arXiv:1411.4722v1 [math.PR] 18 Nov 2014

# ASYMPTOTICS FOR SPARSE EXPONENTIAL RANDOM GRAPH MODELS

MEI YIN AND LINGJIONG ZHU

ABSTRACT. We study the asymptotics for sparse exponential random graph models where the parameters may depend on the number of vertices of the graph. We obtain a variational principle for the limiting free energy, an associated concentration of measure, the asymptotics for the mean and variance of the limiting probability distribution, and phase transitions in the edge-triangle model. Similar analysis is done for directed sparse exponential random graph models parametrized by edges and outward stars.

## 1. INTRODUCTION

Exponential random graphs are a class of graph ensembles of fixed vertex number $n$ defined by analogy with the Boltzmann ensemble of statistical mechanics. Let $\{\epsilon_p\}$ be a set of local features of a single graph, for example the number of edges or copies of any finite subgraph, as well as more complicated characteristics including the degree sequence or degree distribution, and combinations thereof. These quantities play a role similar to energy in statistical mechanics. Let $\{\beta_p\}$ be a set of inverse temperature parameters whose values we are free to choose. By varying these parameters, one could analyze the influence of different local features on the global structure of the graph. Let $\mathcal{G}_n$ be the set of all possible graphs (undirected and with no self-loops or multiple edges in the simplest case) on $n$ vertices. The $k$-parameter family of exponential random graphs is defined by assigning a probability $\mathbb{P}^{(n)}(G_n)$ to every graph $G_n$ in $\mathcal{G}_n$:

$$\mathbb{P}^{(n)}(G_n) = Z_n(\beta_1^{(n)}, \beta_2^{(n)}, \dots, \beta_k^{(n)})^{-1} \exp\left[\sum_{p=1}^{k} \beta_p^{(n)} \epsilon_p(G_n)\right], \tag{1.1}$$

where $Z_n$ is the partition function,

$$Z_n(\beta_1^{(n)}, \beta_2^{(n)}, \dots, \beta_k^{(n)}) = \sum_{G_n \in \mathcal{G}_n} \exp\left[\sum_{p=1}^{k} \beta_p^{(n)} \epsilon_p(G_n)\right]. \tag{1.2}$$

These rather general models are widely used to model real-world networks, such as the Internet, the World Wide Web, social networks, and biological networks, as they are able to capture a wide variety of common network tendencies by representing a complex global structure through a set of tractable local features, see e.g. Newman [23] and Wasserman and Faust [32]. They are particularly useful when one wants to construct models that resemble observed networks as closely as possible but without going into details of the specific process underlying network formation.

*Date*: 17 November 2014. *Revised:* 17 November 2014.
2000 *Mathematics Subject Classification.* 05C80,82B26,05C35.
*Key words and phrases.* sparse random graphs, exponential random graphs, phase transitions.

Since real-world networks are often very large in size, a pressing objective is to understand the asymptotics of the limiting partition function $Z_n$, the limiting probability distribution $\mathbb{P}^{(n)}(G_n)$, and the limiting free energy $\psi := \lim_{n\to\infty} \frac{1}{\gamma_n} \log Z_n$, for some proper scaling $\gamma_n \to \infty$ as $n \to \infty$, which encodes essential information about the structure of the limiting probability measure. By differentiating $\psi$ with respect to appropriate parameters $\beta_p$, averages of various quantities of interest may be derived. In particular, a phase transition occurs when $\psi$ is non-analytic, since it is the generating function for the limiting expectations of other random variables. Computation of $\psi$ is also important in statistics because it is crucial for carrying out maximum likelihood estimates and Bayesian inference of unknown parameters.

Exponential models have been extensively studied over the last decades. We refer to Besag [4], Snijders et al. [31], Rinaldo et al. [30], and Fienberg [11, 12] for history and a review of developments. In recent years, exponential random graph models and their variations have received (exponentially) growing attention, where the emphasis has been made on the variational principle of the limiting free energy, concentration of the limiting probability distribution, phase transitions and asymptotic structures, see e.g. Chatterjee and Varadhan [10], Chatterjee and Diaconis [9], Radin and Yin [25], Lubetzky and Zhao [22], Radin and Sadun [26, 27], Radin et al. [28], Kenyon et al. [17], Yin [34], Yin et al. [35], Aristoff and Zhu [2, 3], and Zhu [36].

However, in the real world, most networks data are sparse, see e.g. Golub et al. [13], Guyon et al. [14], Hromádka et al. [16] etc. For example, a gene network is sparse since a regulatory pathway involves only a small number of genes; the neural representation of sounds in the auditory cortex of unanesthetized animals is sparse, since the fraction of neurons active at a given instant is small; many biomedical signals have sparse depictions when expressed in a proper basis, see Ye and Liu [33]. Therefore it is important to understand sparse exponential random graph models. Nevertheless, all previous investigations have been centered on dense graphs (number of edges comparable to the square of number of vertices) except some partial results in a very recent paper by Chatterjee and Dembo [8] where very strong assumptions are imposed. A systematic study in the sparse regime is currently lacking, and this will be the main focus of the present paper.

The rest of this paper is organized as follows. In Section 2 we introduce notation and results concerning the theory of graph limits and its use in (undirected) exponential random graph models. In Section 3 we analyze the asymptotic features of the undirected exponential model parametrized by various subgraph densities. Our main results are: a variational principle for the limiting free energy (Theorem 1), an associated concentration of measure (Theorem 2) indicating that almost all large graphs lie near the maximizing set, and the mean and variance of the limiting probability distribution under a scaling assumption about the parameters (Propositions 5-6). We then resort to the large deviations result of Chatterjee and Dembo [8] and obtain exact asymptotics for the limiting partition function of edge-(single)-star model (and beyond) in the sparse regime (Theorem 7). Lastly, we specialize to the edge-triangle model and show the existene of countably many first-order phase transitions (Proposition 10). In Section 4 we analyze the asymptotic features of the directed exponential model parametrized by edges and multiple outward stars. Our main results are: a variational principle for the limiting free energy (Theorem 11), and the mean and variance of the limiting probability distribution under different scaling assumptions about the parameters (Propositions 13-16). We then specialize to the edge-(single)-star model and show the existence of first- and second-order phase transitions.

## 2. Background

We present some background on the theory of graph limits and its use in (undirected) exponential random graph models. Following the earlier work of Aldous [1] and Hoover [15], Lovász and coauthors (V. T. Sós, B. Szegedy, C. Borgs, J. Chayes, K. Vesztergombi, etc.) have constructed a unified and elegant theory of graph limits in a sequence of papers [5, 6, 7, 19, 21]. See also the recent book of Lovász [20] for a comprehensive account and references. This emerging theory has provided a new set of tools for representing and studying the asymptotic behavior of graphs, and has become the object of intense research in many fields, such as discrete mathematics, statistical mechanics, and probability.

Here are the basics of this beautiful theory. Any undirected graph $G_n$ that has no self-loops or multiple edges, irrespective of the number of vertices, may be represented as an element $h^{G_n}$ of a single abstract space $\mathcal{W}$ that consists of all symmetric measurable functions from $[0,1]^2$ into $[0,1]$, by defining

$$h^{G_n}(x,y) = \begin{cases} 1, & \text{if } (\lceil nx \rceil, \lceil ny \rceil) \text{ is an edge in } G_n; \\ 0, & \text{otherwise.} \end{cases} \tag{2.1}$$

A sequence of graphs $\{G_n\}_{n\geq 1}$ is said to converge to a function $h \in \mathcal{W}$ (referred to as a "graph limit" or "graphon") if for every finite simple graph $H$ with vertex set $V(H) = [k] = \{1, \ldots, k\}$ and edge set $E(H)$,

$$\lim_{n\to\infty} t(H, h^{G_n}) = t(H, h), \tag{2.2}$$

where

$$t(H,h) = \int_{[0,1]^k} \prod_{\{i,j\}\in E(H)} h(x_i, x_j) dx_1 \cdots dx_k, \tag{2.3}$$

and so by construction,

$$t(H, h^{G_n}) = t(H, G_n) := \frac{|\text{hom}(H, G_n)|}{|V(G_n)|^{|V(H)|}}, \tag{2.4}$$

the homomorphism density of $H$ in $G_n$. It was shown in Lovász and Szegedy [21] that every function in $\mathcal{W}$ arises as the limit of a certain graph sequence. Intuitively, the interval $[0,1]$ represents a "continuum" of vertices, and $h(x,y)$ denotes the probability of putting an edge between $x$ and $y$. For example, for the Erdős-Rényi random graph $G(n,\rho)$, the "graphon" is represented by the function that is identically equal to $\rho$ on $[0,1]^2$.

This "graphon" interpretation enables us to capture the notion of convergence in terms of subgraph densities by an explicit metric on $\mathcal{W}$, the so-called "cut distance":

$$d_\square(f,h) = \sup_{S,T\subseteq[0,1]} \left| \int_{S\times T} (f(x,y) - h(x,y))\, dx\, dy \right| \tag{2.5}$$

for $f, h \in \mathcal{W}$. A non-trivial complication is that the topology induced by the cut metric is well defined only up to measure preserving transformations of $[0,1]$ (and up to sets of Lebesgue measure zero), which in the context of finite graphs may be thought of as vertex relabeling. The solution is to work instead on an appropriate quotient space. To that end, an equivalence relation $\sim$ is introduced in $\mathcal{W}$. We say that $f \sim h$ if $f(x,y) = h_\sigma(x,y) := h(\sigma x, \sigma y)$ for some measure preserving bijection $\sigma$ of $[0,1]$. Let $\tilde{h}$ (referred to as a "reduced graphon") denote the equivalence class of $h$ in $(\mathcal{W}, d_\square)$. Since $d_\square$ is invariant under $\sigma$, one can then define on the resulting quotient space $\tilde{\mathcal{W}}$ the natural distance $\delta_\square$ by $\delta_\square(\tilde{f}, \tilde{h}) = \inf_{\sigma_1,\sigma_2} d_\square(f_{\sigma_1}, h_{\sigma_2})$, where the infimum ranges over all measure preserving bijections $\sigma_1$ and $\sigma_2$, making $(\tilde{\mathcal{W}}, \delta_\square)$ into a metric space. With some abuse of notation we also refer to $\delta_\square$ as the "cut distance". The space $(\tilde{\mathcal{W}}, \delta_\square)$ enjoys many important properties that are essential for the study of exponential random graph models. For example, it is a compact space and homomorphism densities $t(H, \cdot)$ are continuous functions on it.

## 3. Undirected Graphs

Consider undirected graphs $G_n$ on $n$ vertices, where a graph is represented by a matrix $X = (X_{ij})_{1 \le i < j \le n}$ with each $X_{ij} \in \{0,1\}$. Here, $X_{ij} = 1$ means there is an edge between vertex $i$ and vertex $j$; otherwise, $X_{ij} = 0$. Give the set of such graphs the probability

$$\mathbb{P}^{(n)}(G_n) = Z_n(\beta_1^{(n)}, \beta_2^{(n)}, \dots, \beta_k^{(n)})^{-1} \exp\left[n^2\left(\sum_{p=1}^k \beta_p^{(n)} t(H_p, G_n)\right)\right], \tag{3.1}$$

where $Z_n(\beta_1^{(n)}, \beta_2^{(n)}, \dots, \beta_k^{(n)})$ is the appropriate normalization.

We are interested in the sparse graph, i.e., the probability that there is an edge between vertex $i$ and vertex $j$ goes to 0 as $n \to \infty$. This requires that $\beta_p^{(n)} \to -\infty$, for some $1 \le p \le k$, as $n \to \infty$. Let us assume that

$$\beta_p^{(n)} = \beta_p \alpha_n, \quad p = 1, 2, \dots, k, \tag{3.2}$$

where $\alpha_n \to \infty$ as $n \to \infty$.

**Theorem 1.** *Assume that $H_1$ denotes a single edge.*

$$\psi(\beta_1, \dots, \beta_k) := \lim_{n\to\infty} \frac{1}{n^2 \alpha_n} \log Z_n = \sup_{\substack{h:[0,1]^2 \to [0,1] \\ h(x,y) = h(y,x)}} \{\beta_1 t(H_1, h) + \dots + \beta_k t(H_k, h)\}. \tag{3.3}$$

*Proof.* We can compute that

$$\begin{aligned} Z_n(\beta_1^{(n)}, \dots, \beta_k^{(n)}) &= 2^{\binom{n}{2}} \mathbb{E}\left[e^{n^2(\beta_1^{(n)} t(H_1, G_n) + \dots + \beta_k^{(n)} t(H_k, G_n))}\right] \\ &= \sum_{(x_{ij})_{1 \le i < j \le n} \in \{0,1\}^{n^2}} e^{n^2(\beta_1^{(n)} t(H_1, x) + \dots + \beta_k^{(n)} t(H_k, x))}. \end{aligned} \tag{3.4}$$

On the one hand, we have

$$Z_n(\beta_1^{(n)}, \dots, \beta_k^{(n)}) \le 2^{\binom{n}{2}} e^{n^2 \alpha_n \max_{(x_{ij})_{1 \le i < j \le n} \in \{0,1\}^{n^2}} \{\beta_1 t(H_1, x) + \dots + \beta_k t(H_k, x)\}}, \tag{3.5}$$

and on the other hand,

$$Z_n(\beta_1^{(n)}, \dots, \beta_k^{(n)}) \ge e^{\alpha_n n^2 \max_{(x_{ij})_{1 \le i < j \le n} \in \{0,1\}^{n^2}} \{\beta_1 t(H_1, x) + \dots + \beta_k t(H_k, x)\}}. \tag{3.6}$$

The proof is complete by noticing that

$$\begin{aligned}&\lim_{n\to\infty}\max_{(x_{ij})_{1\le i<j\le n}\in\{0,1\}^{n^2}}\{\beta_1 t(H_1,x)+\cdots+\beta_k t(H_k,x)\} \\ &= \sup_{\substack{h:[0,1]^2\to[0,1]\\ h(x,y)=h(y,x)}}\{\beta_1 t(H_1,h)+\cdots+\beta_k t(H_k,h)\}.\end{aligned} \tag{3.7}$$

First, it is clear that the LHS is less than or equal to the RHS in (3.7). Second, there exists a graphon $h_*$, so that

$$\sum_{p=1}^{k}\beta_p t(H_p,h_*)=\sup_{h:[0,1]^2\to[0,1],h(x,y)=h(y,x)}\left\{\sum_{p=1}^{k}\beta_p t(H_p,h)\right\}. \tag{3.8}$$

For a fixed graphon $h_*$, we can find a sequence of graphs that converge to $h_*$ in the cut norm as $n\to\infty$. Hence we proved (3.7). □

Let $\tilde{H}$ be the subset of $\tilde{W}$ where $\sum_{p=1}^{k}\beta_p t(H_p,\tilde{h})$ is maximized. By the compactness of $\tilde{W}$ and the continuity of $\sum_{p=1}^{k}\beta_p t(H_p,\cdot)$, $\tilde{H}$ is a non-empty compact set. Let $G_n$ be a random graph on $n$ vertices drawn from the sparse exponential random graph model $\mathbb{P}^{(n)}$ (3.1). The following theorem shows that for $n$ large, $\tilde{G}_n$ must lie close to $\tilde{H}$ with high probability. In particular, if $\tilde{H}$ is a singleton set, then the theorem gives a weak law of large numbers for $G_n$.

