Statistics 291: Lecture 17 (March 26, 2024)

Multi-OGP

Instructor: Mark Sellke

Scribe: Sabarish Sainathan

1 Review of problem 2.4 on Homework 3

Recall that this problem asks us to show that measures that admit shattering decompositions have exponentially long mixing times. Precisely, suppose μ_β shatters into $\mathscr{C}_1, \ldots, \mathscr{C}_m$ with radii $r\sqrt{N}$ and separation $s\sqrt{N}$.

Here's a sketch of how one might show this:

(a) Sample $X_0 \sim \mu_\beta$. With probability at least $1 - e^{-cN}$ for some $c > 0$, $X_0 \in \mathscr{C}_j$ for some j.

(b) Let X_t be a Langevin dynamics started at X_0. Define the stopping time $\tau = \inf\{t : d(X_t, \mathscr{C}_j) > s\sqrt{N}/3\}$, where $r\sqrt{N}$ is the cluster diameter. We intend to show that $\mathbb{P}[\tau \geq e^{cN}] \geq 1 - e^{-cN}$.

(c) By problem 2.2, X_t does not move substantially during small intervals: for some $\delta > 0$, $\mathbb{P}[\|x_t - x_\tau\| \leq s\sqrt{N}/10 \ \forall t \in [\tau, \tau + \delta]] \geq 1 - e^{-N} \geq \frac{1}{2}$ for large N.

(d) This implies that conditionally on the event $\tau < e^{cN}$, the expected time that X_t is not in cluster C_j durng the interval $[0, e^{cN}]$ is at least $\delta/2$. By the separation property, X_t is also distance at least $s\sqrt{N}/10$ from each of the other clusters during the interval $[\tau, \tau + \delta]$. However since X_0 is started from the stationary distribution, we have $X_t \sim \mu_\beta$ for all fixed t, so the expected time spent outside any cluster on $t \in [0, T]$ should be at most $e^{-10c}T$. This implies that $\mathbb{P}[\tau \geq e^{cN}] \geq 1 - e^{-cN}$ completing the proof.

2 Overview

As in the last class, we'll continue studying the maximum independent set problem in the Erdös-Rényi (ER) graph $G(n, d/n)$ in the setting $1 \ll d \ll n$.

We previously proved the upper bound in the following theorem of Frieze:

Theorem 2.1. *$|MaxIndSet(G(n, d/n))| = (2 \pm o(1))\Phi$ with high probability, where $\Phi = n \log d/d$.*

We further showed a computational hardness result illustrated by the following theorem — namely, that a class of local algorithms do not find maximum independent sets.

Theorem 2.2 ([4])**.** *If $\mathscr{A}$ is an R-local algorithm to compute independent sets, then*

$$\mathbb{E}[|\mathscr{A}(G(n, d/n))|] \leq (1 + 1/\sqrt{2} + o(1))\Phi$$

Our goal for today is the following improvement.

Theorem 2.3 ([5])**.** $\mathbb{E}|\mathscr{A}(G(n,d/n))| \le (1+o(1))\Phi$.

Recall our proof strategy for the first of these theorems. We showed (1) that the expected number of independent sets with overlap within a small multiplicative interval of Φ decays exponentially, implying by Markov's inequality that such independent sets are exponentially unlikely. We also showed (2) that R-local algorithms asymptotically very likely generate independent sets with overlap within this interval by constructing pairs of graphs with a particular correlation structure for the ω_v random variables defined in the last lecture. (1) and (2) generate a contradiction.

This proof strategy uses the overlap gap property. The tighter upper bound of [5] leverages a symmetric multi-overlap gap property: the overlaps of collections of r independent sets exhibit an overlap gap property violated by R-local algorithms accepting distributions with particular correlation structures.

In particular, we can construct such a collection of r ER graphs, each pair of which are p-correlated:

(a) Generate $\tilde{G}_0 \sim G(n, pd/n)$.

(b) $\tilde{G}_i \overset{iid}{\sim} G(n, 1-\frac{1-d/n}{1-pd/n})$, for each $i \in [r]$. Note that the edge probability is asymptotically $(1-p)d/n$.

(c) Return the collection of p-correlated graphs $G_i = \tilde{G}_0 \cup \tilde{G}_i$. Each possible edge of G_i exists with probability $pd/+(1-pd/n)\cdot\left(1-\frac{1-d/n}{1-pd/n}\right) = d/n$, and that the correlation of this edge existing between each pair of graphs is p.

In the next section, we will sketch a cleaner proof of Theorem 2.3 by Wein [6], leveraging an asymmetric version of multi-OGP.

3 Asymmetric multi-OGP proof

Suppose $\mathscr{A}$ is R-local and $\mathbb{E}|\mathscr{A}((G(n,d/n)))| \ge (1+3\epsilon)\Phi$. By concentration,

$$\mathbb{P}\{|\mathscr{A}(G(n,d/n))| \ge (1+2\epsilon)\Phi\} \ge 1-e^{-\delta n}$$

for small $\delta > 0$.

