The large deviation principle for the Erdős-Rényi random graph

Sourav Chatterjee

(Courant Institute, NYU)

joint work with S. R. S. Varadhan

Main objective: how to count graphs with a given property

- Only consider finite undirected graphs without self-loops in this talk.
- $2^{n(n-1)/2}$ such graphs on n vertices.
- Question: Given a property P and an integer n, roughly how many of these graphs have property P?
- For example, P may be: #triangles $\geq tn^3$, where t is a given constant.
- To make any progress, need to assume some regularity on P. For example, we may demand that P be continuous with respect to some metric.
- What metric? What space?

Another motivation

- Let $G(n, p)$ be the Erdős-Rényi random graph on n vertices where each edge is added independently with probability p.
- Number of triangles in $G(n, p)$ roughly $\binom{n}{3} p^3 \sim n^3 p^3 / 6$.
- What if, just by chance, #triangles turns out to be $\approx tn^3$ where $t > p^3/6$? What would the graph look like, conditional on this rare event?

An abstract topological space of graphs

- Beautiful unifying theory developed by Lovász and coauthors (listed in order of frequency: V. T. Sós, B. Szegedy, C. Borgs, J. Chayes, K. Vesztergombi, A. Schrijver and M. Freedman). Related to earlier works of Aldous, Hoover, Kallenberg.
- Let G_n be a sequence of simple graphs whose number of nodes tends to infinity.
- For every fixed simple graph H, let $|\hom(H,G)|$ denote the number of homomorphisms of H into G (i.e. edge-preserving maps $V(H) \to V(G)$, where $V(H)$ and $V(G)$ are the vertex sets).
- This number is normalized to get the homomorphism density

$$t(H,G) := \frac{|\hom(H,G)|}{|V(G)|^{|V(H)|}}.$$

This gives the probability that a random mapping $V(H) \to V(G)$ is a homomorphism.

Abstract space of graphs contd.

- Suppose that $t(H, G_n)$ tends to a limit $t(H)$ for every H.
- Then Lovász & Szegedy proved that there is a natural "limit object" in the form of a function $f \in \mathcal{W}$, where $\mathcal{W}$ is the space of all measurable functions from $[0,1]^2$ into $[0,1]$ that satisfy $f(x,y) = f(y,x)$ for all x, y.
- Conversely, every such function arises as the limit of an appropriate graph sequence.
- This limit object determines all the limits of subgraph densities: if H is a simple graph with k vertices, then

$$t(H, f) = \int_{[0,1]^k} \prod_{(i,j)\in E(H)} f(x_i, x_j)\, dx_1 \cdots dx_k.$$

- A sequence of graphs $\{G_n\}_{n\geq 1}$ is said to converge to f if for every finite simple graph H,

$$\lim_{n\to\infty} t(H, G_n) = t(H, f).$$

Example

- For any fixed graph H,

$$t(H, G(n,p)) \to p^{|E(H)|} \text{ almost surely as } n \to \infty.$$

- On the other hand, if f is the function that is identically equal to p, then $t(H, f) = p^{|E(H)|}$.
- Thus, the sequence of random graphs $G(n,p)$ converges almost surely to the non-random limit function $f(x,y) \equiv p$ as $n \to \infty$.

Abstract space of graphs contd.

- The elements of $\mathcal{W}$ are sometimes called 'graphons'.
- A finite simple graph G on n vertices can also be represented as a graphon f^G is a natural way:

$$f^G(x,y) = \begin{cases} 1 & \text{if } (\lceil nx \rceil, \lceil ny \rceil) \text{ is an edge in } G, \\ 0 & \text{otherwise.} \end{cases}$$

- Note that this allows *all* simple graphs, irrespective of the number of vertices, to be represented as elements of the single abstract space $\mathcal{W}$.
- So, what is the topology on this space?

The cut metric

- For any $f, g \in \mathcal{W}$, Frieze and Kannan defined the cut distance:

$$d_\square(f,g) := \sup_{S,T \subseteq [0,1]} \left| \int_{S\times T} [f(x,y) - g(x,y)]\, dxdy \right|.$$

- Introduce an equivalence relation on $\mathcal{W}$: say that $f \sim g$ if $f(x,y) = g_\sigma(x,y) := g(\sigma x, \sigma y)$ for some measure preserving bijection σ of $[0,1]$.
- Denote by $\widetilde{g}$ the closure in $(\mathcal{W}, d_\square)$ of the orbit $\{g_\sigma\}$.
- The quotient space is denoted by $\widetilde{\mathcal{W}}$ and τ denotes the natural map $g \to \widetilde{g}$.
- Since $d_\square$ is invariant under σ one can define on $\widetilde{\mathcal{W}}$ the natural distance $\delta_\square$ by

$$\delta_\square(\widetilde{f}, \widetilde{g}) := \inf_\sigma d_\square(f, g_\sigma) = \inf_\sigma d_\square(f_\sigma, g) = \inf_{\sigma_1,\sigma_2} d_\square(f_{\sigma_1}, g_{\sigma_2})$$

making $(\widetilde{\mathcal{W}}, \delta_\square)$ into a metric space.