**Theorem 2.** *Let $\tilde{H}$ be defined as above. Then for any $\eta>0$ there exist $C,\gamma>0$ such that for all $n$ large enough,*

$$\mathbb{P}^{(n)}(\delta_\square(\tilde{G}_n,\tilde{H})>\eta)\le Ce^{-n^2\alpha_n\gamma}. \tag{3.9}$$

*Proof.* Take any $\eta>0$. Let $\tilde{A}$ be the subset of $\tilde{H}$ consisting of reduced graphons that are at least $\eta$-distance away from $\tilde{H}$,

$$\tilde{A}=\{\tilde{h}:\delta_\square(\tilde{h},\tilde{H})>\eta\}. \tag{3.10}$$

By the compactness of $\tilde{W}$ and $\tilde{H}$ and the continuity of $\sum_{p=1}^{k}\beta_p t(H_p,\cdot)$, it follows that

$$\gamma':=\sup_{\tilde{h}\in\tilde{W}}\sum_{p=1}^{k}\beta_p t(H_p,\tilde{h})-\sup_{\tilde{h}\in\tilde{A}}\sum_{p=1}^{k}\beta_p t(H_p,\tilde{h})>0. \tag{3.11}$$

Then

$$\begin{aligned}\mathbb{P}^{(n)}(\tilde{G}_n\in\tilde{A}) &= e^{-n^2\alpha_n\psi_n}\sum_{\delta_\square(\tilde{G}_n,\tilde{H})>\eta}e^{n^2\alpha_n\sum_{p=1}^{k}\beta_p t(H_p,G_n)} \\ &\le e^{-n^2\alpha_n\psi_n}2^{\binom{n}{2}}e^{n^2\alpha_n\sup_{\delta_\square(\tilde{G}_n,\tilde{H})>\eta}\sum_{p=1}^{k}\beta_p t(H_p,G_n)}.\end{aligned} \tag{3.12}$$

Notice that

$$\lim_{n\to\infty}\sup_{\delta_\square(\tilde{G}_n,\tilde{H})>\eta}\sum_{p=1}^{k}\beta_p t(H_p,G_n)=\sup_{\tilde{h}\in\tilde{A}}\sum_{p=1}^{k}\beta_p t(H_p,\tilde{h}). \tag{3.13}$$

First, it is clear that the LHS is less than or equal to the RHS. Second, for any $\epsilon>0$, there exists a reduced graphon $\tilde{h}_*\in\tilde{A}$ so that

$$\sum_{p=1}^{k}\beta_p t(H_p,\tilde{h}_*)\ge\sup_{\tilde{h}\in\tilde{A}}\sum_{p=1}^{k}\beta_p t(H_p,\tilde{h})-\epsilon. \tag{3.14}$$

For a fixed reduced graphon $\tilde{h}_*$, we can find a sequence of graphs $G_n$ that converge to $\tilde{h}_*$ in the cut norm as $n \to \infty$. This implies that

$$\limsup_{n\to\infty} \frac{\log \mathbb{P}^{(n)}(\tilde{G}_n \in \tilde{A})}{n^2\alpha_n} \leq -\sup_{\tilde{h}\in\tilde{W}} \sum_{p=1}^{k} \beta_p t(H_p, \tilde{h}) + \sup_{\tilde{h}\in\tilde{A}} \sum_{p=1}^{k} \beta_p t(H_p, \tilde{h}) = -\gamma'. \tag{3.15}$$

Hence (3.9) holds for any $\gamma < \gamma'$. □

When $\psi = 0$, $h \equiv 0$ is an optimal graphon and in the limit $n \to \infty$, we have an empty graph. That translates to sparse graphs before we take the limit $n \to \infty$. One natural question to ask is for what set of parameters $(\beta_1, \ldots, \beta_k)$ we will get $\psi(\beta_1, \ldots, \beta_k) = 0$. Note that if $\psi = 0$ then $h \equiv 0$ is an optimal graphon and therefore

$$0 = \psi(\beta_1, \ldots, \beta_k) = \sup_{0\leq x\leq 1} \{\beta_1 x + \beta_2 x^2 + \cdots + \beta_k x^k\}. \tag{3.16}$$

So it is interesting to understand when we have

$$\sup_{0\leq x\leq 1} \{\beta_1 x + \beta_2 x^2 + \cdots + \beta_k x^k\} = 0. \tag{3.17}$$

**Remark 3.** *Three trivial observations.*

*(i) If (3.17) holds, then we must have* $\beta_1 + \beta_2 + \cdots + \beta_k < 0$.

*(ii) If (3.17) holds and* $\beta_1 = \cdots = \beta_{\ell-1} = 0$, *then we must have* $\beta_\ell < 0$.

*(iii) If* $\beta_1 < 0$ *and* $\beta_1 + \sum_{j:\beta_j>0} j < 0$, *then (3.17) holds.*

*(i) is true because otherwise* $x = 1$ *is more optimal than* $x = 0$. *(ii) is true since when* $x > 0$ *is very small,* $\beta_\ell x$ *is the dominating term if* $\beta_1 = \cdots = \beta_{\ell-1} = 0$. *(iii) is true since*

$$\ell'(x) = \beta_1 + \sum_{j>1:\beta_j<0} j\beta_j x^{j-1} + \sum_{j:\beta_j>0} j\beta_j x^{j-1} \leq \beta_1 + \sum_{j:\beta_j>0} j, \qquad 0 \leq x \leq 1. \tag{3.18}$$

**Remark 4.** *Let us derive the sufficient and necessary conditions for*

$$\sup_{0\leq x\leq 1} \{\beta_1 x + \beta_2 x^2 + \beta_3 x^3\} = 0, \qquad \beta_1, \beta_2, \beta_3 \neq 0. \tag{3.19}$$

*First, we must have* $\beta_1 < 0$. *We can compute that*

$$\ell'(x) = \beta_1 + 2\beta_2 x + 3\beta_3 x^2. \tag{3.20}$$

*Thus* $\ell'(0) = \beta_1 < 0$. *If* $\beta_3 > 0$, *then as* $x$ *increases from* $0$ *to* $\infty$, $\ell'(x)$ *changes from being negative to positive. Thus as* $x$ *increases from* $0$ *to* $\infty$, $\ell(x)$ *first decreases and then increases. Hence when* $\beta_3 > 0$, $\sup_{0\leq x\leq 1} \ell(x) = 0$ *if and only if* $\beta_1+\beta_2+\beta_3 \leq 0$. *If* $\beta_3 < 0$ *and* $\beta_2 \leq 0$, *then* $\ell'(x) \leq 0$ *for any* $x \geq 0$ *and* $\sup_{0\leq x\leq 1} \ell(x) = 0$. *If* $\beta_3 < 0$ *and* $\beta_2 > 0$, *then there are two positive roots of* $\ell'(x) = 0$, *given by* $x = \frac{-\beta_2 \pm \sqrt{\beta_2^2 - 3\beta_1\beta_3}}{3\beta_3}$ *if* $\beta_2^2 > 3\beta_1\beta_3$. *If* $1 \leq \frac{-\beta_2 + \sqrt{\beta_2^2 - 3\beta_1\beta_3}}{3\beta_3}$, *then on the interval* $[0, 1]$, $\ell(x)$ *first decreases and then increases. Hence,* $\sup_{0\leq x\leq 1} \ell(x) = 0$ *if and only if* $\beta_1 + \beta_2 + \beta_3 \leq 0$. *If* $1 > \frac{-\beta_2 + \sqrt{\beta_2^2 - 3\beta_1\beta_3}}{3\beta_3}$, *then on the interval* $[0, 1]$, $\ell(x)$ *first decreases and then increases and finally decreases. Therefore,* $\sup_{0\leq x\leq 1} \ell(x) = 0$ *if and only if* $\ell(\frac{-\beta_2 + \sqrt{\beta_2^2 - 3\beta_1\beta_3}}{3\beta_3}) \leq 0$. *Finally if* $\beta_3 < 0$, $\beta_2 > 0$ *and* $\beta_2^2 \leq 3\beta_1\beta_3$

*then* $\ell'(x) \leq 0$ *and* $\sup_{0\leq x\leq 1}\ell(x) = 0$. *Hence, to summarize, the sufficient and necessary condition for* (3.19) *is that* $(\beta_1, \beta_2, \beta_3)$ *belongs to the set*

$$\begin{aligned}
&\{\beta_1 < 0, \beta_3 > 0, \beta_1 + \beta_2 + \beta_3 \leq 0\} \cup \{\beta_1 < 0, \beta_3 < 0, \beta_2 \leq 0\} \\
&\quad \cup \{\beta_1 < 0, \beta_3 < 0, \beta_2 > 0, \beta_2^2 \leq 3\beta_1\beta_3\} \\
&\quad \cup \Bigg\{\beta_1 < 0, \beta_3 < 0, \beta_2 > 0, \beta_2^2 > 3\beta_1\beta_3, \\
&\qquad\qquad 1 > \frac{-\beta_2 + \sqrt{\beta_2^2 - 3\beta_1\beta_3}}{3\beta_3}, \ell\left(\frac{-\beta_2 + \sqrt{\beta_2^2 - 3\beta_1\beta_3}}{3\beta_3}\right) \leq 0\Bigg\} \\
&\quad \cup \Bigg\{\beta_1 < 0, \beta_3 < 0, \beta_2 > 0, \beta_2^2 > 3\beta_1\beta_3, \\
&\qquad\qquad 1 \leq \frac{-\beta_2 + \sqrt{\beta_2^2 - 3\beta_1\beta_3}}{3\beta_3}, \beta_1 + \beta_2 + \beta_3 \leq 0\Bigg\}.
\end{aligned}$$

For a sparse random graph, $P^{(n)}(X_{ij} = 1) \to 0$ as $n \to \infty$. It will be interesting to know how sparse the random graph is and how fast $P^{(n)}(X_{ij} = 1)$ converges to zero as $n \to \infty$.

**Proposition 5.** *Assume that* $\beta_1, \ldots, \beta_k$ *are all negative and* $H_1$ *denotes a single edge. Let us further assume that* $\lim_{n\to\infty} n^2 e^{2\alpha_n\beta_1} = 0$ *and* $\lim_{n\to\infty} \frac{\alpha_n}{n} = 0$.

$$\lim_{n\to\infty} \frac{\mathbb{P}^{(n)}(X_{1i} = 1)}{e^{2\beta_1\alpha_n}} = 1. \tag{3.21}$$

*Proof.* By symmetry,

$$\begin{aligned}
\mathbb{P}^{(n)}(X_{1i} = 1) &= \mathbb{E}^{(n)}[X_{1i}] \qquad (3.22) \\
&= \frac{1}{\binom{n}{2}} \mathbb{E}^{(n)}\left[\sum_{1\leq i<j\leq n} X_{ij}\right] \\
&= \frac{1}{\binom{n}{2}} \frac{\mathbb{E}[\sum_{1\leq i<j\leq n} X_{ij} e^{\sum_{p=1}^k \beta_p^{(n)} n^2 t(H_p, X)}]}{\mathbb{E}[e^{\sum_{p=1}^k \beta_p^{(n)} n^2 t(H_p, X)}]} \\
&= \frac{1}{\binom{n}{2}} \frac{2^{\binom{n}{2}} \mathbb{E}[\sum_{1\leq i<j\leq n} X_{ij} e^{\alpha_n n^2 \sum_{p=1}^k \beta_p t(H_p, X)}]}{2^{\binom{n}{2}} \mathbb{E}[e^{\alpha_n n^2 \sum_{p=1}^k \beta_p t(H_p, X)}]}.
\end{aligned}$$

First, let us analyze the denominator. It is clear that

$$2^{\binom{n}{2}} \mathbb{E}[e^{\alpha_n n^2 \sum_{p=1}^k \beta_p t(H_p, X)}] \geq 2^{\binom{n}{2}} \frac{1}{2^{\binom{n}{2}}} = 1. \tag{3.23}$$

On the other hand, since $\beta_i$'s are negative,

$$\begin{aligned}
2^{\binom{n}{2}} \mathbb{E}[e^{\alpha_n n^2 \sum_{p=1}^k \beta_p t(H_p, X)}] &\leq 2^{\binom{n}{2}} \mathbb{E}[e^{\alpha_n n^2 \beta_1 t(H_1, X)}] \qquad (3.24) \\
&= 2^{\binom{n}{2}} \mathbb{E}[e^{2\alpha_n\beta_1 \sum_{i<j} X_{ij}}] \\
&= 2^{\binom{n}{2}} \left(\frac{1 + e^{2\alpha_n\beta_1}}{2}\right)^{\binom{n}{2}} \\
&\to 1,
\end{aligned}$$

as $n \to \infty$ since we assumed that $\lim_{n\to\infty} n^2 e^{2\alpha_n\beta_1} = 0$.