We'll generate a sequence of ER graphs in the following way. Start with $G_0 \sim G(n,d/n)$. Considering an ordering $\{e_k\}$ of the $\binom{n}{2}$ possible edges, where we allow $k > \binom{n}{2}$ to wrap around the indices mod $\binom{n}{2}$. In step $t \in \mathbb{N}$, we re-sample the event that e_t exists in the graph independently to generate the graph G_t given G_{t-1}.

The intuition here is that if you change one edge, you change a sub-linear number of i.i.d. factors (defined in the last lecture). This means that the sequence of graphs maps to an (intuitively) continuous path in the space of independent sets. We'll trace out this path in independent set space until we get to an independent set with sufficiently small overlap with our last marked independent set. We'll mark this independent set and continue tracing out the path to obtain a sequence of sufficiently separated independent sets S_i.

This collection of times and sets will be termed the "forbidden structure," and has the following precise definition:

Definition 3.1. (Forbidden structure with (k,n,ϵ).) Let $0 = t_0 < t_1 < \cdots < t_k \le n^3$ be a sequence of times such that the independent sets $S_i := \mathscr{A}(G_{t_i})$ satisfies

$$\left| S_{j+1} \setminus \bigcup_{i \le j} S_i \right| \in \left[\frac{\epsilon\Phi}{4}, \frac{\epsilon\Phi}{2}\right] \tag{1}$$

for each $j \in [k]$. (Intuitively, about Φ new vertices appear which haven't been used yet.) Then the tuple $(t_0, S_0, \ldots, t_k, S_k)$ is a forbidden structure.

We start by showing that R-local algorithms generate forbidden structures.

Lemma 3.2. *Take $1 \ll \epsilon^{-1} \ll d$. Suppose $\mathscr{A}$ is R-local and $\mathbb{E}|\mathscr{A}(G(n,d/n))| \geq (1+3\epsilon)\Phi$. Then with probability at least $1-e^{-\delta n}$ for some $\delta > 0$, the forbidden structure exists for $k = \epsilon^{-3}$.*

Proof. We'll apply a discrete version of the intermediate value theorem to show the interpolation sequence of ER graphs will generate a collection of S_i such that (1) is satisfied (without yet assuring that $t_k < n^3$). R-locality of the algorithm $\mathscr{A}$ implies that $|\mathscr{A}(G_t)\Delta\mathscr{A}(G_{t+1})| \leq \frac{\epsilon\Phi}{10}$ for all t with high probability, since re-sampling edge e_{t+1} affects at most $O(d^R)$ many iid factors, which is vanishingly small compared to Φ.

We shall use the fact that $\left|\mathscr{A}(G_{t_j+\binom{n}{2}}) \cap \left(\bigcup_{i\leq j} S_i\right)\right| \leq \Phi/2$. Indeed, choosing $\epsilon^{-2} \leq 2\sqrt{d}/\log d$, the forbidden structure upper bound (1) implies that

$$\left|\bigcup_{i\leq j} S_i\right| \leq j \cdot \frac{\epsilon\Phi}{2} \leq k \cdot \frac{\epsilon\Phi}{2} \leq \frac{n}{\sqrt{d}}.$$

It follows that the restriction of $G_{t_j+\binom{n}{2}}$ to the set of vertices $\bigcup_{i\leq j} S_i$ is isomorphic to $G(\leq n/\sqrt{d}, d/n)$ (by independence of the graph $G_{t_j+\binom{n}{2}}$ with the independent sets $S_0,\ldots,S_j$). The upper bound of the maximum independent set identified by an R-local algorithm on this subgraph is smaller than $\Phi/2$ whp. Note that we used the fact that $t_k \leq n^3$ here.

Now fix some $j < k$. The difference $|\mathscr{A}(G_{t_k+\binom{n}{2}})\backslash\bigcup_{i\leq j} S_i| \geq \Phi/2$ whp, which by our earlier symmetric difference discretization shows that there must exist some time t_{j+1} such that $S_{j+1} = \mathscr{A}(G_{t_{j+1}})$ satisfies (1). □

By Markov's inequality, the following lemma is sufficient to guarantee the exponentially unlikely existence of a forbidden structure. Below we use *forbidden structure* to mean any $(t_0, S_0, \ldots, t_k, S_k)$ as above, in which each S_j is an independent set of size at least $(1+2\epsilon)\Phi$ in G_{t_j}.

Lemma 3.3. *Let $\mathscr{C}$ be the collection of forbidden structures. Then $\mathbb{E}|\mathscr{C}| \leq e^{-\delta n}$, where $\delta := \delta(d,\epsilon) > 0$.*

Proof. Let $b_j \cdot \Phi = \left|S_j \backslash \bigcup_{i<j} S_i\right|$ (number of new vertices) and $a_j\Phi = |S_j|$ (clearly $a_j > b_j$).

The number of time sequences is at most n^{3k}, since each time is at most n^3. This is upper bounded by $n^{3/\epsilon^3} = e^{o(n)}$ given our choice of k.