Cut metric and graph limits

To any finite graph G, we associate the natural graphon f^G and its orbit $\widetilde{G} = \tau f^G = \widetilde{f}^G \in \widetilde{\mathcal{W}}$. One of the key results of the is the following:

Theorem (Borgs, Chayes, Lovász, Sós & Vesztergombi)

A sequence of graphs $\{G_n\}_{n\geq 1}$ converges to a limit $f \in \mathcal{W}$ if and only if $\delta_\Box(\widetilde{G}_n, \widetilde{f}) \to 0$ as $n \to \infty$.

Our result

- For any Borel set $\widetilde{A} \subseteq \widetilde{\mathcal{W}}$, let
$$\widetilde{A}_n := \{\widetilde{h} \in \widetilde{A} : \widetilde{h} = \widetilde{G} \text{ for some } G \text{ on } n \text{ vertices}\}.$$
- Let $I(u) := \frac{1}{2}u \log u + \frac{1}{2}(1-u)\log(1-u)$.
- For any $\widetilde{h} \in \widetilde{\mathcal{W}}$, let $I(\widetilde{h}) := \iint I(h(x,y))dxdy$, where h is any element of $\widetilde{h}$.

Theorem (Chatterjee & Varadhan, 2010)

The function I is well-defined and lower-semicontinuous on $\widetilde{\mathcal{W}}$. If $\widetilde{F}$ is a closed subset of $\widetilde{\mathcal{W}}$ then

$$\limsup_{n\to\infty} n^{-2} \log |\widetilde{F}_n| \le - \inf_{\widetilde{h}\in\widetilde{F}} I(\widetilde{h})$$

and if $\widetilde{U}$ is an open subset of $\widetilde{\mathcal{W}}$, then

$$\liminf_{n\to\infty} n^{-2} \log |\widetilde{U}_n| \ge - \inf_{\widetilde{h}\in\widetilde{U}} I(\widetilde{h}).$$

Erdős-Rényi graphs

- Counting graphs can be related to finding large deviation probabilities for Erdős-Rényi random graphs.
- For example,

$$\begin{aligned}&\#\text{graphs on } n \text{ vertices satisfying } P\\ &= 2^{n(n-1)/2}\mathbb{P}(G(n, 1/2) \text{ satisfies } P).\end{aligned}$$

- Indeed, the main result in our paper is stated as a large deviation principle for the Erdős-Rényi graph, which can be easily proved to be equivalent to the graph counting principle stated before.

Large deviation principle for ER graphs

- The random graph $G(n, p)$ induces probability distribution $\widetilde{\mathbb{P}}_{n,p}$ on the space $\widetilde{\mathcal{W}}$ through the map $G \to \widetilde{G}$.
- Let $I_p(u) := \frac{1}{2} u \log \frac{u}{p} + \frac{1}{2}(1-u) \log \frac{1-u}{1-p}$.
- For $\widetilde{h} \in \widetilde{\mathcal{W}}$, let $I_p(\widetilde{h}) := \iint I_p(h(x, y))dxdy$, where h is any element of $\widetilde{h}$.

Theorem (Chatterjee & Varadhan, 2010)

For any closed set $\widetilde{F} \subseteq \widetilde{\mathcal{W}}$,

$$\limsup_{n\to\infty} \frac{1}{n^2} \log \widetilde{\mathbb{P}}_{n,p}(\widetilde{F}) \le - \inf_{\widetilde{h}\in\widetilde{F}} I_p(\widetilde{h}).$$

and for any open set $\widetilde{U} \subseteq \widetilde{\mathcal{W}}$,

$$\liminf_{n\to\infty} \frac{1}{n^2} \log \widetilde{\mathbb{P}}_{n,p}(\widetilde{U}) \ge - \inf_{\widetilde{h}\in\widetilde{U}} I_p(\widetilde{h}).$$

Method of proof

- The LDP can be proved by standard techniques for the weak topology on $\widetilde{\mathcal{W}}$. (Fenchel-Legendre transforms, Gärtner-Ellis theorem, etc.)
- However, the weak topology is not very interesting. For example, subgraph counts are not continuous with respect to the weak topology.
- The LDP for the topology of the cut metric does not follow via standard methods.