Next, let us analyze the numerator. On the one hand,

$$
\begin{aligned}
&2^{\binom{n}{2}}\mathbb{E}\left[\sum_{1\leq i<j\leq n}X_{ij}e^{\alpha_n n^2\sum_{p=1}^k\beta_p t(H_p,X)}\right] \qquad (3.25)\\
&\geq 2^{\binom{n}{2}}\sum_{1\leq i<j\leq n}\mathbb{E}\left[e^{\alpha_n n^2\sum_{p=1}^k\beta_p t(H_p,X)}\Big|X_{ij}=1,X_{i'j'}=0,(i',j')\neq(i,j)\right]\\
&\qquad\qquad\cdot\mathbb{P}\left(X_{ij}=1,X_{i'j'}=0,(i',j')\neq(i,j)\right)\\
&=e^{2\alpha_n\beta_1}\sum_{1\leq i<j\leq n}\mathbb{E}\left[e^{\alpha_n n^2\sum_{p=2}^k\beta_p t(H_p,X)}\Big|X_{ij}=1,X_{i'j'}=0,(i',j')\neq(i,j)\right]\\
&=\sum_{1\leq i<j\leq n}e^{2\alpha_n\beta_1+\alpha_n n^2\sum_{p=2}^k\beta_p c_p n^{-v(H_p)}}\\
&=\binom{n}{2}e^{2\alpha_n\beta_1+\alpha_n n^2\sum_{p=2}^k\beta_p c_p n^{-v(H_p)}},
\end{aligned}
$$

where $v(H_p)\geq 3$ denotes the number of vertices of $H_p$ and $c_p\geq 0$ is a constant that only depends on $H_p$.

On the other hand,

$$
\begin{aligned}
2^{\binom{n}{2}}\mathbb{E}\left[\sum_{1\leq i<j\leq n}X_{ij}e^{\alpha_n n^2\sum_{p=1}^k\beta_p t(H_p,X)}\right]&\leq 2^{\binom{n}{2}}\mathbb{E}\left[\sum_{1\leq i<j\leq n}X_{ij}e^{\alpha_n n^2\beta_1 t(H_1,X)}\right] \qquad (3.26)\\
&=2^{\binom{n}{2}}\mathbb{E}\left[\sum_{1\leq i<j\leq n}X_{ij}e^{2\alpha_n\beta_1\sum_{i<j}X_{ij}}\right]\\
&=2^{\binom{n}{2}}\frac{1}{2\alpha_n}\frac{\partial}{\partial\beta_1}\mathbb{E}\left[e^{2\alpha_n\beta_1\sum_{i<j}X_{ij}}\right]\\
&=2^{\binom{n}{2}}\frac{1}{2\alpha_n}\frac{\partial}{\partial\beta_1}\left(\frac{1+e^{2\alpha_n\beta_1}}{2}\right)^{\binom{n}{2}}\\
&=e^{2\alpha_n\beta_1}\binom{n}{2}\left(1+e^{2\alpha_n\beta_1}\right)^{\binom{n}{2}-1}.
\end{aligned}
$$

Putting everything together, we proved the desired result. □

**Proposition 6.** *Assume that $\beta_1,\ldots,\beta_k$ are all negative and $H_1$ denotes a single edge. Let us further assume that $\lim_{n\to\infty}n^2e^{2\alpha_n\beta_1}=0$ and $\lim_{n\to\infty}\frac{\alpha_n}{n}=0$.*

$$
\lim_{n\to\infty}\frac{\mathbb{P}^{(n)}(X_{1i}=1,X_{1j}=1)}{e^{4\beta_1\alpha_n}}=1,\quad i\neq j. \qquad (3.27)
$$

*Proof.* By symmetry,

$$\begin{aligned}
&\mathbb{P}^{(n)}(X_{1i}=1, X_{1j}=1) = \mathbb{E}^{(n)}[X_{1i}X_{1j}] \qquad (3.28)\\
&= \frac{4}{(n+1)n(n-1)(n-2)}\left[\mathbb{E}^{(n)}\left[\left(\sum_{1\le i<j\le n} X_{ij}\right)^2\right] - \mathbb{E}^{(n)}\left[\sum_{1\le i<j\le n} X_{ij}\right]\right]\\
&= \frac{4\mathbb{E}\left[\left[\left(\sum_{1\le i<j\le n} X_{ij}\right)^2 - \sum_{1\le i<j\le n} X_{ij}\right] e^{\sum_{p=1}^k \beta_p^{(n)} n^2 t(H_p,X)}\right]}{(n+1)n(n-1)(n-2)\mathbb{E}[e^{\sum_{p=1}^k \beta_p^{(n)} n^2 t(H_p,X)}]}\\
&= \frac{4\cdot 2^{\binom{n}{2}}\mathbb{E}\left[\left[\left(\sum_{1\le i<j\le n} X_{ij}\right)^2 - \sum_{1\le i<j\le n} X_{ij}\right] e^{\alpha_n n^2\sum_{p=1}^k \beta_p t(H_p,X)}\right]}{(n+1)n(n-1)(n-2)2^{\binom{n}{2}}\mathbb{E}[e^{\alpha_n n^2\sum_{p=1}^k \beta_p t(H_p,X)}]}.
\end{aligned}$$

Under the assumption $\lim_{n\to\infty} n^2 e^{2\beta_1\alpha_n} = 0$, we have $2^{\binom{n}{2}}\mathbb{E}[e^{\alpha_n n^2\sum_{p=1}^k \beta_p t(H_p,X)}] \to 1$ as $n\to\infty$. Moreover, it is clear that

$$\begin{aligned}
&2^{\binom{n}{2}}\mathbb{E}\left[\left[\left(\sum_{1\le i<j\le n} X_{ij}\right)^2 - \sum_{1\le i<j\le n} X_{ij}\right] e^{\alpha_n n^2\sum_{p=1}^k \beta_p t(H_p,X)}\right] \qquad (3.29)\\
&\ge 2^{\binom{n}{2}}\frac{1}{2}\sum_{1\le i<j\le n}\ \sum_{1\le i'<j'\le n:(i',j')\ne(i,j)}\\
&\mathbb{E}\left[(2^2-2)e^{\alpha_n n^2\sum_{p=1}^k \beta_p t(H_p,X)}\middle| X_{ij}=X_{i'j'}=1, X_{i''j''}=0, (i'',j'')\ne(i,j) \text{ and } (i'',j'')\ne(i',j')\right]\\
&\qquad\cdot\mathbb{P}\left(X_{ij}=X_{i'j'}=1, X_{i''j''}=0, (i'',j'')\ne(i,j) \text{ and } (i'',j'')\ne(i',j')\right)\\
&= e^{4\alpha_n\beta_1}\sum_{1\le i<j\le n}\ \sum_{1\le i'<j'\le n:(i',j')\ne(i,j)}\\
&\mathbb{E}\left[e^{\alpha_n n^2\sum_{p=2}^k \beta_p t(H_p,X)}\middle| X_{ij}=X_{i'j'}=1, X_{i''j''}=0, (i'',j'')\ne(i,j) \text{ and } (i'',j'')\ne(i',j')\right]\\
&= \sum_{1\le i<j\le n}\ \sum_{1\le i'<j'\le n:(i',j')\ne(i,j)} e^{4\alpha_n\beta_1+\alpha_n n^2\sum_{p=2}^k \beta_p c_p n^{-v(H_p)}}\\
&= \frac{(n+1)n(n-1)(n-2)}{4}e^{4\alpha_n\beta_1+\alpha_n n^2\sum_{p=2}^k \beta_p c_p n^{-v(H_p)}},
\end{aligned}$$

where $v(H_p) \geq 3$ denotes the number of vertices of $H_p$ and $c_p \geq 0$ is a constant that only depends on $H_p$. And

$$\begin{aligned}
&2^{\binom{n}{2}}\mathbb{E}\left[\left[\left(\sum_{1\leq i<j\leq n}X_{ij}\right)^2-\sum_{1\leq i<j\leq n}X_{ij}\right]e^{\alpha_n n^2\sum_{p=1}^k\beta_p t(H_p,X)}\right] \qquad (3.30)\\
&\leq 2^{\binom{n}{2}}\mathbb{E}\left[\left[\left(\sum_{1\leq i<j\leq n}X_{ij}\right)^2-\sum_{1\leq i<j\leq n}X_{ij}\right]e^{\alpha_n n^2\beta_1 t(H_1,X)}\right]\\
&= 2^{\binom{n}{2}}\mathbb{E}\left[\left[\left(\sum_{1\leq i<j\leq n}X_{ij}\right)^2-\sum_{1\leq i<j\leq n}X_{ij}\right]e^{2\alpha_n\beta_1\sum_{i<j}X_{ij}}\right]\\
&= 2^{\binom{n}{2}}\left(\frac{1}{4\alpha_n^2}\frac{\partial^2}{\partial\beta_1^2}\mathbb{E}\left[e^{2\alpha_n\beta_1\sum_{i<j}X_{ij}}\right]-\frac{1}{2\alpha_n}\frac{\partial}{\partial\beta_1}\mathbb{E}\left[e^{2\alpha_n\beta_1\sum_{i<j}X_{ij}}\right]\right)\\
&= 2^{\binom{n}{2}}\left(\frac{1}{4\alpha_n^2}\frac{\partial^2}{\partial\beta_1^2}\left(\frac{1+e^{2\alpha_n\beta_1}}{2}\right)^{\binom{n}{2}}-\frac{1}{2\alpha_n}\frac{\partial}{\partial\beta_1}\left(\frac{1+e^{2\alpha_n\beta_1}}{2}\right)^{\binom{n}{2}}\right)\\
&= e^{4\alpha_n\beta_1}\frac{(n+1)n(n-1)(n-2)}{4}\left(1+e^{2\alpha_n\beta_1}\right)^{\binom{n}{2}-2}.
\end{aligned}$$

From the assumptions $\lim_{n\to\infty} n^2 e^{2\alpha_n\beta_1}=0$ and $\lim_{n\to\infty}\frac{\alpha_n}{n}=0$, our claim follows. □

In Chatterjee and Dembo [8], when $|\beta_1^{(n)}|+\cdots+|\beta_k^{(n)}|$ does not grow too fast, then $\frac{\log Z_n}{n^2}$ can be approximated by

$$L_n := \sup_{x\in\mathcal{P}_n}\left\{\beta_1^{(n)}t(H_1,x)+\cdots+\beta_k^{(n)}t(H_k,x)-\frac{I(x)}{n^2}\right\}, \qquad (3.31)$$

where $\mathcal{P}_n := \{(x_{ij})_{1\leq i<j\leq n} : x_{ij}\in[0,1], 1\leq i<j\leq n\}$.

Chatterjee and Dembo [8] showed that

$$\begin{aligned}
-\frac{cB}{n} &\leq \frac{\log Z_n}{n^2}-L_n \qquad (3.32)\\
&\leq CB^{8/5}n^{-1/5}(\log n)^{1/5}\left(1+\frac{\log B}{\log n}\right)+CB^2n^{-1/2},
\end{aligned}$$

where $B := 1+|\beta_1^{(n)}|+\cdots+|\beta_k^{(n)}|$ and $c$ and $C$ are constants only depending on $H_1,\ldots,H_k$.

By considering $h(x,y)=x_{ij}$ for any $[\frac{i-1}{n},\frac{i}{n}]\times[\frac{j-1}{n},\frac{j}{n}]$, we have

$$L_n \leq \sup_{h:[0,1]^2\to[0,1],h(x,y)=h(y,x)}\left\{\sum_{p=1}^k\beta_p^{(n)}t(H_p,h)-\frac{1}{2}\iint_{[0,1]^2}I(h(x,y))dxdy\right\}. \qquad (3.33)$$

It was proved in Chatterjee and Diaconis [9] that when $H_p$, $p \geq 2$ are stars or when $\beta_p$'s are non-negative for any $p \geq 2$,

$$\sup_{h:[0,1]^2\to[0,1],h(x,y)=h(y,x)} \left\{\sum_{p=1}^{k} \beta_p^{(n)} t(H_p,h) - \frac{1}{2}\iint_{[0,1]^2} I(h(x,y))dxdy\right\} \tag{3.34}$$
$$= \sup_{0\leq x\leq 1}\left\{\alpha_n\beta_1 x + \alpha_n\beta_2 x^2 + \cdots + \alpha_n\beta_k x^k - \frac{1}{2}I(x)\right\}.$$

On the other hand, by considering $x_{ij} \equiv x$,

$$L_n \geq \sup_{0\leq x\leq 1}\left\{\alpha_n\beta_1 x + \alpha_n\beta_2 x^2 + \cdots + \alpha_n\beta_k x^k - \frac{1}{2}I(x)\right\}. \tag{3.35}$$

Therefore, for edge-star model or when $\beta_i$'s are non-negative, $2 \leq i \leq k$,

$$L_n = \sup_{0\leq x\leq 1}\left\{\alpha_n\beta_1 x + \alpha_n\beta_2 x^2 + \cdots + \alpha_n\beta_k x^k - \frac{1}{2}I(x)\right\}. \tag{3.36}$$

Let us analyze the edge and $p$-star model in more detail. The analysis also works for edge and $H_2$ model where the number of edges in $H_2$ is $p$ and $\beta_2 \geq 0$.

$$L_n = \sup_{0\leq x\leq 1}\left\{\alpha_n\beta_1 x + \alpha_n\beta_2 x^p - \frac{1}{2}I(x)\right\}. \tag{3.37}$$

We know that

$$\psi(\beta_1,\beta_2) = 0, \quad \text{when } (\beta_1,\beta_2) \in \{\beta_1+\beta_2 \leq 0\} \cap \{\beta_1 \leq 0\}. \tag{3.38}$$

This represents the regime in which we expect sparse random graphs.

