Phase transitions in (generalized) exponential random graphs

Mei Yin
Department of Mathematics, University of Denver

The exponential random graph model has been a topic of continued research interest. The past few years especially has witnessed (exponentially) growing attention in exponential models and their variations. Emphasis has been made on the variational principle of the limiting normalization constant (free energy density), concentration of the limiting probability distribution, phase transitions, and asymptotic structures.

This presentation is based on joint work with several collaborators, including Sukhada Fadnavis (Harvard University), Richard Kenyon (Brown University), Charles Radin (University of Texas at Austin), and Alessandro Rinaldo (Carnegie Mellon University), and will focus on the phenomenon of phase transitions in (generalized) exponential random graphs. Research supported under NSF grant DMS-1308333.

What is an exponential random graph? Consider the set $\mathcal{G}_n$ of all simple graphs G_n on n vertices ("simple" means undirected, with no loops or multiple edges). The k-parameter family of exponential random graphs is defined by assigning a probability mass function $\mathbb{P}_n^\beta(G_n)$ to every simple graph $G_n \in \mathcal{G}_n$:

$$\mathbb{P}_n^\beta(G_n) = \exp\left(n^2\left(\beta_1 t(H_1, G_n) + \cdots + \beta_k t(H_k, G_n) - \psi_n^\beta\right)\right), \tag{1}$$

where $\beta = (\beta_1, ..., \beta_k)$ are k real parameters, $H_1, ..., H_k$ are pre-chosen finite simple graphs (and we take H_1 to be a single edge), $t(H_i, G_n)$ is the density of graph homomorphisms (the probability that a random vertex map $V(H_i) \to V(G_n)$ is edge-preserving),

$$t(H_i, G_n) = \frac{|\hom(H_i, G_n)|}{|V(G_n)|^{|V(H_i)|}}, \tag{2}$$

and ψ_n^β is the normalization constant,

$$\psi_n^\beta = \frac{1}{n^2}\log \sum_{G_n \in \mathcal{G}_n} \exp\left(n^2\left(\beta_1 t(H_1, G_n) + \cdots + \beta_k t(H_k, G_n)\right)\right). \tag{3}$$

Why are we interested? The popularity of exponential family of random graphs lies in the fact that they capture a wide variety of common network tendencies by representing a complex global structure through a set of tractable local features. They are particularly useful when one wants to simulate observed networks as closely as possible, but without going into details of the specific process underlying network formation.

What has been done? Exponential random graphs have been widely studied since the pioneering work on the independent case by Erdős and Rényi. The theoretical foundations for these models were originally laid by Besag, who applied methods of statistical analysis and demonstrated the powerful Markov-Gibbs equivalence (Hammersley-Clifford theorem) in the context of spatial data. Building on Besag's work, further investigations quickly followed. Holland and Leinhardt derived the exponential family of distributions for networks in the directed case. Frank and Strauss showed that the random graph edges form a Markov random field when the local network features are given by counts of various triangles and stars. Newer developments will be referenced later!

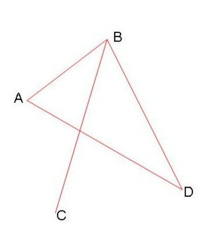
Complication: Graphs with different numbers of vertices belong to different probability spaces!

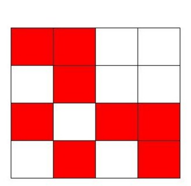
Solution: Theory of graph limits (graphons)! (Lovász and coauthors; earlier work of Aldous and Hoover)

Graphon space $\mathcal{W}$ is the space of all symmetric measurable functions $h(x, y)$ from $[0,1]^2$ into $[0,1]$ (referred to as a "graph limit" or "graphon"). The interval $[0,1]$ represents a "continuum" of vertices, and $h(x,y)$ denotes the probability of putting an edge between x and y.

- Erdős-Rényi graph $G(n, \rho)$, $h(x, y) = \rho$.
- Any $G_n \in \mathcal{G}_n$,

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





Two graphons $f \sim h$ if $f(x,y) = h_\sigma(x,y) := h(\sigma x, \sigma y)$ for some measure preserving bijection σ of $[0,1]$, which in the context of finite graphs may be thought of as vertex relabeling. Reduced graphon space $\tilde{\mathcal{W}}$ is the resulting quotient space of $\mathcal{W}$ mod the equivalence relation $\sim$.