Szemerédi's lemma

- Let $G = (V, E)$ be a simple graph of order n.
- For any $X, Y \subseteq V$, let $e_G(X, Y)$ be the number of X-Y edges of G and let
$$\rho_G(X, Y) := \frac{e_G(X, Y)}{|X||Y|}$$
- Call a pair (A, B) of disjoint sets $A, B \subseteq V$ ϵ-regular if all $X \subseteq A$ and $Y \subseteq B$ with $|X| \geq \epsilon|A|$ and $|Y| \geq \epsilon|B|$ satisfy $|\rho_G(X, Y) - \rho_G(A, B)| \leq \epsilon$.
- A partition $\{V_0, \ldots, V_K\}$ of V is called an ϵ-regular partition of G if it satisfies the following conditions: (i) $|V_0| \leq \epsilon n$; (ii) $|V_1| = |V_2| = \cdots = |V_K|$; (iii) all but at most ϵK^2 of the pairs (V_i, V_j) with $1 \leq i < j \leq K$ are ϵ-regular.

Theorem (Szemerédi's lemma)

Given $\epsilon > 0$, $m \geq 1$ there exists $M = M(\epsilon, m)$ such that every graph of order $\geq M$ admits an ϵ-regular partition $\{V_0, \ldots, V_K\}$ for some $K \in [m, M]$.

Finishing the proof using Szemerédi's lemma

- Suppose G is a graph of order n with ϵ-regular partition $\{V_0, \ldots, V_K\}$.
- Let G' be the random graph with independent edges where a vertex $u \in V_i$ is connected to a vertex $v \in V_j$ with probability $\rho_G(V_i, V_j)$.
- Using Szemerédi's regularity lemma, one can prove that $\delta_\square(G, G') \simeq 0$ with high probability if K and n are appropriately large and ϵ is small.
- The chance that $G(n,p) \approx G'$ is computed by approximating the probability density of G' with respect to $G(n,p)$ as in the proof of Cramér's theorem.
- Combining the two steps gives the probability that $G(n,p) \approx G$, leading to the LDP for $G(n,p)$, and finally leading to the graph counting theorem.

Conditional distributions

Theorem

Take any $p \in (0,1)$. Let $\widetilde{F}$ be a closed subset of $\widetilde{\mathcal{W}}$ satisfying

$$\inf_{\widetilde{h} \in \widetilde{F}^o} I_p(\widetilde{h}) = \inf_{\widetilde{h} \in \widetilde{F}} I_p(\widetilde{h}) > 0.$$

Let $\widetilde{F}^*$ be the subset of $\widetilde{F}$ where I_p is minimized. Then $\widetilde{F}^*$ is non-empty and compact, and for each n, and each $\epsilon > 0$,

$$\mathbb{P}(\delta_\square(G(n,p), \widetilde{F}^*) \geq \epsilon \mid G(n,p) \in \widetilde{F}) \leq e^{-C(\epsilon, \widetilde{F})n^2}$$

where $C(\epsilon, \widetilde{F})$ is a positive constant depending only on ϵ and $\widetilde{F}$.

Proof: Follows from the compactness of $\widetilde{\mathcal{W}}$ (a deep result of Lovász and Szegedy, involving recursive applications of Szemerédi's lemma and martingales).

Large deviations for triangle counts

- Let $T_{n,p}$ be the number of triangles in $G(n,p)$.
- Objective: to evaluate the limit

$$\lim_{n\to\infty} \frac{1}{n^2} \log \mathbb{P}(T_{n,p} \geq (1+\epsilon)\mathbb{E}(T_{n,p}))$$

  as a function of p and ϵ.
- In the sparse case ($p \to 0$), a long-standing conjecture about matching the upper and lower bounds was recently resolved by Chatterjee and also shortly afterward by DeMarco & Kahn.
- For fixed p, exact evaluation of limit due to Chatterjee & Dey: for a certain explicit set of (p,t),

$$\lim_{n\to\infty} \frac{1}{n^2} \log \mathbb{P}(T_{n,p} \geq tn^3) = -I_p((6t)^{1/3}),$$

  when $I_p(u) := \frac{1}{2} u \log \frac{u}{p} + \frac{1}{2}(1-u) \log \frac{1-u}{1-p}$.
- Unfortunately, the result does not cover all values of (p,t).

Large deviations for triangle counts contd.

- Recall: $\mathcal{W}$ is the space of symmetric measurable functions from $[0,1]^2$ into $[0,1]$.
- For each $f \in \mathcal{W}$, let
$$T(f) := \frac{1}{6}\int_0^1\int_0^1\int_0^1 f(x,y)f(y,z)f(z,x)\,dx\,dy\,dz$$
and let $I_p(f) = \iint I_p(f(x,y))dxdy$.
- For each $p \in (0,1)$ and $t \geq 0$, let
$$\phi(p,t) := \inf\{I_p(f) : f \in \mathcal{W},\ T(f) \geq t\}. \tag{1}$$

Theorem (Chatterjee & Varadhan, 2010)

For each $p \in (0,1)$ and each $t \geq 0$,
$$\lim_{n\to\infty} \frac{1}{n^2} \log \mathbb{P}(T_{n,p} \geq tn^3) = -\phi(p,t).$$
Moreover, the infimum is attained in the variational problem (1).