**Theorem 7.** *Take $p \geq 3$. When $\beta_1 < 0$ and $\beta_1 + \beta_2 \leq 0$,*

$$\lim_{n\to\infty} \frac{L_n}{e^{2\alpha_n\beta_1}} = \frac{1}{2}. \tag{3.39}$$

*Moreover, if we further assume that $\lim_{n\to\infty} \frac{\alpha_n^{8/5}(\log n)^{1/5}e^{2\alpha_n|\beta_1|}}{n^{1/5}} = 0$, then*

$$\lim_{n\to\infty} \frac{\log Z_n}{n^2 e^{2\alpha_n\beta_1}} = \frac{1}{2}. \tag{3.40}$$

*When $\beta_1 = 0$ and $\beta_2 < 0$,*

$$\lim_{n\to\infty} \frac{L_n}{\gamma_n e^{\gamma_n}} = \frac{1-p}{2p}, \tag{3.41}$$

*where $\gamma_n$ is uniquely defined via the equation $2\alpha_n\beta_2 e^{(p-1)\gamma_n} p = \gamma_n$ and $\gamma_n \to -\infty$ as $n \to \infty$. Moreover, if we further assume that $\lim_{n\to\infty} \frac{\alpha_n^{8/5}(\log n)^{1/5}e^{|\gamma_n|}}{|\gamma_n| n^{1/5}} = 0$, then*

$$\lim_{n\to\infty} \frac{\log Z_n}{n^2 \gamma_n e^{\gamma_n}} = \frac{1-p}{2p}. \tag{3.42}$$

*Proof.* The optimization problem (3.37) has been well studied in Radin and Yin [25] and Aristoff and Zhu [2]. In Radin and Yin [25], it was proved that there exists a phase transition curve $\beta_2^{(n)} = q(\beta_1^{(n)})$ below which the maximizer of (3.37) is either unique or the smaller one. They also proved that $\beta_1^{(n)} + q(\beta_1^{(n)}) \to 0$ as $n \to \infty$. Moreover, in Aristoff and Zhu [2], it was shown that the phase transition curve always lies above the curve $\beta_1^{(n)} + \beta_2^{(n)} = 0$ when $p \geq 3$. Therefore, the optimizer in (3.37) is either the unique or the smaller solution to the equation

$$\alpha_n\beta_1 + \alpha_n\beta_2 p(x^*)^{p-1} = \frac{1}{2}\log\left(\frac{x^*}{1-x^*}\right). \tag{3.43}$$

In the sparse regime we consider, $(\beta_1, \beta_2) \in \{\beta_1 + \beta_2 \leq 0\} \cap \{\beta_1 \leq 0\}$, at least one of $|\alpha_n\beta_1|$ and $|\alpha_n\beta_2|$ go to $\infty$ as $n \to \infty$, so we must have $x^* \to 0$ as $n \to \infty$, see either [25] or [2]. We can rewrite (3.43) as

$$(1-x^*)e^{2\alpha_n\beta_2 p(x^*)^{p-1}} = \frac{x^*}{e^{2\alpha_n\beta_1}}. \tag{3.44}$$

For the edge and $p$-star model, if $\beta_2 \leq 0$, then we have $\frac{x^*}{e^{2\alpha_n\beta_1}} \leq 1$ and thus

$$2\alpha_n\beta_2 p(x^*)^{p-1} = 2\alpha_n\beta_2 p e^{2(p-1)\alpha_n\beta_1}\left(\frac{x^*}{e^{2\alpha_n\beta_1}}\right)^{p-1} \to 0, \tag{3.45}$$

as $n \to \infty$. Hence, we conclude that

$$\lim_{n\to\infty}\frac{x^*}{e^{2\alpha_n\beta_1}} = 1. \tag{3.46}$$

Now, let us consider the more general case $\beta_1 < 0$ and $\beta_1 + \beta_2 \leq 0$. By letting $y^* = \frac{x^*}{e^{2\alpha_n\beta_1}}$, we can rewrite (3.44) as

$$(1 - y^* e^{2\alpha_n\beta_1})e^{2\alpha_n\beta_2 e^{2\alpha_n\beta_1(p-1)}p(y^*)^{p-1}} = y^*. \tag{3.47}$$

Let us define $F(y) = (1 - ye^{2\alpha_n\beta_1})e^{2\alpha_n\beta_2 e^{2\alpha_n\beta_1(p-1)}py^{p-1}} - y$. We can compute that $F(0) = 1$ and

$$F(1) = (1 - e^{2\alpha_n\beta_1})e^{2\alpha_n\beta_2 e^{2\alpha_n\beta_1(p-1)}p} - 1 < 0, \tag{3.48}$$

for any sufficiently large $n$. This can be seen through the following argument. First we notice that when $\beta_1 + \beta_2 \leq 0$, $F(1) \leq (1 - e^{2\alpha_n\beta_1})e^{-2\alpha_n\beta_1 e^{2\alpha_n\beta_1(p-1)}p} - 1$. Take $z = e^{2\alpha_n\beta_1}$, it suffices to show that for $z$ sufficiently close to 0, $(1-z)z^{-pz^{p-1}} < 1$, which is equivalent to $pz^{p-1}\log z > \log(1-z)$. But this is clear when $p \geq 3$ since the derivative on the left tends to 0 whereas the derivative on the right tends to $-1$ as $z$ approaches $0+$. Moreover, we can compute that

$$\begin{aligned} F'(y) &= -e^{2\alpha_n\beta_1}e^{2\alpha_n\beta_2 e^{2\alpha_n\beta_1(p-1)}py^{p-1}} \\ &\quad + (1 - ye^{2\alpha_n\beta_1})2\alpha_n\beta_2 e^{2\alpha_n\beta_1(p-1)}p(p-1)y^{p-2}e^{2\alpha_n\beta_2 e^{2\alpha_n\beta_1(p-1)}py^{p-1}} - 1, \end{aligned} \tag{3.49}$$

and for sufficiently large $n$ we have $F'(y) < 0$ for any $0 \leq y \leq 1$. Since we know that $y^*$ is the unique or the smaller solution of $F(y) = 0$, we conclude that $y^*$ is the unique solution on the interval $(0, 1)$. Furthermore, we can also check that $F(1) \to 0$ and $F'(1) \to -1$ as $n \to \infty$. Hence, we conclude that $y^* \to 1$ as $n \to \infty$ and therefore

$$\lim_{n\to\infty}\frac{x^*}{e^{2\alpha_n\beta_1}} = 1. \tag{3.50}$$

Hence, by using (3.43), we get

$$\begin{aligned} L_n &= \alpha_n\beta_1 x^* + \alpha_n\beta_2(x^*)^p - \frac{1}{2}x^*\log x^* - \frac{1}{2}(1-x^*)\log(1-x^*) \\ &= \alpha_n\beta_1 x^* + \alpha_n\beta_2(x^*)^p - \frac{1}{2}x^*\log\left(\frac{x^*}{1-x^*}\right) - \frac{1}{2}\log(1-x^*) \\ &= \alpha_n\beta_2(x^*)^p(1-p) - \frac{1}{2}\log(1-x^*) \\ &= \alpha_n\beta_2(x^*)^p(1-p) + \frac{1}{2}x^* + O((x^*)^2). \end{aligned} \tag{3.51}$$

Therefore, $\lim_{n\to\infty}\frac{L_n}{x^*} = \frac{1}{2}$ and

$$\lim_{n\to\infty}\frac{L_n}{e^{2\alpha_n\beta_1}} = \frac{1}{2}. \tag{3.52}$$

Chatterjee and Dembo [8] showed that

$$\begin{aligned} &-\frac{c\alpha_n(|\beta_1|+|\beta_2|)}{n} \\ &\leq \frac{\log Z_n}{n^2} - L_n \\ &\leq C\alpha_n^{8/5}(|\beta_1|+|\beta_2|)^{8/5}n^{-1/5}(\log n)^{1/5}\left(1+\frac{\log\alpha_n+\log(|\beta_1|+|\beta_2|)}{\log n}\right) \\ &\qquad + C(|\beta_1|+|\beta_2|)^2\alpha_n^2 n^{-1/2}. \end{aligned} \tag{3.53}$$

Therefore, under the further assumption that $\lim_{n\to\infty}\frac{\alpha_n^{8/5}(\log n)^{1/5}e^{2\alpha_n|\beta_1|}}{n^{1/5}} = 0$, we have

$$\left|\frac{\log Z_n}{n^2 e^{2\alpha_n\beta_1}} - \frac{L_n}{e^{2\alpha_n\beta_1}}\right| \to 0, \qquad \text{as } n\to\infty, \tag{3.54}$$

and hence

$$\lim_{n\to\infty}\frac{\log Z_n}{n^2 e^{2\alpha_n\beta_1}} = \frac{1}{2}. \tag{3.55}$$

The boundary of phases are $\{\beta_1 = 0, \beta_2 < 0\}$ and $\{\beta_1+\beta_2 = 0, \beta_1 < 0\}$. Along the curve $\{\beta_1 = 0, \beta_2 < 0\}$, we have

$$\alpha_n\beta_2 p(x^*)^{p-1} = \frac{1}{2}\log\left(\frac{x^*}{1-x^*}\right). \tag{3.56}$$

Let $\gamma_n < 0$ be defined via the equation $2\alpha_n\beta_2 e^{(p-1)\gamma_n}p = \gamma_n$. First, let us check that $\gamma_n$ is well defined. Let us consider the function $F(x) = 2\alpha_n\beta_2 e^{(p-1)x}p - x$. Then $F(0) < 0$ and $F(-\infty) = \infty$. Moreover, $F'(x) = 2\alpha_n\beta_2 e^{(p-1)x}(p-1)p - 1 < 0$. Thus $F(x) = 0$ has a unique negative solution, which is denoted by $\gamma_n$. For any fixed $x < 0$, $F(x) = 2\alpha_n\beta_2 e^{(p-1)x}p - x < 0$ for sufficiently large $n$. Therefore, $\gamma_n \to -\infty$ as $n\to\infty$. We can rewrite (3.56) as

$$\left(1 - \frac{x^*}{e^{\gamma_n}}e^{\gamma_n}\right)e^{2\alpha_n\beta_2 p(\frac{x^*}{e^{\gamma_n}})^{p-1}e^{(p-1)\gamma_n} - \gamma_n} = \frac{x^*}{e^{\gamma_n}}, \tag{3.57}$$

which is equivalent to

$$\left(1 - \frac{x^*}{e^{\gamma_n}}e^{\gamma_n}\right)e^{\gamma_n((\frac{x^*}{e^{\gamma_n}})^{p-1}-1)} = \frac{x^*}{e^{\gamma_n}}. \tag{3.58}$$

Let $y^* = \frac{x^*}{e^{\gamma_n}}$. The equation reduces to

$$(1 - y^* e^{\gamma_n}) \, e^{\gamma_n ((y^*)^{p-1} - 1)} = y^*. \tag{3.59}$$

Let us define the function $G(y) = (1 - ye^{\gamma_n}) \, e^{\gamma_n (y^{p-1}-1)} - y$. We can check that $G(1) = -e^{\gamma_n}$ and $G(0) = e^{-\gamma_n}$. Moreover, $G'(y) = -e^{\gamma_n} e^{\gamma_n (y^{p-1}-1)} + (1 - ye^{\gamma_n}) \, \gamma_n (p-1) y^{p-2} e^{\gamma_n (y^{p-1}-1)} - 1 < 0$ on $(0,1)$. Therefore $y^*$ is the unique solution of $G(y) = 0$ on $(0,1)$. Since $\gamma_n \to -\infty$ as $n \to \infty$, we have $G(1) = -e^{\gamma_n} \to 0$ and $G'(1) = -e^{\gamma_n} + (1 - e^{\gamma_n})\gamma_n (p-1) - 1 \to -\infty$ as $n \to \infty$. Thus $y^* \to 1$ as $n \to \infty$. Hence, we conclude that

$$\lim_{n \to \infty} \frac{x^*}{e^{\gamma_n}} = 1, \tag{3.60}$$

where $\gamma_n$ is uniquely defined via the equation $2\alpha_n \beta_2 e^{(p-1)\gamma_n} p = \gamma_n$ and $\gamma_n \to -\infty$ as $n \to \infty$. Moreover, we have

$$\lim_{n \to \infty} \frac{L_n}{\gamma_n e^{\gamma_n}} = \frac{1-p}{2p} + \lim_{n \to \infty} \frac{1}{2\gamma_n} = \frac{1-p}{2p}. \tag{3.61}$$

If we further assume that $\lim_{n \to \infty} \frac{\alpha_n^{8/5} (\log n)^{1/5} e^{|\gamma_n|}}{|\gamma_n| n^{1/5}} = 0$, then

$$\left| \frac{\log Z_n}{n^2 \gamma_n e^{\gamma_n}} - \frac{L_n}{\gamma_n e^{\gamma_n}} \right| \to 0, \tag{3.62}$$

as $n \to \infty$ and

$$\lim_{n \to \infty} \frac{\log Z_n}{n^2 \gamma_n e^{\gamma_n}} = \frac{1-p}{2p}. \tag{3.63}$$

Along the curve $\{\beta_1 + \beta_2 = 0, \beta_1 < 0\}$, we have

$$\alpha_n \beta_1 - \alpha_n \beta_1 p (x^*)^{p-1} = \frac{1}{2} \log \left( \frac{x^*}{1 - x^*} \right), \tag{3.64}$$

and the asymptotic estimates follow exactly as in the case $\beta_1 + \beta_2 \leq 0$ and $\beta_1 < 0$ discussed earlier. □

**Remark 8.** *Intuitively, the edge density is given by*

$$\mathbb{P}^{(n)}(X_{ij} = 1) = \mathbb{E}^{(n)}[X_{ij}] = \frac{n^2}{2\binom{n}{2}} \frac{\partial}{\partial \beta_1^{(n)}} \frac{\log Z_n}{n^2} \text{ ``}\simeq\text{''} x^*, \tag{3.65}$$

*where $x^*$ is the maximizer in* (3.36) *and $x^* \simeq e^{2\alpha_n \beta_1}$. This is consistent with the results in Proposition 5.*

Let us recall that in Theorem 1, we proved that

$$\psi(\beta_1, \ldots, \beta_k) = \sup_{\substack{h:[0,1]^2 \to [0,1] \\ h(x,y) = h(y,x)}} \left\{ \beta_1 t(H_1, h) + \cdots + \beta_k t(H_k, h) \right\}, \tag{3.66}$$

where $H_1$ denotes a single edge. For the edge-(single)-star model, i.e., $H_1$ is an edge and $H_2$ is a $p$-star, it is easy to see that the optimal graphon is uniform.

**Proposition 9.** *For the edge-(single)-star model,*

$$\psi(\beta_1,\beta_2) = \sup_{\substack{h:[0,1]^2\to[0,1]\\ h(x,y)=h(y,x)}} \left\{ \beta_1 \int_0^1 \int_0^1 h(x,y)dxdy + \beta_2 \int_0^1 \left( \int_0^1 h(x,y)dy \right)^p dx \right\} \tag{3.67}$$
$$= \max_{0\leq x\leq 1} \{\beta_1 x + \beta_2 x^p\}.$$

*Proof.* By optimizing over constant $h$, it is clear that $\psi(\beta_1,\beta_2) \geq \max_{0\leq x\leq 1}\{\beta_1 x + \beta_2 x^p\}$. On the other hand,

$$\begin{aligned}
&\psi(\beta_1,\beta_2) \qquad (3.68)\\
&= \sup_{\substack{h:[0,1]^2\to[0,1]\\ h(x,y)=h(y,x)}} \left\{ \int_0^1 \left[ \beta_1 \left( \int_0^1 h(x,y)dy \right) + \beta_2 \left( \int_0^1 h(x,y)dy \right)^p \right] dx \right\} \\
&\leq \int_0^1 \left[ \max_{0\leq x\leq 1} \{\beta_1 x + \beta_2 x^p\} \right] dx \\
&= \max_{0\leq x\leq 1} \{\beta_1 x + \beta_2 x^p\}.
\end{aligned}$$

□

Next, let us consider the edge-triangle model, i.e., $H_1$ is an edge and $H_2$ is a triangle.