Large deviation: The limiting normalization constant

$$\psi^\beta := \lim_{n\to\infty} \psi_n^\beta = \max_{\tilde{h} \in \tilde{\mathcal{W}}} \left(\beta_1 t(H_1, \tilde{h}) + \cdots + \beta_k t(H_k, \tilde{h}) - \frac{1}{2}\iint_{[0,1]^2} I(\tilde{h}) dx dy\right), \tag{4}$$

where $I : [0,1] \to \mathbb{R}$ is the function

$$I(u) = u \log u + (1-u)\log(1-u). \tag{5}$$

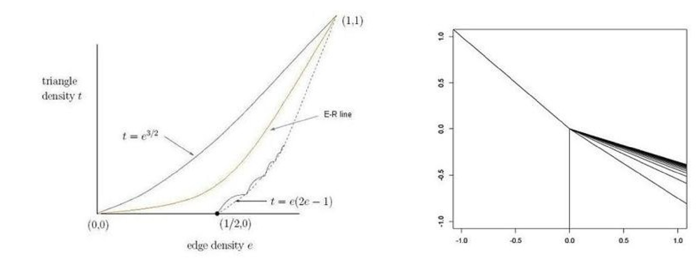
Let $\tilde{H}$ be the subset of $\tilde{\mathcal{W}}$ where ψ^β is maximized. G_n lies close to $\tilde{H}$ with high probability for large n. (Chatterjee and Varadhan; Chatterjee and Diaconis)

Since the limiting normalization constant is the generating function for the limiting expectations of other random variables on the graph space such as expectations and correlations of homomorphism densities, a phase transition occurs when ψ^β is non-analytic or when $\tilde{H}$ is not a singleton set.

The edge-triangle model is a 2-parameter exponential random graph model obtained by taking H_1 to be a single edge and H_2 to be a triangle. For each fixed edge density e in between 0 and 1, there is an upper bound and a lower bound for the triangle density t (up to insignificant errors $o(1)$), and every value of t between the upper bound and the lower bound is (asymptotically) feasible. The best possible upper bound $t \le e^{3/2} + o(1)$ can be easily derived by applying Hölder's inequality, and is a sharp bound, attainable at a complete subgraph on $(e^{1/2} + o(1))n$ vertices. The lower bound is trickier. The trivial lower bound $t \ge 0$ is attainable when $e \le 1/2 - o(1)$ at a complete bipartite graph with $1 - 2e$ fraction of edges randomly deleted, and is a sharp bound by Túran's theorem. For $e \ge 1/2$, the optimal bound was obtained much later by Razborov, who established that, using the flag algebra calculus, for $1 - 1/k \le e \le 1 - 1/(k+1)$ with $k \ge 2$,

$$t \ge \frac{(k-1)\left(k - 2\sqrt{k(k - e(k+1))}\right)\left(k + \sqrt{k(k-e(k+1))}\right)^2}{k^2(k+1)^2} - o(1). \tag{6}$$

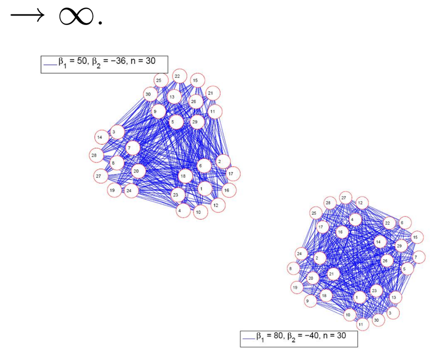
All the curve segments describing the lower boundary are strictly convex, and the boundary points of those segments are precisely the Turán graphons with k classes $v_k = \left(\frac{k}{k+1}, \frac{k(k-1)}{(k+1)^2}\right)$ for $k = 1, 2, ...$. The normal lines of the convex hull of these curves will be essential later!



