(Continuing from the previous session)

We are trying to prove the following:

**Lemma 1.** For any arbitrary \( \alpha \in (\alpha_d, \alpha_c) \), where \( \alpha_d \) is the 2-core percolation and \( \alpha_c \) will be defined later, the following happens with high probability:

There exists a partition of the hypercube \( \{0,1\}^n \) into sets \( \Omega_1, \Omega_2, \cdots, \Omega_N \) such that

- \( N = e^{n(\Sigma \pm \epsilon)} \), for some \( \Sigma > 0 \) which we call "complexity".
- \( \mu(x) = \sum_{l=1}^{N} w_l \mu_l(x) \), with \( \mu_l(.) \) supported on \( \Omega_l \).
- \( w_l = e^{-n(\Sigma \pm \epsilon)} \).
- \( \frac{\mu(\partial \epsilon \Omega_l)}{\mu(\Omega_l)} \left[ 1 - \mu(\Omega_l) \right] \leq e^{-nC(\epsilon)} \), for small enough \( \epsilon \) and where \( C > 0 \). Also \( \partial \epsilon \Omega_l \equiv \{ x : 0 < d(x, \Omega_l) \leq n\epsilon \} \).

See figure 1.

![Figure 1: \( \partial \epsilon \Omega_l \equiv \{ x : 0 < d(x, \Omega_l) \leq n\epsilon \} \)](image)

- Each of the \( \mu_l(.) \) cannot be further decomposed.

Physicists usually refer to the above partitioning as "Pure States Decomposition".

We will discuss some points before going through the proof of the above lemma. The first point is that why this is important.

- If \( \alpha > \alpha_d \), then there is no correlation decay for \( \mu \). Hence, difficulties arise for BP.
This correlation originally appears because far apart variables know that the whole system is in a certain state \( l \). This is usually referred as “Proliferation” of pure states.

If we restrict to a single state, then for each of them we have correlation decay and consequently BP works. Now if we can put all this information together, we may hope to do something.

We need a second ingredient to be used together with the core size computation. Consider factor graph \( G \) from \( G_n(\Lambda, P) \) in which the degrees of variable (resp. function) nodes come from the distribution \( \Lambda \) (resp. \( P \)). Also, let \( \tilde{H}_{\tilde{m} \times \tilde{n}} \) denote the adjacency matrix of \( G \) with \( \tilde{m} \) function nodes and \( \tilde{n} \) variable nodes.

We first state the following lemma.

**Lemma 2.** Let \( Z_n(w) \) be the cardinality of the set \( \{ x : w(x) = w; \tilde{H}x = 0 \} \). Recall that \( w(x) \) is the “weight” of \( x \). We have:

\[
\mathbb{E}Z_n(w) = \sum_{l=0}^E \frac{l!(E-l)!}{E!} \times \text{coeff} \left[ \prod_{l=1}^{l_{\max}} (1 + xy_l^{\tilde{n}_i}) x^{w_i} y_l^{\tilde{n}_i} \right] \times \text{coeff} \left[ \prod_{k=2}^{k_{\max}} q_k(Z)^{\tilde{n}_k} Z^l \right].
\]

where \( E = \tilde{n} \sum_l \Lambda_l = \tilde{m} \sum_k P_k k = \# \text{ of the edges.} \)

**Proof** [of lemma 2] The probability of any factor graph is \(( E! )^{-1}\). Therefore,

\[
\mathbb{E}Z_n(w) = \frac{1}{E!} \# \{ \text{graph/solution pairs} \}.
\]

A graph/solution pair can be identified with a valid ”colored graph”, in which each edge can be colored either RED or BLUE.

A colored graph is ”valid” if:

- All the edges incoming in the same variable node have the same color.
- Each function node has an even number of RED edges.
- There are \( w \) RED variable nodes.

Now, if \( l \) is the number of RED edges in the graph, then:

\[
\text{coeff} \left[ \prod_{l=1}^{l_{\max}} (1 + xy_l^{\tilde{n}_i}) x^{w_i} y_l^{\tilde{n}_i} \right] = \# \text{ ways of coloring half edges on } v\text{-node side, and}
\]

\[
\text{coeff} \left[ \prod_{k=2}^{k_{\max}} q_k(Z)^{\tilde{n}_k} Z^l \right] = \# \text{ ways of coloring half edges on } f\text{-node side.}
\]

After a (boring) calculation, one can find the asymptotic behavior.

By the above, we have

\[
\mathbb{E}Z_n(\tilde{n}w) \sim e^{\phi(w)}.
\]

If \( \Lambda_0 = \Lambda_1 = 0 \), i.e. the minimum variable node degree is at least 2, then \( \phi(w) = Aw + O(w^2) \), where \( A = A(\Lambda, P) \).

Two possible cases can be considered. See figure [2]

Let us apply this to XORSAT problem. We define \( \tilde{H} \) be the restriction of \( H \) to the core \( K_2(G) \); \( \tilde{n} = |K_2(G)| \cong nQ_\ast(\alpha) \) and \( Q_\ast > 0 \) for \( \alpha > \alpha_d \). Here, we can observe three different cases. See figure [2]

2
Figure 2: In the case $A < 0$, any two solutions have hamming distance $\geq n(w_* - \epsilon)$.

Note that conditional on its "degree profile", $K_2(G)$ is distributed essentially as $G_n(\Lambda, P)$. Furthermore, $\Lambda$ and $P$ are exponentially concentrated around typical degree profiles $\Lambda^*$ and $P^*$ where:

$$
P^*_k = 1,
\Lambda^*_{l} = (l^\lambda - 1 - \lambda)^{-1}l!/l^\lambda, \text{ for } l \geq 2,
\Lambda^*_{l} = 0, \text{ for } l \leq 1,
\lambda = k\alpha Q_{k}^{-1};
$$

Thus we could apply the previous calculation.

Hence, for $\alpha_d < \alpha < \alpha_c$, any two solutions of the subsystem have distance at least $\tilde{n}[w_*(\alpha) - \epsilon]$.

Now, let us see what does this mean for the solutions of the original system. As we had mentioned before, some clusters appear. See figure 4. This gives us the desired decomposition.

For BP we have the following: (See figure 5)

$$
\mu_{i \rightarrow a}(x_i) \propto \prod_{b \in \partial i \setminus a} \hat{\nu}_{b \rightarrow i}(x_i)
$$

3
and

$$\nu_{\nu-i}(x_i) \propto \sum_{\{x_j\}} \mathbb{I}(x_i \oplus x_{j(1)} \oplus x_{j(2)} \oplus x_{j(k-1)} = 0)\nu_{j(1)} \rightarrow a(x_{j(1)}) \cdots \nu_{j(k-1)} \rightarrow a(x_{j(k-1)}).$$
Figure 3: Three cases for different $\alpha_c \geq \alpha \geq \alpha_d$. 
Figure 4: The $\Omega_l$ gives us the desired decomposition.
Figure 5: Belief Propagation.