**Proposition 10.** *For the edge-triangle model,*

$$\begin{aligned}
&\psi(\beta_1,\beta_2) \qquad (3.69)\\
&= \sup_{\substack{h:[0,1]^2\to[0,1]\\ h(x,y)=h(y,x)}} \left\{ \beta_1 \iint_{[0,1]^2} h(x,y)dxdy + \beta_2 \iiint_{[0,1]^3} h(x,y)h(y,z)h(z,x)dxdydz \right\} \\
&= \begin{cases} \beta_1+\beta_2 & \textit{if } \beta_2 \geq 0 \textit{ and } \beta_1+\beta_2 \geq 0 \textit{ or if } \beta_2<0 \textit{ and } \beta_1+3\beta_2 \geq 0 \\ 0 & \textit{if } \beta_2 \geq 0 \textit{ and } \beta_1+\beta_2 < 0 \textit{ or if } \beta_2<0 \textit{ and } \beta_1 \leq 0 \\ \frac{l+1}{l+2}\beta_1 + \frac{l(l+1)}{(l+2)^2}\beta_2 & \textit{if } \beta_2<0 \textit{ and } a_l\beta_2 < \beta_1 \leq a_{l+1}\beta_2 \\ & \qquad \textit{where } a_l = -\frac{l(3l+5)}{(l+1)(l+2)} \textit{ for } l\geq 0 \end{cases}.
\end{aligned}$$

*Proof.* When $\beta_2 \geq 0$, by generalized Hölder's inequality,

$$\sup_{\substack{h:[0,1]^2\to[0,1]\\ h(x,y)=h(y,x)}} \left\{ \beta_1 \iint_{[0,1]^2} h(x,y)dxdy + \beta_2 \iiint_{[0,1]^3} h(x,y)h(y,z)h(z,x)dxdydz \right\} \tag{3.70}$$
$$\leq \sup_{\substack{h:[0,1]^2\to[0,1]\\ h(x,y)=h(y,x)}} \left\{ \beta_1 \iint_{[0,1]^2} h(x,y)dxdy + \beta_2 \iiint_{[0,1]^2} h(x,y)^3 dxdy \right\}$$
$$\leq \sup_{0\leq x\leq 1} \{\beta_1 x + \beta_2 x^3\}.$$

The other direction $\psi(\beta_1,\beta_2) \geq \sup_{0\leq x\leq 1}\{\beta_1 x + \beta_2 x^3\}$ is trivial.

When $\beta_2 < 0$ and $\beta_1 \leq 0$, it is clear that $\psi(\beta_1, \beta_2) = 0$. When $\beta_2 < 0$ and $\beta_1 > 0$, we need to do a more careful analysis. We have

$$\psi(\beta_1, \beta_2) = \sup_{0 \leq \epsilon \leq 1} \left\{ \beta_1 \epsilon + \beta_2 \tau(\epsilon) \right\}, \tag{3.71}$$

where

$$\tau(\epsilon) = \inf_{\substack{h:[0,1]^2 \to [0,1] \\ h(x,y)=h(y,x), \iint_{[0,1]^2} h(x,y)dxdy=\epsilon}} \iiint_{[0,1]^3} h(x,y)h(y,z)h(z,x)dxdydz \tag{3.72}$$

is the smallest possible triangle density given the edge density $\epsilon$. It is well known that $\tau(\epsilon) = 0$ for $0 \leq \epsilon \leq \frac{1}{2}$ and $\tau(\epsilon)$ is a nontrivial scallop curve defined as

$$\tau(\epsilon) = \frac{(\ell-1)(\ell - 2\sqrt{\ell(\ell - \epsilon(\ell+1))})(\ell + \sqrt{\ell(\ell-\epsilon(\ell+1))})^2}{\ell^2(\ell+1)^2}, \tag{3.73}$$

where $\ell = \lfloor 1/(1-\epsilon) \rfloor$ is the integer so that $\epsilon \in [1 - \frac{1}{\ell}, 1 - \frac{1}{\ell+1}]$ for $\frac{1}{2} \leq \epsilon \leq 1$, see e.g. Pikhurko and Raborov [24] and Razborov [29]. Therefore,

$$\psi(\beta_1, \beta_2) = \max\left\{ \frac{\beta_1}{2}, \sup_{\frac{1}{2} \leq \epsilon \leq 1} \{\beta_1 \epsilon + \beta_2 \tau(\epsilon)\} \right\} = \sup_{\frac{1}{2} \leq \epsilon \leq 1} \{\beta_1 \epsilon + \beta_2 \tau(\epsilon)\}. \tag{3.74}$$

Assume $\beta_1 = a\beta_2$ for some $a < 0$. Then

$$\psi(\beta_1, \beta_2) = \beta_2 \min_{\frac{1}{2} \leq \epsilon \leq 1} \{a\epsilon + \tau(\epsilon)\}. \tag{3.75}$$

The rest of the proof follows a similar line of reasoning as in the proof of Theorem 3.3 of [35]. The derivative of $a\epsilon + \tau(\epsilon)$ with respect to $\epsilon$ is given by

$$a + \frac{3(l-1)}{l(l+1)} \left( l + \sqrt{l(l - \epsilon(l+1))} \right). \tag{3.76}$$

It is a decreasing function of $\epsilon$ on each subinterval $[1 - \frac{1}{l}, 1 - \frac{1}{l+1}]$, hence we further conclude that the minimizer can only be obtained at the connection points $\epsilon_l = \frac{l}{l+1}$. Consider two adjacent connection points $(\epsilon_l, \tau(\epsilon_l))$ and $(\epsilon_{l+1}, \tau(\epsilon_{l+1}))$, where

$$(\epsilon_l, \tau(\epsilon_l)) = \left( \frac{l}{l+1}, \frac{l(l-1)}{(l+1)^2} \right) \text{ and } (\epsilon_{l+1}, \tau(\epsilon_{l+1})) = \left( \frac{l+1}{l+2}, \frac{l(l+1)}{(l+2)^2} \right). \tag{3.77}$$

Let $L_l$ be the line segment joining these two points. The slope of the line passing through $L_l$ is

$$\frac{l(3l+5)}{(l+1)(l+2)} = -a_l. \tag{3.78}$$

It is clear that $a_l$ is a decreasing function of $l$ and $a_l \to -3$ as $l \to \infty$. More importantly, if $a > a_l$, then $a\epsilon_l + \tau(\epsilon_l) < a\epsilon_{l+1} + \tau(\epsilon_{l+1})$; if $a = a_l$, then $a\epsilon_l + \tau(\epsilon_l) = a\epsilon_{l+1} + \tau(\epsilon_{l+1})$; and if $a < a_l$, then $a\epsilon_l + \tau(\epsilon_l) > a\epsilon_{l+1} + \tau(\epsilon_{l+1})$. Decreasing $a$ thus moves the location of the minimizer upward along the scallop curve, with sudden jumps happening at special angles $a = a_l$, which correspond to first-order phase transitions. To illustrate the phase diagram, we refer to Figure 1. □

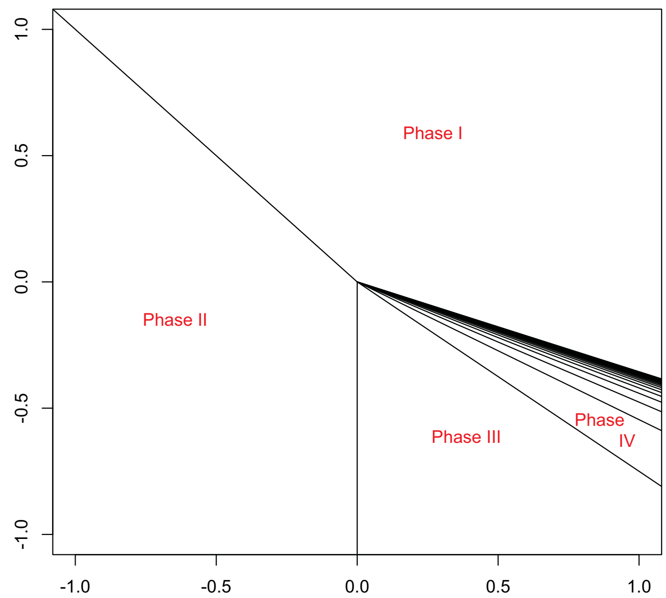


FIGURE 1. This is a plot of the phase diagram for the edge and triangle model. The horizontal axis denotes $\beta_1$ and the vertical axis denotes $\beta_2$. There are countably many phases with boundaries given by $\{\beta_1+\beta_2=0,\beta_1<0\}$, $\{\beta_1=0,\beta_2<0\}$, $\{\beta_1=a_l\beta_2,\beta_1>0\}$, $\ell=1,2,\ldots$, and $\{\beta_1=-3\beta_2,\beta_1>0\}$.

## 4. Directed Graphs

Consider directed graphs on $n$ vertices, where a graph is represented by a matrix $X=(X_{ij})_{1\leq i,j\leq n}$ with each $X_{ij}\in\{0,1\}$. Here, $X_{ij}=1$ means there is a directed edge from vertex $i$ to vertex $j$; otherwise, $X_{ij}=0$. Give the set of such graphs the probability

$$\mathbb{P}^{(n)}(X)=Z_n(\beta_1^{(n)},\beta_2^{(n)},\ldots,\beta_k^{(n)})^{-1}\exp\left[n^2\left(\sum_{p=1}^{k}\beta_p^{(n)}s_p(X)\right)\right], \tag{4.1}$$

where

$$s_p(X):=n^{-p-1}\sum_{1\leq i,j_1,j_2,\ldots,j_p\leq n}X_{ij_1}X_{ij_2}\cdots X_{ij_p}. \tag{4.2}$$

Here, $Z_n(\beta_1^{(n)},\beta_2^{(n)},\ldots,\beta_k^{(n)})$ is the appropriate normalization. Note $s_p(X)$, defined in (4.2), represents outward directed $p$-star homomorphism densities of $X$. When $p=1$, it represents the directed edge homomorphism density of $X$. It is easy to see that $s_p(X)$ has the alternative expression

$$s_p(X)=n^{-p-1}\sum_{i=1}^{n}\left(\sum_{j=1}^{n}X_{ij}\right)^p. \tag{4.3}$$

We allow $X_{ii}$ to equal 1 for ease of notation. It is not hard to see that without this simplification, our main results still hold.

We are interested in the sparse graph, i.e., the probability that there is a directed edge from vertex $i$ to vertex $j$ goes to 0 as $n \to \infty$. This requires that $\beta_p^{(n)} \to -\infty$, for some $1 \leq p \leq k$, as $n \to \infty$. Let us assume that

$$\beta_p^{(n)} = \beta_p \alpha_n, \quad p = 1, 2, \ldots, k, \tag{4.4}$$

where $\alpha_n \to \infty$ as $n \to \infty$.

**Theorem 11.**

$$\psi(\beta_1, \ldots, \beta_k) := \lim_{n\to\infty} \frac{1}{n^2\alpha_n} \log Z_n = \sup_{0\leq x\leq 1} \{\beta_1 x + \beta_2 x^2 + \cdots + \beta_k x^k\}, \tag{4.5}$$

*Proof.* It is therefore easy to compute that

$$\begin{aligned} Z_n(\beta_1^{(n)}, \ldots, \beta_k^{(n)}) &= 2^{n^2}\mathbb{E}\left[e^{n^2(\sum_{p=1}^k \beta_p^{(n)} s_p(X))}\right] \\ &= 2^{n^2}\mathbb{E}\left[e^{\sum_{p=1}^k \beta_p^{(n)} n^{-p+1} \sum_{i=1}^n (\sum_{j=1}^n X_{ij})^p}\right] \\ &= 2^{n^2}\left(\mathbb{E}\left[e^{\sum_{p=1}^k \beta_p^{(n)} n^{-p+1} (\sum_{j=1}^n X_{ij})^p}\right]\right)^n \\ &= \left[\sum_{j=0}^n \binom{n}{j} e^{n\sum_{p=1}^k \beta_p^{(n)} (\frac{j}{n})^p}\right]^n, \end{aligned} \tag{4.6}$$

where $\mathbb{E}$ denotes the expectation under which $X_{ij}$ are i.i.d. $\mathbb{P}(X_{ij} = 0) = \mathbb{P}(X_{ij} = 1) = \frac{1}{2}$.

On the one hand, we have

$$\begin{aligned} Z_n(\beta_1^{(n)}, \ldots, \beta_k^{(n)}) &\leq \left[2^n e^{\max_{0\leq j\leq n}\{n\sum_{p=1}^k \beta_p^{(n)}(\frac{j}{n})^p\}}\right]^n \\ &= 2^{n^2} e^{n^2\alpha_n \max_{0\leq j\leq n}\{\sum_{p=1}^k \beta_p (\frac{j}{n})^p\}} \\ &\leq 2^{n^2} e^{n^2\alpha_n \max_{0\leq x\leq 1}\{\sum_{p=1}^k \beta_p x^p\}}. \end{aligned} \tag{4.7}$$

Therefore, $\limsup_{n\to\infty} \frac{1}{n^2\alpha_n} \log Z_n \leq \max_{0\leq x\leq 1}\{\sum_{p=1}^k \beta_p x^p\}$.