Chatterjee and Diaconis demonstrated that, as $n \to \infty$ and $\beta_2 \to -\infty$, the exponential model will begin to exhibit a peculiar extremal behavior, in the sense that a typical random graph from such a distribution will be close to a random subgraph of a complete bipartite graph. We extend this result by letting the size of the network n grow unbounded and the natural parameters (β_1, β_2) diverge along generic straight lines. We elucidate the relationship between all possible directions along which the natural parameters can diverge and the way the model tends to place most of its mass on graph configurations that resemble complete multipartite graphs for large enough n.

Double asymptotic framework: Take $\beta_1 = a\beta_2 + b$. Fix a and b.

- Let $n \to \infty$ and then let $\beta_2 \to -\infty$.
- Let $\beta_2 \to -\infty$ and then let $n \to \infty$.



Conclusion: G_n exhibits quantized behavior, jumping from one complete multipartite structure to another, and the jumps happen precisely at the normal lines of a polyhedral set with infinitely many facets. (Y, Rinaldo, and Fadnavis; related work in Handcock; Rinaldo et al.)

The exponential family of random graphs discussed above assumes no prior knowledge of the graph before sampling. But in many situations partial information of the graph is already known beforehand. For example, practitioners might be told that the edge density of the graph is close to 1/2 or the triangle density is close to 1/4 or the adjacency matrix of the graph obeys a certain form. A natural question to ask then is what would be a typical random graph drawn from an exponential model subject to these constraints? Or perhaps more importantly will there be a similar phase transition phenomenon as in the unconstrained exponential model?

Without loss of generality, we assume that the edge density of the graph is approximately known. Let $e : 0 \le e \le 1$ be a real parameter that signifies an "ideal" edge density. Take $\alpha > 0$. The conditional normalization constant $\psi_{n,\alpha}^{e,\beta}$ is analogously defined as the normalization constant for the unconstrained exponential random graph model,

$$\psi_{n,\alpha}^{e,\beta} = \frac{1}{n^2}\log \sum_{G_n \in \mathcal{G}_n : |e(G_n) - e| < \alpha} \exp\left(n^2\left(\beta_1 t(H_1, G_n) + \cdots + \beta_k t(H_k, G_n)\right)\right), \tag{7}$$

the difference being that we are only taking into account graphs G_n whose edge density $e(G_n)$ is within an α neighborhood of e. Correspondingly, the associated conditional probability mass function $\mathbb{P}_{n,\alpha}^{e,\beta}(G_n)$ is given by

$$\mathbb{P}_{n,\alpha}^{e,\beta}(G_n) = \exp(-n^2\psi_{n,\alpha}^{e,\beta})\exp\left(n^2\left(\beta_1 t(H_1, G_n) + \cdots + \beta_k t(H_k, G_n)\right)\right)\mathbb{1}_{|e(G_n) - e| < \alpha}. \tag{8}$$

We perform two limit operations on $\psi_{n,\alpha}^{e,\beta}$. First we take n to infinity, then we shrink the interval around e by letting α go to zero:

These two operations ensure that we are examining the asymptotics of exponentially weighted large graphs with edge density sufficiently close to e.

Large deviation: The limiting conditional normalization constant

$$\psi^{e,\beta} := \lim_{n\to\infty} \psi_{n,\alpha}^{e,\beta} = \max_{\tilde{h} \in \tilde{\mathcal{W}} : e(\tilde{h}) = e} \left(\beta_1 t(H_1, \tilde{h}) + \cdots + \beta_k t(H_k, \tilde{h}) - \frac{1}{2}\iint_{[0,1]^2} I(\tilde{h}) dx dy\right), \tag{10}$$

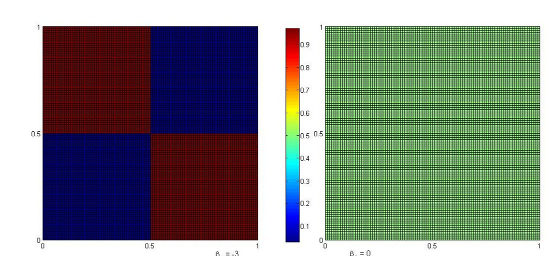
where $I : [0,1] \to \mathbb{R}$ is the function

$$I(u) = u \log u + (1-u)\log(1-u). \tag{11}$$

Let $\tilde{H}$ be the subset of $\{\tilde{\mathcal{W}} : e(\cdot) = e\}$ where $\psi^{e,\beta}$ is maximized. G_n lies close to $\tilde{H}$ with high conditional probability for large n. As in the unconstrained model, a phase transition occurs when $\psi^{e,\beta}$ is non-analytic or when $\tilde{H}$ is not a singleton set. (Kenyon and Y)