The ‘replica symmetric’ phase

Theorem (Chatterjee & Varadhan, 2010)

Let $h_p(t) := I_p((6t)^{1/3})$. *Let* $\hat{h}_p$ *be the convex minorant of* h_p. *If* t *is a point where* $h_p(t) = \hat{h}_p(t)$, *then* $\phi(p,t) = h_p(t)$. *Moreover, for such* (p,t), *the conditional distribution of* $G(n,p)$ *given* $T_{n,p} \geq tn^3$ *is indistinguishable from the law of* $G(n,(6t)^{1/3})$ *in the large* n *limit.*

Remarks: This result recovers the result of Chatterjee & Dey and gives more. However, the theorem of Chatterjee & Dey gives an error bound of order $n^{-1/2}$, which is impossible to obtain via Szemerédi’s lemma.

'Replica symmetry breaking'

The following theorem shows that given any t, for all p small enough, the conditional distribution of $G(n,p)$ given $T_{n,p} \geq tn^3$ does not resemble that of an Erdős-Rényi graph.

Theorem (Chatterjee & Varadhan, 2010)

Let $\widetilde{C}$ denote the set of constant functions in $\widetilde{\mathcal{W}}$ (representing all Erdős-Rényi graphs). For each t, there exists $p' > 0$ and $\epsilon > 0$ such that for all $p < p'$,

$$\lim_{n\to\infty} \mathbb{P}(\delta_{\square}(G(n,p), \widetilde{C}) > \epsilon \mid T_{n,p} \geq tn^3) = 1.$$

The double phase transition

Theorem (Chatterjee & Varadhan, 2010)

There exists $p_0 > 0$ *such that if* $p \le p_0$*, then there exists* $p^3/6 < t' < t'' < 1/6$ *such that the replica symmetric picture holds when* $t \in (p^3/6, t') \cup (t'', 1/6)$*, but there is a non-empty subset of* (t', t'') *where replica symmetry breaks down.*

The small p limit

The following theorem says that when t is fixed and p is very small, then conditionally on the event $\{T_{n,p} \ge tn^3\}$ the graph $G(n,p)$ must look like a clique.

Theorem (Chatterjee & Varadhan, 2010)

For each t,

$$\lim_{p\to 0} \frac{\phi(p,t)}{\log(1/p)} = \frac{(6t)^{2/3}}{2}.$$

Moreover, if

$$\chi_t(x,y) := 1_{\{\max\{x,y\} \le (6t)^{1/3}\}}$$

is the graphon representing a clique with triangle density t, then for each $\epsilon > 0$,

$$\lim_{p\to 0}\lim_{n\to\infty} \mathbb{P}(\delta_\square(\widetilde{G(n,p)}, \widetilde{\chi}_t) \ge \epsilon \mid T_{n,p} \ge tn^3) = 0.$$

Lower tails

- Given a fixed simple graph H,

$$\lim_{u\to 0}\lim_{n\to\infty}\frac{\log\mathbb{P}(t(H,G(n,p))\le u)}{n^2} = -\frac{1}{2(\chi(H)-1)}\log\frac{1}{1-p},$$

where $\chi(H)$ is the chromatic number of H.

- Closely related to the Erdős-Stone theorem from extremal graph theory.
- In fact, the precise result implies the following: given that $t(H,G(n,p))$ is very small (or zero), the graph $G(n,p)$ looks like a complete $(\chi(H)-1)$-partite graph with $(1-p)$-fraction of edges randomly deleted.
- However, if $t(H,G(n,p))$ is just a little bit below its expected value, the graph continues to look like an Erdős-Rényi graph as in the upper tail case.

Open questions

- There are many questions that remain unresolved, even in the simple example of upper tails for triangle counts. For example:
- What is the set of optimal solutions of the variational problem defining the rate function in the broken replica symmetry phase (i.e. where the optimizer is not a constant)?
- Is the solution unique in the quotient space $\widetilde{\mathcal{W}}$, or can there exist multiple solutions?
- Is it possible to explicitly compute a nontrivial solution for at least some values of (p, t) in the broken replica symmetry region?
- Is it possible to even numerically evaluate or approximate a solution using a computer?
- What is the full characterization of the replica symmetric phase? What is the phase boundary?
- What happens in the sparse case where p and t are both allowed to tend to zero?

Acknowledgment

Special thanks to: Amir Dembo, who suggested the problem to me in 2005. An old manuscript due to Bolthausen, Comets and Dembo (2003) provided a partial solution to the question but was never published.