On the other hand, suppose that $\max_{0\leq x\leq 1}\{\sum_{p=1}^k \beta_p x^p\} = \sum_{p=1}^k \beta_p x_*^p$ for some $0 \leq x_* \leq 1$. There exists such an $x_*$ since the maximum of a continuous function on a compact set is achieved though it may not be unique. Then, for any $\epsilon > 0$, for sufficiently large $n$, there exists some $j_* \in \{0, 1, \ldots, n\}$ so that

$$\sum_{p=1}^k \beta_p \left(\frac{j_*}{n}\right)^p \geq \max_{0\leq x\leq 1}\left\{\sum_{p=1}^k \beta_p x^p\right\} - \epsilon. \tag{4.8}$$

Therefore, we have

$$\begin{aligned} Z_n(\beta_1^{(n)}, \ldots, \beta_k^{(n)}) &\geq \left[e^{n\sum_{p=1}^k \beta_p^{(n)}(\frac{j_*}{n})^p}\right]^n \\ &\geq e^{n^2\alpha_n(\max_{0\leq x\leq 1}\{\sum_{p=1}^k \beta_p x^p\} - \epsilon)}. \end{aligned} \tag{4.9}$$

Therefore, $\liminf_{n\to\infty} \frac{1}{n^2\alpha_n} \log Z_n \geq \max_{0\leq x\leq 1}\{\sum_{p=1}^k \beta_p x^p\} - \epsilon$. Since it holds for any $\epsilon > 0$, together with the upper bound, we proved (4.5). ∎

**Remark 12.** *When the parameters $\beta_1, \beta_2, \dots, \beta_k$ are negative, it gives the sparse random graphs. We have already computed that*

$$(Z_n)^{\frac{1}{n}} = \sum_{j=0}^{n} \binom{n}{j} e^{n\alpha_n \sum_{p=1}^{k} \beta_p (\frac{j}{n})^p}, \tag{4.10}$$

*and* $\lim_{n\to\infty} \frac{1}{n^2\alpha_n} \log Z_n = 0$. *That indicates that when the parameters $\beta_1, \beta_2, \dots, \beta_k$ are negative, $\frac{1}{n^2\alpha_n}$ is not the optimal scaling for $\log Z_n$ as $n \to \infty$.*

*(i) When $\beta_1, \beta_2, \dots, \beta_k$ are negative,*

$$(Z_n)^{\frac{1}{n}} \leq \sum_{j=0}^{n} \binom{n}{j} e^{n\alpha_n \beta_1 (\frac{j}{n})} = (1 + e^{\beta_1 \alpha_n})^n. \tag{4.11}$$

*Therefore, we always have* $\limsup_{n\to\infty} (Z_n)^{\frac{1}{n^2}} \leq 1$. *On the other hand,* $Z_n^{\frac{1}{n}} \geq \binom{n}{0} e^{n\alpha_n \sum_{p=1}^{k} \beta_p (\frac{0}{n})^p} = 1$. *Thus,*

$$\lim_{n\to\infty} (Z_n)^{\frac{1}{n^2}} = 1. \tag{4.12}$$

*(ii) Furthermore, if we assume that* $\lim_{n\to\infty} n e^{\beta_1 \alpha_n} = 0$, *then*

$$\lim_{n\to\infty} (Z_n)^{\frac{1}{n}} = 1. \tag{4.13}$$

*(iii) If instead we assume that* $\lim_{n\to\infty} n e^{\beta_1 \alpha_n} = \lambda \in (0, \infty)$, *then as will be shown in Proposition 15,*

$$\lim_{n\to\infty} (Z_n)^{\frac{1}{n}} = e^{\lambda}. \tag{4.14}$$

*(iv) We can get more precise asymptotics. Let us assume that* $\lim_{n\to\infty} n e^{\beta_1 \alpha_n} = 0$ *and* $\lim_{n\to\infty} \frac{\alpha_n}{n} = 0$. *On the one hand,*

$$\frac{1}{n^2} \log Z_n \leq \log(1 + e^{\beta_1 \alpha_n}). \tag{4.15}$$

*Therefore,*

$$\limsup_{n\to\infty} \frac{\log Z_n}{n^2 e^{\beta_1 \alpha_n}} \leq 1. \tag{4.16}$$

*On the other hand,*

$$(Z_n)^{\frac{1}{n}} \geq \sum_{j=0}^{1} \binom{n}{j} e^{n\alpha_n \sum_{p=1}^{k} \beta_p (\frac{j}{n})^p} = 1 + n e^{n\alpha_n \sum_{p=1}^{k} \beta_p (\frac{1}{n})^p} \tag{4.17}$$

*Therefore,*

$$\begin{aligned} \liminf_{n\to\infty} \frac{\log Z_n}{n^2 e^{\beta_1 \alpha_n}} &\geq \liminf_{n\to\infty} \frac{\log(1 + n e^{n\alpha_n \sum_{p=1}^{k} \beta_p (\frac{1}{n})^p})}{n e^{\beta_1 \alpha_n}} \\ &= \liminf_{n\to\infty} \frac{\log(1 + n e^{n\alpha_n \sum_{p=1}^{k} \beta_p (\frac{1}{n})^p})}{n e^{n\alpha_n \sum_{p=1}^{k} \beta_p (\frac{1}{n})^p}} \frac{n e^{n\alpha_n \sum_{p=1}^{k} \beta_p (\frac{1}{n})^p}}{n e^{\beta_1 \alpha_n}} \\ &= 1, \end{aligned} \tag{4.18}$$

*where we used the assumptions that* $\lim_{n\to\infty} n e^{\beta_1 \alpha_n} = 0$ *and* $\lim_{n\to\infty} \frac{\alpha_n}{n} = 0$. *Hence, we conclude that*

$$\lim_{n\to\infty} \frac{\log Z_n}{n^2 e^{\beta_1 \alpha_n}} = 1. \tag{4.19}$$

For a sparse random graph, $\mathbb{P}^{(n)}(X_{ij}=1)\to 0$ as $n\to\infty$. It will be interesting to know how sparse the random graph is and how fast $\mathbb{P}^{(n)}(X_{ij}=1)$ converges to zero as $n\to\infty$.

**Proposition 13.** *Assume that $\beta_1,\ldots,\beta_k$ are all negative. Let us further assume that $\lim_{n\to\infty} ne^{\beta_1\alpha_n}=0$ and $\lim_{n\to\infty}\frac{\alpha_n}{n}=0$.*

$$\lim_{n\to\infty}\frac{\mathbb{P}^{(n)}(X_{1i}=1)}{e^{\beta_1\alpha_n}}=1. \tag{4.20}$$

*Proof.* By symmetry,

$$\begin{aligned}\mathbb{P}^{(n)}(X_{1i}=1)&=\mathbb{E}^{(n)}[X_{1i}] \qquad (4.21)\\ &=\frac{1}{n}\mathbb{E}^{(n)}\left[\sum_{i=1}^{n}X_{1i}\right]\\ &=\frac{1}{n}\frac{\mathbb{E}[\sum_{i=1}^{n}X_{1i}e^{\sum_{p=1}^{k}\beta_p^{(n)}n^{-p+1}(\sum_{i=1}^{n}X_{1i})^p}]}{\mathbb{E}[e^{\sum_{p=1}^{k}\beta_p^{(n)}n^{-p+1}(\sum_{i=1}^{n}X_{1i})^p}]}\\ &=\frac{1}{n}\frac{\sum_{j=0}^{n}\binom{n}{j}je^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{j}{n})^p}}{\sum_{j=0}^{n}\binom{n}{j}e^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{j}{n})^p}}.\end{aligned}$$

The denominator converges to 1 as $n\to\infty$ from the assumption $\lim_{n\to\infty}ne^{\beta_1\alpha_n}=0$. For the numerator, it is clear that

$$\sum_{j=0}^{n}\binom{n}{j}je^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{j}{n})^p}\geq ne^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{1}{n})^p}. \tag{4.22}$$

On the other hand,

$$\begin{aligned}\sum_{j=0}^{n}\binom{n}{j}je^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{j}{n})^p}&\leq\sum_{j=0}^{n}\binom{n}{j}je^{j\alpha_n\beta_1} \qquad (4.23)\\ &=\frac{1}{\alpha_n}\frac{\partial}{\partial\beta_1}\sum_{j=0}^{n}\binom{n}{j}e^{j\alpha_n\beta_1}\\ &=\frac{1}{\alpha_n}\frac{\partial}{\partial\beta_1}(1+e^{\alpha_n\beta_1})^n\\ &=e^{\alpha_n\beta_1}n(1+e^{\alpha_n\beta_1})^{n-1}.\end{aligned}$$

Finally, from the assumptions $\lim_{n\to\infty}ne^{\beta_1\alpha_n}=0$ and $\lim_{n\to\infty}\frac{\alpha_n}{n}=0$, we conclude that

$$\lim_{n\to\infty}\frac{\sum_{j=0}^{n}\binom{n}{j}je^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{j}{n})^p}}{ne^{\beta_1\alpha_n}}=1. \tag{4.24}$$

□

**Proposition 14.** *Assume that $\beta_1,\ldots,\beta_k$ are all negative. Let us further assume that $\lim_{n\to\infty} ne^{\beta_1\alpha_n}=0$ and $\lim_{n\to\infty}\frac{\alpha_n}{n}=0$.*

$$\lim_{n\to\infty}\frac{\mathbb{P}^{(n)}(X_{1i}=1,X_{1j}=1)}{e^{2\beta_1\alpha_n}}=1,\quad i\neq j. \tag{4.25}$$

*Proof.* By symmetry,

$$\begin{aligned}
&\mathbb{P}^{(n)}(X_{1i}=1,X_{1j}=1) \\
&=\mathbb{E}^{(n)}[X_{1i}X_{1j}] \\
&=\frac{1}{n(n-1)}\left[\mathbb{E}^{(n)}\left[\left(\sum_{i=1}^{n}X_{1i}\right)^2\right]-\mathbb{E}^{(n)}\left[\sum_{i=1}^{n}X_{1i}\right]\right] \\
&=\frac{1}{n(n-1)}\frac{\mathbb{E}\left[\left[(\sum_{i=1}^{n}X_{1i})^2-\sum_{i=1}^{n}X_{1i}\right]e^{\sum_{p=1}^{k}\beta_p^{(n)}n^{-p+1}(\sum_{i=1}^{n}X_{1i})^p}\right]}{\mathbb{E}[e^{\sum_{p=1}^{k}\beta_p^{(n)}n^{-p+1}(\sum_{i=1}^{n}X_{1i})^p}]} \\
&=\frac{1}{n(n-1)}\frac{\sum_{j=0}^{n}\binom{n}{j}(j^2-j)e^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{j}{n})^p}}{\sum_{j=0}^{n}\binom{n}{j}e^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{j}{n})^p}}.
\end{aligned} \tag{4.26}$$

Under the assumption $\lim_{n\to\infty}ne^{\beta_1\alpha_n}=0$, we have $\sum_{j=0}^{n}\binom{n}{j}e^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{j}{n})^p}\to 1$ as $n\to\infty$. Moreover, it is clear that

$$\begin{aligned}
\sum_{j=0}^{n}\binom{n}{j}(j^2-j)e^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{j}{n})^p}&\geq\binom{n}{2}(2^2-2)e^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{2}{n})^p} \\
&=n(n-1)e^{2\alpha_n\beta_1+n\alpha_n\sum_{p=2}^{k}\beta_p(\frac{2}{n})^p},
\end{aligned} \tag{4.27}$$

and

$$\begin{aligned}
&\sum_{j=0}^{n}\binom{n}{j}(j^2-j)e^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{j}{n})^p} \\
&\leq\sum_{j=0}^{n}\binom{n}{j}(j^2-j)e^{j\alpha_n\beta_1} \\
&=\frac{1}{\alpha_n^2}\frac{\partial^2}{\partial\beta_1^2}\sum_{j=0}^{n}\binom{n}{j}e^{j\alpha_n\beta_1}-\frac{1}{\alpha_n}\frac{\partial}{\partial\beta_1}\sum_{j=0}^{n}\binom{n}{j}e^{j\alpha_n\beta_1} \\
&=\frac{1}{\alpha_n^2}\frac{\partial^2}{\partial\beta_1^2}(1+e^{\alpha_n\beta_1})^n-\frac{1}{\alpha_n}\frac{\partial}{\partial\beta_1}(1+e^{\alpha_n\beta_1})^n \\
&=n(n-1)(1+e^{\alpha_n\beta_1})^{n-2}e^{2\alpha_n\beta_1}.
\end{aligned} \tag{4.28}$$

From the assumptions $\lim_{n\to\infty}ne^{\beta_1\alpha_n}=0$ and $\lim_{n\to\infty}\frac{\alpha_n}{n}=0$, our claim follows. □

For the directed exponential random graph model, under the assumptions $\beta_i$, $1\leq i\leq k$ are all negative and $\lim_{n\to\infty}ne^{\beta_1\alpha_n}=0$ and $\lim_{n\to\infty}\frac{\alpha_n}{n}=0$, we showed that $\lim_{n\to\infty}\frac{\mathbb{P}^{(n)}(X_{1i}=1)}{e^{\beta_1\alpha_n}}=1$. What if we have $\lim_{n\to\infty}ne^{\beta_1\alpha_n}=\lambda\in(0,\infty)$? If that is the case, then $\lim_{n\to\infty}\frac{\alpha_n}{n}=0$ is automatically satisfied. We have the following result.

**Proposition 15.** *Assume that $\beta_i$, $1\leq i\leq k$ are all negative and $\lim_{n\to\infty}ne^{\beta_1\alpha_n}=\lambda\in(0,\infty)$. Then, we have*

$$\lim_{n\to\infty}\frac{\mathbb{P}^{(n)}(X_{1i}=1)}{\lambda n^{-1}}=1. \tag{4.29}$$

$$\lim_{n\to\infty}\frac{\mathbb{P}^{(n)}(X_{1i}=1,X_{1j}=1)}{\lambda^2 n^{-2}}=1,\quad i\neq j. \tag{4.30}$$

*Moreover, the degree of any vertex is asymptotically Poisson with parameter* $\lambda$*, i.e.,*

$$\sum_{i=1}^{n}X_{1i}\to \mathit{Poisson}(\lambda), \tag{4.31}$$

*in distribution as* $n\to\infty$*.*

*Proof.* By symmetry,

$$\begin{aligned} n\mathbb{P}^{(n)}(X_{1i}=1)&=\mathbb{E}^{(n)}\left[\sum_{i=1}^{n}X_{1i}\right] \\ &=\frac{\sum_{j=0}^{n}j\binom{n}{j}e^{\alpha_n\beta_1 j+\sum_{p=2}^{k}\alpha_n\beta_p\frac{j^p}{n^{p-1}}}}{\sum_{j=0}^{n}\binom{n}{j}e^{\alpha_n\beta_1 j+\alpha_n\sum_{p=2}^{k}\beta_p\frac{j^p}{n^{p-1}}}}. \end{aligned}\tag{4.32}$$

First, let us analyze the denominator. Since we assumed that $\lim_{n\to\infty}ne^{\alpha_n\beta_1}=\lambda$, for any fixed $M$,

$$\begin{aligned} \sum_{j=0}^{n}\binom{n}{j}e^{\alpha_n\beta_1 j+\alpha_n\sum_{p=2}^{k}\beta_p\frac{j^p}{n^{p-1}}}&\geq\sum_{j=0}^{M}\binom{n}{j}e^{\alpha_n\beta_1 j+\alpha_n\sum_{p=2}^{k}\beta_p\frac{j^p}{n^{p-1}}} \\ &\to\sum_{j=0}^{M}\frac{\lambda^j}{j!}, \end{aligned}\tag{4.33}$$

as $n\to\infty$. Since it's true for any $M$, let $M\to\infty$, and we obtain an asymptotic lower bound $\sum_{j=0}^{\infty}\frac{\lambda^j}{j!}=e^{\lambda}$.