We focus on the constrained edge-triangle model. Let e and β_1 be arbitrary but fixed. We are particularly interested in the asymptotics of $\psi^{e,\beta}$ when β_2 is negative, the so-called "repulsive" region. Naturally, varying β_2 allows one to adjust the influence of the triangle density of the graph on the probability distribution. The more negative the β_2, the more unlikely that graphs with a large number of triangles will be observed. When β_2 approaches negative infinity, the most probable graph would likely be triangle free. At the other extreme, when β_2 is zero, the edge-triangle model reduces to the well-studied Erdős-Rényi model, where edges between different vertex pairs are independently included. The structure of triangle free graphs and disordered Erdős-Rényi graphs are apparently quite different, and thus a phase transition is expected as β_2 decays from 0 to $-\infty$. In fact, it is believed that, quite generally, "repulsive" models exhibit a transition qualitatively like the solid/fluid transition, in that a region of parameter space depicting emergent multipartite structure, which is in imitation of the structure of solids, is separated by a phase transition from a region of disordered graphs, which resemble fluids.

The special strip $e = 1/2$: Let β_1 be arbitrary but fixed. As β_2 decreases from 0 to $-\infty$, a typical graph G_n drawn from the constrained "repulsive" edge-triangle model jumps from being Erdős-Rényi to almost complete bipartite, skipping a large portion of the $e = \frac{1}{2}$ line. (Kenyon and Y)



Despite their flexibility, conventionally used exponential random graphs cannot directly model weighted networks as the underlying probability space consists of simple graphs only. An alternative interpretation for simple graphs is such that the edge weights are iid and satisfy a Bernoulli distribution. Generalizations?

Consider the set $\mathcal{G}_n$ of all edge-weighted undirected labeled graphs on n vertices, where the edge weights x_{ij} between vertex i and vertex j are iid real random variables uniformly distributed on $(0,1)$. The common distribution for the edge weights yields probability measure $\mathbb{P}_n$ on $\mathcal{G}_n$. Give the set of such graphs the probability

$$\mathbb{P}_n^\beta(G_n) = \exp\left(n^2\left(\beta_1 t(H_1, G_n) + \cdots + \beta_k t(H_k, G_n) - \psi_n^\beta\right)\right)\mathbb{P}_n(G_n). \tag{12}$$

Large deviation: The limiting normalization constant

$$\psi^\beta := \lim_{n\to\infty} \psi_n^\beta = \max_{\tilde{h} \in \tilde{\mathcal{W}}} \left(\beta_1 t(H_1, \tilde{h}) + \cdots + \beta_k t(H_k, \tilde{h}) - \frac{1}{2}\iint_{[0,1]^2} I(\tilde{h}) dx dy\right), \tag{13}$$

where I, which does not admit closed-form expression, is the Cramér function associated with μ:

$$I(u) = \sup_{\theta \in \mathbb{R}} \left(\theta u - \log \frac{e^\theta - 1}{\theta}\right). \tag{14}$$

Let $\tilde{H}$ be the subset of $\tilde{\mathcal{W}}$ where ψ^β is maximized. G_n lies close to $\tilde{H}$ with high probability for large n. As for simple graphs, a phase transition occurs when ψ^β is non-analytic or when $\tilde{H}$ is not a singleton set. (Chatterjee and Varadhan; Y)

The edge-2-star model is a 2-parameter exponential random graph model obtained by taking H_1 to be a single edge and H_2 to be a 2-star. We concentrate on the asymptotics of ψ^{β_1,β_2} when β_2 is positive, the so-called "attractive" region. The parameter space consists of a single phase with a first order phase transition across the indicated curve and a second order phase transition at the critical point, similar to the transition between liquid and gas in equilibrium materials.