes} \rightarrow \{P_k\} \text{ at } f\text{-nodes} \]

2. **Edge Perspective Degree Distribution**

   Pick a uniformly random edge; what is the prob. that the adjacent
v-node (f-node) has degree $\Theta(k^2)$?

$G_n(k, \alpha) \rightarrow v$-node: Poisson($k\alpha$).

f-node: always k-1

$D_n(\lambda, \rho) \rightarrow v$-node: $\lambda_e = \frac{\lambda \rho}{\sum \lambda \rho}$

f-node: $\gamma_k = \frac{\rho P_k}{\sum_k k \cdot P_k}$

"Proof"

There are $n \cdot \lambda_e$ edges adjacent to a $v$-node of degree $\lambda$. Normalizing by the total number of edges we get $\lambda_e$

3. SMALL LOOPS

Pick a uniformly random vertex $i$.

Then the shortest loop through $i$ has, with length $\Theta(\log n)$. 

Proof: Imagine growing the neighborhood of a vertex $v$-node $i$.

$e_1, e_2, e_3$ are approx iid in $[m]$.

$i_1, i_2, i_3$ are approx iid in $[n]$. (In fact, one cannot have $i_1 = i_2, i_3 = i_4, i_5 = i_6$ or $i_e = i$, but ...)

$\Rightarrow$ why they are distinct.

We assume, for sake of simplicity $D_n(N, P)$ with $\Lambda = 1, P = 1$. 
At distance $t$ we choose $\Theta(b^t)$ new nodes. The probability that two of them coincide start to be non-negligible when

$$\frac{(b^t)^2}{n} = 1 \Rightarrow t = \Theta(\log n)$$

The same proof shows the following: 

**Proposition**

Let $B_i(t)$ be the radius $t$ neighborhood around node $i$ in $G = \mathcal{D}_n(\Lambda, \mathcal{P})$ where $\Lambda_e = 1$, $\mathcal{P}_k = 1$ (regular graph). Then $B_i(t)$ is whp an $e \times k$ regular tree of $t$-generations.
What about "irregular" ensembles?

Define the \textsc{treeensemble} $T_t(V,P)$ as follows:

\begin{align*}
\text{Proposition} \quad \text{Let } B_i(t) \text{ be the radius of neighborhood around a uniformly random vertex } i \text{ in } G \overset{d}{=} D_n(V,P).
\end{align*}

Then, $B_i(t) \overset{d}{\rightarrow} T_t(V,P)$ as $n \rightarrow \infty$.

What is the analogous statement (and tree model) for $G_{\lambda}(k,\alpha)$?