Moreover, since $\beta_p$'s are negative,

$$\begin{aligned} \sum_{j=0}^{n}\binom{n}{j}e^{\alpha_n\beta_1 j+\alpha_n\sum_{p=2}^{k}\beta_p\frac{j^p}{n^{p-1}}}&\leq\sum_{j=0}^{n}\binom{n}{j}e^{\alpha_n\beta_1 j} \\ &=(1+e^{\alpha_n\beta_1})^n \\ &\to e^{\lambda}, \end{aligned}\tag{4.34}$$

as $n\to\infty$.

Next, let us analyze the numerator. On the one hand, for any fixed $M$,

$$\begin{aligned} \sum_{j=0}^{n}j\binom{n}{j}e^{\alpha_n\beta_1 j+\sum_{p=2}^{k}\alpha_n\beta_p\frac{j^p}{n^{p-1}}}&\geq\sum_{j=0}^{M}j\binom{n}{j}e^{\alpha_n\beta_1 j+\sum_{p=2}^{k}\alpha_n\beta_p\frac{j^p}{n^{p-1}}} \\ &\to\sum_{j=0}^{M-1}\frac{\lambda^{j+1}}{j!}, \end{aligned}\tag{4.35}$$

as $n\to\infty$. Since it's true for any $M$, let $M\to\infty$, and we obtain an asymptotic lower bound $\sum_{j=0}^{\infty}\frac{\lambda^{j+1}}{j!}=\lambda e^{\lambda}$.

On the other hand,

$$\begin{aligned}\sum_{j=0}^{n} j\binom{n}{j} e^{\alpha_n\beta_1 j+\sum_{p=2}^{k}\alpha_n\beta_p\frac{j^p}{n^{p-1}}} &\leq \sum_{j=0}^{n} j\binom{n}{j} e^{\alpha_n\beta_1 j} \\ &= \frac{1}{\alpha_n}\frac{\partial}{\partial\beta_1}(1+e^{\alpha_n\beta_1})^n \\ &= ne^{\alpha_n\beta_1}(1+e^{\alpha_n\beta_1})^{n-1} \\ &\to \lambda e^{\lambda}.\end{aligned} \tag{4.36}$$

Again by symmetry,

$$\begin{aligned}n(n-1)\mathbb{P}^{(n)}(X_{1i}=1, X_{1j}=1) &= \mathbb{E}^{(n)}\left[\left(\sum_{i=1}^{n} X_{1i}\right)^2\right] - \mathbb{E}^{(n)}\left[\sum_{i=1}^{n} X_{1i}\right] \\ &= \frac{\sum_{j=0}^{n}(j^2-j)\binom{n}{j}e^{\alpha_n\beta_1 j+\sum_{p=2}^{k}\alpha_n\beta_p\frac{j^p}{n^{p-1}}}}{\sum_{j=0}^{n}\binom{n}{j}e^{\alpha_n\beta_1 j+\sum_{p=2}^{k}\alpha_n\beta_p\frac{j^p}{n^{p-1}}}}.\end{aligned} \tag{4.37}$$

The denominator converges to $e^{\lambda}$ as $n\to\infty$ from the assumption $\lim_{n\to\infty} ne^{\beta_1\alpha_n} = \lambda$. For the numerator, it is clear that for any fixed $M$,

$$\begin{aligned}\sum_{j=0}^{n}(j^2-j)\binom{n}{j}e^{\alpha_n\beta_1 j+\sum_{p=2}^{k}\alpha_n\beta_p\frac{j^p}{n^{p-1}}} &\geq \sum_{j=0}^{M}(j^2-j)\binom{n}{j}e^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{j}{n})^p} \\ &\to \sum_{j=0}^{M-2}\frac{\lambda^{j+2}}{j!},\end{aligned} \tag{4.38}$$

as $n\to\infty$. Since it's true for any $M$, let $M\to\infty$, and we obtain an asymptotic lower bound $\sum_{j=0}^{\infty}\frac{\lambda^{j+2}}{j!} = \lambda^2 e^{\lambda}$.

On the other hand,

$$\begin{aligned}\sum_{j=0}^{n}(j^2-j)\binom{n}{j}e^{\alpha_n\beta_1 j+\sum_{p=2}^{k}\alpha_n\beta_p\frac{j^p}{n^{p-1}}} &\leq \sum_{j=0}^{n}(j^2-j)\binom{n}{j}e^{\alpha_n\beta_1 j} \\ &= \frac{1}{\alpha_n^2}\frac{\partial^2}{\partial\beta_1^2}(1+e^{\alpha_n\beta_1})^n - \frac{1}{\alpha_n}\frac{\partial}{\partial\beta_1}(1+e^{\alpha_n\beta_1})^n \\ &= n(n-1)(1+e^{\alpha_n\beta_1})^{n-2}e^{2\alpha_n\beta_1} \\ &\to \lambda^2 e^{\lambda}.\end{aligned} \tag{4.39}$$

Lastly, for any fixed $j\in\mathbb{N}\cup\{0\}$,

$$\begin{aligned}\mathbb{P}^{(n)}\left(\sum_{i=1}^{n} X_{1i} = j\right) &= \mathbb{E}^{(n)}\left[1_{\sum_{i=1}^{n} X_{1i}=j}\right] \\ &= \frac{\binom{n}{j}e^{\alpha_n\beta_1 j+\sum_{p=2}^{k}\alpha_n\beta_p\frac{j^p}{n^{p-1}}}}{\sum_{j=0}^{n}\binom{n}{j}e^{\alpha_n\beta_1 j+\alpha_n\sum_{p=2}^{k}\beta_p\frac{j^p}{n^{p-1}}}}.\end{aligned} \tag{4.40}$$

Under the assumption $\lim_{n\to\infty} ne^{\beta_1\alpha_n} = \lambda$, the denominator converges to $e^{\lambda}$ and the numerator converges to $\frac{\lambda^j}{j!}$ as $n\to\infty$. Hence, we proved the desired result.

□

Another natural question to ask is what if $\liminf_{n\to\infty}\frac{\alpha_n}{n}>0$? If that is the case, then $\lim_{n\to\infty} ne^{\beta_1\alpha_n}=0$ is automatically satisfied.

**Proposition 16.** *Assume that $\beta_1,\ldots,\beta_k$ are all negative and $\liminf_{n\to\infty}\frac{\alpha_n}{n}>\frac{\log 2}{|\beta_1|}$, then*

$$\lim_{n\to\infty}\frac{\mathbb{P}^{(n)}(X_{1i}=1)}{e^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{1}{n})^p}}=1. \tag{4.41}$$

*Proof.* Let us recall that

$$\mathbb{P}^{(n)}(X_{1i}=1)=\frac{1}{n}\frac{\sum_{j=0}^{n}\binom{n}{j}je^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{j}{n})^p}}{\sum_{j=0}^{n}\binom{n}{j}e^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{j}{n})^p}}. \tag{4.42}$$

Under the implied assumption $\lim_{n\to\infty} ne^{\beta_1\alpha_n}=0$, the denominator

$$\lim_{n\to\infty}\sum_{j=0}^{n}\binom{n}{j}e^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{j}{n})^p}=1. \tag{4.43}$$

For the numerator, on the one hand,

$$\sum_{j=0}^{n}\binom{n}{j}je^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{j}{n})^p}\geq\binom{n}{1}e^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{1}{n})^p}=ne^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{1}{n})^p}. \tag{4.44}$$

On the other hand, since $\beta_p$'s are negative,

$$\sum_{j=2}^{n}\binom{n}{j}je^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{j}{n})^p}\leq\sum_{j=2}^{n}\binom{n}{j}je^{n\alpha_n\beta_1\frac{2}{n}}\leq 2^n ne^{2\alpha_n\beta_1}. \tag{4.45}$$

Moreover,

$$\frac{2^n ne^{2\alpha_n\beta_1}}{ne^{n\alpha_n\sum_{p=1}^{k}\beta_p(\frac{1}{n})^p}}=\frac{2^n e^{2\alpha_n\beta_1}}{e^{\alpha_n\beta_1+\sum_{p=2}^{k}\alpha_n\frac{\beta_p}{n^{p-1}}}}=e^{n[\log 2+\frac{\alpha_n}{n}\beta_1-\frac{\alpha_n}{n}\sum_{p=2}^{k}\frac{\beta_p}{n^{p-1}}]}\to 0, \tag{4.46}$$

as $n\to\infty$ under the assumption $\liminf_{n\to\infty}\frac{\alpha_n}{n}>\frac{\log 2}{|\beta_1|}$. Hence, we proved the desired result. □

Let us recall that in Theorem 11, we proved that

$$\psi(\beta_1,\ldots,\beta_k):=\lim_{n\to\infty}\frac{1}{n^2\alpha_n}\log Z_n=\sup_{0\leq x\leq 1}\{\beta_1 x+\beta_2 x^2+\cdots+\beta_k x^k\}. \tag{4.47}$$

Now, let us consider the edge-(single)-star model, i.e. $H_1$ is an edge and $H_2$ is an outward $p$-star. Thus, by Theorem 11,

$$\psi(\beta_1,\beta_2)=\sup_{0\leq x\leq 1}\{\beta_1 x+\beta_2 x^p\}.$$

**Proposition 17.**

$$\psi(\beta_1,\beta_2)=\begin{cases}\left[p^{\frac{-1}{p-1}}-p^{\frac{-p}{p-1}}\right]\frac{\beta_1^{\frac{p}{p-1}}}{(-\beta_2)^{\frac{1}{p-1}}} & \text{if } \beta_1+\beta_2 p<0 \text{ and } \beta_1>0\\ 0 & \text{if } \beta_1+\beta_2\leq 0 \text{ and } \beta_1\leq 0\\ \beta_1+\beta_2 & \text{otherwise}\end{cases}. \tag{4.48}$$

*There are second-order phase transitions across the phase transition curves $\{\beta_1=0,\beta_2<0\}$ and $\{\beta_1+p\beta_2=0,\beta_2<0\}$. There is a first-order phase transition across the phase transition curve $\{\beta_1+\beta_2=0,\beta_2>0\}$.*

*Proof.* Let us define

$$\ell(x) = \beta_1 x + \beta_2 x^p, \qquad 0 \leq x \leq 1. \tag{4.49}$$

Then, $\ell''(x) = \beta_2 p(p-1)x^{p-2}$ is always non-negative if $\beta_2$ is and is always non-positive if $\beta_2$ is. Therefore, the local maximizer on the interval $[0,1]$ is unique. Note that if $\beta_1\beta_2 < 0$, then, $\ell'(x) = \beta_1 + \beta_2 p x^{p-1} = 0$ gives $x = \left(\frac{-\beta_1}{\beta_2 p}\right)^{\frac{1}{p-1}}$. The maximizer should be achieved in the set $\left\{0, 1, \left(\frac{-\beta_1}{\beta_2 p}\right)^{\frac{1}{p-1}}\right\}$. The maximizer is achieved at $\left(\frac{-\beta_1}{\beta_2 p}\right)^{\frac{1}{p-1}}$ only if $\ell''(x) < 0$ and $\left(\frac{-\beta_1}{\beta_2 p}\right)^{\frac{1}{p-1}} < 1$ which is equivalent to $\beta_1 > 0$ and $\beta_1 + \beta_2 p < 0$. One can compute that

$$\ell\left(\left(\frac{-\beta_1}{\beta_2 p}\right)^{\frac{1}{p-1}}\right) = \beta_1\left(\frac{-\beta_1}{\beta_2 p}\right)^{\frac{1}{p-1}} + \beta_2\left(\frac{-\beta_1}{\beta_2 p}\right)^{\frac{p}{p-1}} = \left[p^{\frac{-1}{p-1}} - p^{\frac{-p}{p-1}}\right]\frac{\beta_1^{\frac{p}{p-1}}}{(-\beta_2)^{\frac{1}{p-1}}} \tag{4.50}$$

Otherwise, the maximizer is achieved at either 0 or 1. To summarize, we have

$$\psi(\beta_1,\beta_2) = \begin{cases} \left[p^{\frac{-1}{p-1}} - p^{\frac{-p}{p-1}}\right]\frac{\beta_1^{\frac{p}{p-1}}}{(-\beta_2)^{\frac{1}{p-1}}} & \text{if } \beta_1 + \beta_2 p < 0 \text{ and } \beta_1 > 0 \\ 0 & \text{if } \beta_1 + \beta_2 \leq 0 \text{ and } \beta_1 \leq 0 \\ \beta_1 + \beta_2 & \text{otherwise} \end{cases}. \tag{4.51}$$

There are three phases, with boundaries given by

(1) $\{\beta_1 = 0, \beta_2 < 0\}$.
(2) $\{\beta_1 + \beta_2 = 0, \beta_2 > 0\}$.
(3) $\{\beta_1 + p\beta_2 = 0, \beta_2 < 0\}$.

The three phases are illustrated in Figure 2.