Similarly, the number of possible values of the sequence $(a_0, b_0, \ldots, a_k, b_k, c)$ is at most $n^{3k+3} \leq e^{o(n)}$ since each value is at most n. Not here that c is defined such that $c\Phi = \left|\bigcup_{j\leq k} S_j\right| \leq C(\epsilon)\cdot\Phi$.

Since the cardinalities of these sequences are sub-exponential, we can simply condition $|\mathscr{C}|$ on them and sum over all possible sequences without affecting the exponential decay of $\mathbb{E}|\mathscr{C}|$.

Indeed, let's compute the first moment for the expected number of forbidden structures conditioned on sequences of times and a's and b's. We will write things up to factors of $(1+o(1))$ in the exponent (for fixed and large d, n). We have the following combinatorial upper bound for $|\mathscr{C}|$, derived by considering the number of iterative constructions of forbidden structures given the constraints:

$$\mathbb{E}[|\mathscr{C}||(t_j, a_j, b_j, \ldots)] \leq \binom{n}{a_0\Phi} \prod_{j\leq k} \binom{n}{b_j\Phi}\binom{c\Phi}{(a_j-b_j)\Phi} \cdot \left(1-\frac{d}{n}\right)^{b_j(a_j-b_j)\Phi^2}.$$

We'll first consider the asymptotics of the expression

$$\leq \binom{n}{b_j\Phi} \cdot \left(1-\frac{d}{n}\right)^{b_j\Phi\cdot(a_j\Phi-b_j\Phi)}.$$

By Stirling's approximation, the combination is $\exp\left(n \cdot \left(\frac{b_j\Phi}{n}\right) \cdot \log\left(\frac{n}{b_j\Phi}\right)\right)$; the log term in the exponent is roughly $\log d$, so the combination grows as roughly $\exp\left(b_j\Phi\log d\right) = \exp(b_j n \log^2 d/d)$.

Since $d/n = o(1)$, the probability term is $\sim \exp\left(-\frac{d}{n}(1+\epsilon)b_j\Phi^2\right) = \exp\left(-(1+\epsilon)b_j\frac{n\log^2 d}{d}\right)$, where we have used the fact that $b_j \le \epsilon\Phi/2$ and $a_j > (1+3\epsilon)\Phi$. Now using $b_j \ge \epsilon\Phi/4$, this quantity is asymptotically upper bounded by $\exp(-\epsilon^2/10n\log^2 d/d)$.

The other binomial coefficient is naively bounded as

$$\binom{c\Phi}{(a_j-b_j)\Phi} \le e^{c\Phi} \le \exp\left(O\left(\frac{n\log d}{d}\right)\right).$$

This is of smaller order by a $\log d$ factor within the exponent.

For each j, the product term is

$$\exp\left(-\frac{\epsilon^2}{10}\frac{n\log^2 d}{d} + O(\frac{n\log d}{d})\right) \le \exp\left(-\epsilon^2\frac{n\log^2 d}{20d}\right).$$

Now considering the entire upper bound for the conditional expectation and taking the expectation over the sequences, we have the upper bound

$$\mathbb{E}[|\mathscr{C}|] \le \exp\left(o(n) + \frac{Cn\log^2 d}{d} - \frac{1}{10\epsilon}\frac{n\log^2 d}{d}\right) \le e^{-\delta n}.$$

□

Remark. What about fixed d, rather than $d >> 1$? [2] showed that the maximum independent set of the regular graph $G_{n,d}$ is whp $n\alpha_*(d) - \log n - c_*(d) \pm O(1)$, where $\alpha_*(d) = 2\log d/d$. This (much more difficult) result uses both first and second moment arguments on 1-RSB clusters, a more involved combinatorial analog of the near-maximal critical points we studied using Kac–Rice.

Remark. (Symmetric binary perceptron.) Recall that this problem asks us to find $\sigma \in \{\pm 1\}^N$, obeying $M = \alpha N$ constraints

$$|\langle \sigma, g^i \rangle| \le \kappa\sqrt{N},$$

where $g_1, \dots, g_M \sim \mathcal{N}(0, I_N)$ iid.

For κ small: solutions exist for $\alpha < \alpha_*(\kappa) \approx \frac{1}{\log_2(1/k)}$.

However algorithms are only known to succeed for $\alpha \le O(\kappa^2)$ [1]. In fact stable algorithms fail for $\alpha \ge Ck^2\log(1/k)$ as shown by [3] using the (symmetric) multi-OGP. This is a rather dramatic gap!

References

[1] Nikhil Bansal and Joel H. Spencer. On-line balancing of random inputs. 4

[2] Jian Ding, Allan Sly, and Nike Sun. Satisfiability threshold for random regular NAE-SAT. 4

[3] David Gamarnik, Eren C. Kızıldau**g**, Will Perkins, and Changji Xu. Algorithms and barriers in the symmetric binary perceptron model. 4

[4] David Gamarnik and Madhu Sudan. Limits of local algorithms over sparse random graphs. 1

[5] Mustazee Rahman and Bálint Virág. Local algorithms for independent sets are half-optimal. *The Annals of Probability*, 45(3), May 2017. 2

[6] Alexander S. Wein. Optimal low-degree hardness of maximum independent set. 2