As $\beta_1 \to 0^+$,

$$\left[p^{\frac{-1}{p-1}} - p^{\frac{-p}{p-1}}\right]\frac{\beta_1^{\frac{p}{p-1}}}{(-\beta_2)^{\frac{1}{p-1}}} \to 0. \tag{4.52}$$

Therefore $\psi(\beta_1,\beta_2)$ is continuous across the phase transition curve

$$\{\beta_1 = 0, \beta_2 < 0\}. \tag{4.53}$$

Moroever,

$$\frac{\partial}{\partial\beta_1}\left[p^{\frac{-1}{p-1}} - p^{\frac{-p}{p-1}}\right]\frac{\beta_1^{\frac{p}{p-1}}}{(-\beta_2)^{\frac{1}{p-1}}} = \left[p^{\frac{-1}{p-1}} - p^{\frac{-p}{p-1}}\right]\frac{\beta_1^{\frac{1}{p-1}}}{(-\beta_2)^{\frac{1}{p-1}}}\frac{p}{p-1} \to 0, \tag{4.54}$$

as $\beta_1 \to 0^+$ and

$$\frac{\partial^2}{\partial\beta_1^2}\left[p^{\frac{-1}{p-1}} - p^{\frac{-p}{p-1}}\right]\frac{\beta_1^{\frac{p}{p-1}}}{(-\beta_2)^{\frac{1}{p-1}}} = \left[p^{\frac{-1}{p-1}} - p^{\frac{-p}{p-1}}\right]\frac{\beta_1^{\frac{1}{p-1}-1}}{(-\beta_2)^{\frac{1}{p-1}}}\frac{p}{p-1}\frac{1}{p-1} \to +\infty, \tag{4.55}$$

as $\beta_1 \to 0^+$. Hence, there is a second-order phase transition across the phase transition curve $\{\beta_1 = 0, \beta_2 < 0\}$. On the other hand, we can check that $\frac{\partial^k}{\partial\beta_2^k}\psi = 0$ as $\beta_1 \to 0^+$ for any fixed $\beta_2 < 0$ for any $k = 0, 1, 2, \ldots$.

There is a first-order phase transition, i.e. $\frac{\partial}{\partial\beta_2}\psi(\beta_1,\beta_2)$ is not continuous when the parameter $\beta_2$ crosses the phase transition curve

$$\{\beta_1+\beta_2=0,\beta_2>0\}. \tag{4.56}$$

First, $\psi(\beta_1,\beta_2)$ is continuous across this curve so there is no zeroth-order phase transition. Second, $\frac{\partial}{\partial\beta_2}\psi$ changes from 1 to 0 when $\beta_2$ decreases across the curve. Hence, there is a first-order phase transition across the curve.

Finally, when $\beta_2=-\frac{\beta_1}{p}$, we have

$$\beta_1+\beta_2=(1-p)\beta_2, \tag{4.57}$$

and

$$\left[\left[p^{\frac{-1}{p-1}}-p^{\frac{-p}{p-1}}\right]\frac{\beta_1^{\frac{p}{p-1}}}{(-\beta_2)^{\frac{1}{p-1}}}\right]=(1-p)\beta_2. \tag{4.58}$$

Therefore $\psi(\beta_1,\beta_2)$ is continuous across the curve $\{\beta_1+p\beta_2=0,\beta_2<0\}$.

Moreover, one can compute that,

$$\begin{aligned}&\frac{\partial}{\partial\beta_2}\left[\left[p^{\frac{-1}{p-1}}-p^{\frac{-p}{p-1}}\right]\frac{\beta_1^{\frac{p}{p-1}}}{(-\beta_2)^{\frac{1}{p-1}}}\right]\Bigg|_{\beta_2=-\frac{\beta_1}{p}} \\ &=\left[p^{\frac{-1}{p-1}}-p^{\frac{-p}{p-1}}\right]\frac{\beta_1^{\frac{p}{p-1}}}{(-\beta_2)^{\frac{p}{p-1}}}\left(\frac{1}{p-1}\right)\Bigg|_{\beta_2=-\frac{\beta_1}{p}} \\ &=1.\end{aligned} \tag{4.59}$$

Hence, $\frac{\partial}{\partial\beta_2}\psi$ is continuous across the curve $\{\beta_1+p\beta_2=0,\beta_2<0\}$. Similarly, one can check that

$$\frac{\partial}{\partial\beta_1}\left[\left[p^{\frac{-1}{p-1}}-p^{\frac{-p}{p-1}}\right]\frac{\beta_1^{\frac{p}{p-1}}}{(-\beta_2)^{\frac{1}{p-1}}}\right]\Bigg|_{\beta_2=-\frac{\beta_1}{p}}=1. \tag{4.60}$$

Hence, $\frac{\partial}{\partial\beta_1}\psi$ is continuous across the curve $\{\beta_1+p\beta_2=0,\beta_2<0\}$.

On the other hand,

$$\begin{aligned}&\frac{\partial^2}{\partial\beta_2^2}\left[\left[p^{\frac{-1}{p-1}}-p^{\frac{-p}{p-1}}\right]\frac{\beta_1^{\frac{p}{p-1}}}{(-\beta_2)^{\frac{1}{p-1}}}\right]\Bigg|_{\beta_2=-\frac{\beta_1}{p}} \\ &=\left[p^{\frac{-1}{p-1}}-p^{\frac{-p}{p-1}}\right]\frac{\beta_1^{\frac{p}{p-1}}}{(-\beta_2)^{\frac{2p-1}{p-1}}}\left(\frac{1}{p-1}\right)\left(\frac{p}{p-1}\right)\Bigg|_{\beta_2=-\frac{\beta_1}{p}} \\ &=\frac{\frac{p}{p-1}}{-\beta_2}.\end{aligned} \tag{4.61}$$

Similarly,

$$\frac{\partial^2}{\partial\beta_1^2}\left[\left[p^{\frac{-1}{p-1}}-p^{\frac{-p}{p-1}}\right]\frac{\beta_1^{\frac{p}{p-1}}}{(-\beta_2)^{\frac{1}{p-1}}}\right]\Bigg|_{\beta_2=-\frac{\beta_1}{p}}=\frac{-1}{p(p-1)\beta_2}. \tag{4.62}$$

Hence there is second-order phase transition across the curve $\{\beta_1+p\beta_2=0,\beta_2<0\}$. □

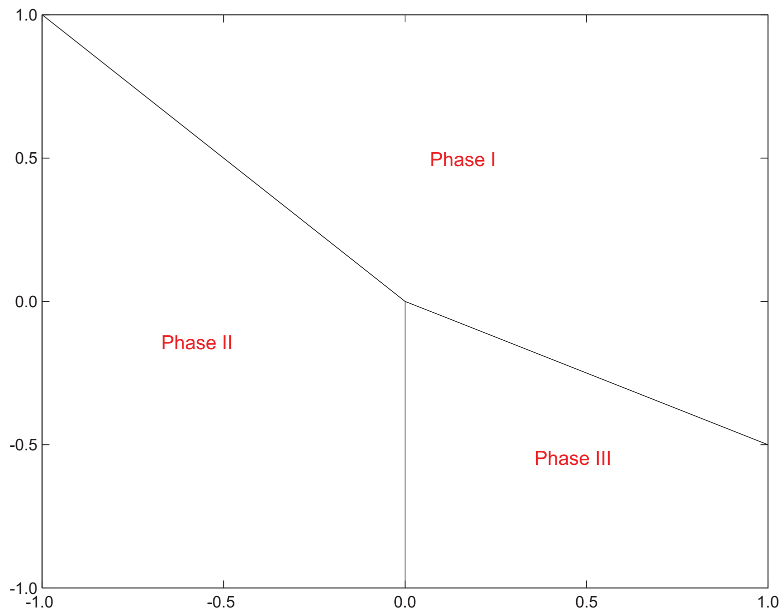


FIGURE 2. This is a plot of the phase diagram for the edge and outward $p$-star model, where $p = 2$. The horizontal axis denotes $\beta_1$ and the vertical axis denotes $\beta_2$. In Phase I, $\psi(\beta_1, \beta_2) = \beta_1 + \beta_2$; in Phase II, $\psi(\beta_1, \beta_2) = 0$; in Phase III, $\psi(\beta_1, \beta_2) = \left[p^{\frac{-1}{p-1}} - p^{\frac{-p}{p-1}}\right] \frac{\beta_1^{\frac{p}{p-1}}}{(-\beta_2)^{\frac{1}{p-1}}}$. The boundaries between the phases are given by $\{\beta_1 = 0, \beta_2 < 0\}$, $\{\beta_1 + \beta_2 = 0, \beta_2 > 0\}$, and $\{\beta_1 + p\beta_2 = 0, \beta_2 < 0\}$. There are second-order phase transitions between Phase II and Phase III, and between Phase III and Phase I. There is a first-order phase transition between Phase I and Phase II.

**Remark 18.** *More generally, we can consider* $\sup_{0\le x\le 1}\{\beta_1 x^q + \beta_2 x^p\}$, $q < p$. *If* $\beta_1 \ge 0$ *and* $\beta_2 \ge 0$, $\psi(\beta_1, \beta_2) = \beta_1 + \beta_2$. *If* $\beta_1 \le 0$ *and* $\beta_2 \le 0$, $\psi(\beta_1, \beta_2) = 0$. *If* $\beta_1\beta_2 < 0$, $\ell'(x) = 0$ *is achieved at* $x = \left(\frac{\beta_1 q}{-\beta_2 p}\right)^{\frac{1}{p-q}}$. *If* $\beta_1 < 0$, $\beta_2 > 0$, *then* $\ell(x)$ *decreases for small positive* $x$ *and* $x = \left(\frac{\beta_1 q}{-\beta_2 p}\right)^{\frac{1}{p-q}}$ *cannot be a maximizer. Thus when* $\beta_1 < 0$ *and* $\beta_2 > 0$, $\psi(\beta_1, \beta_2) = 0$ *if* $\beta_1 + \beta_2 \le 0$ *and* $\psi(\beta_1, \beta_2) = \beta_1 + \beta_2$ *if* $\beta_1 + \beta_2 > 0$. *If* $\beta_1 > 0$ *and* $\beta_2 < 0$ *and* $\frac{\beta_1 q}{-\beta_2 p} \ge 1$, *then* $\psi(\beta_1, \beta_2) = \beta_1 + \beta_2$. *If* $\beta_1 > 0$ *and* $\beta_2 < 0$ *and* $\frac{\beta_1 q}{-\beta_2 p} < 1$, *then* $\psi(\beta_1, \beta_2) = \beta_1 \left(\frac{\beta_1 q}{-\beta_2 p}\right)^{\frac{q}{p-q}} + \beta_2 \left(\frac{\beta_1 q}{-\beta_2 p}\right)^{\frac{p}{p-q}}$.

*To summarize, we have*

$$\psi(\beta_1,\beta_2)=\begin{cases}\left[\left(\frac{q}{p}\right)^{\frac{q}{p-q}}-\left(\frac{q}{p}\right)^{\frac{p}{p-q}}\right]\frac{\beta_1^{\frac{p}{p-q}}}{(-\beta_2)^{\frac{q}{p-q}}} & \textit{if } (\beta_1,\beta_2)\in\{\beta_1 q+\beta_2 p<0\}\cap\{\beta_1>0\}\\ 0 & \textit{if } (\beta_1,\beta_2)\in\{\beta_1+\beta_2<0\}\cap\{\beta_1\leq 0\}\\ \beta_1+\beta_2 & \textit{otherwise}\end{cases}. \tag{4.63}$$

## Acknowledgements

Mei Yin's research was partially supported by NSF grant DMS-1308333.

## References

[1] Aldous, D. (1981). Representations for partially exchangeable arrays of random variables. *J. Multivariate Anal.* **11**, 581-598.
[2] Aristoff, D. and L. Zhu. (2014). On the phase transition curve in a directed exponential random graph model. Preprint.
[3] Aristoff, D. and L. Zhu. (2014). Asymptotic structure and singularities in constrained directed graphs. Preprint.
[4] Besag, J. (1975). Statistical analysis of non-lattice data. *J. R. Stat. Soc., Ser. D. Stat.* **24**, 179-195.
[5] Borgs, C., Chayes, J., Lovász, L., Sós, V.T. and K. Vesztergombi. (2006). Counting graph homomorphisms. In: Klazar, M., Kratochvil, J., Loebl, M., Thomas, R., Valtr, P. (eds.) *Topics in Discrete Mathematics Volume 26*, pp. 315-371. Springer, Berlin.
[6] Borgs, C., Chayes, J.T., Lovász, L., Sós, V.T. and K. Vesztergombi. (2008). Convergent sequences of dense graphs I. Subgraph frequencies, metric properties and testing. *Adv. Math.* **219**, 1801-1851.
[7] Borgs, C., Chayes, J.T., Lovász, L., Sós, V.T. and K. Vesztergombi. (2012). Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Ann. of Math.* **176**, 151-219.
[8] Chatterjee, S. and A. Dembo. Nonlinear large deviations. Preprint, 2014.
[9] Chatterjee, S. and P. Diaconis. (2013). Estimating and understanding exponential random graph models. *Ann. Statist.* **41**, 2428-2461.
[10] Chatterjee, S. and S. R. S. Varadhan. (2011). The large deviation principle for the Erdős-Rényi random graph. *European. J. Combin.* **32**, 1000-1017.
[11] Fienberg, S. E. (2010). Introduction to papers on the modeling and analysis of network data. *Ann. Appl. Statist.* **4**, 1-4.
[12] Fienberg, S. E. (2010). Introduction to papers on the modeling and analysis of network data–II. *Ann. Appl. Statist.* **4**, 533-534.
[13] Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A. and C. D. Bloomfield. (1999). Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. *Science* **286**, 531-537.
[14] Guyon, I., Weston, J., Barnhill, S. and V. Vapnik. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning* **46**, 389-422.
[15] Hoover, D. (1982). Row-column exchangeability and a generalized model for probability. In: Koch, G., Spizzichino, F. (eds.) *Exchangeability in Probability and Statistics*, pp. 281-291. North-Holland, Amsterdam.
[16] Hromádka, T., DeWeese, M. and A. Zador. (2008). Sparse representation of sounds in the unanesthetized auditory cortext. *PLoS Biol* **6**, e16, 2008.
[17] Kenyon, R., Radin, C., Ren K. and L. Sadun. (2014). Multipodal structure and phase transitions in large constraigned graphs. Preprint.
[18] Kenyon, R. and M. Yin. (2014). On the asymptotics of constrained exponential random graphs. Preprint.
[19] Lovász, L. (2009). Very large graphs. *Current Develop. Math.* **2008**, 67-128.

[20] Lovász, L. (2012). *Large Networks and Graph Limits.* American Mathematical Society, Providence.
[21] Lovász, L. and B. Szegedy. (2006). Limits of dense graph sequences. *J. Combin. Theory Ser. B* **96**, 933-957.
[22] Lubetzky, E. and Y. Zhao. (2012). On replica symmetry of large deviations in random graphs. To appear in Random Structures and Algorithms.
[23] Newman, M. E. J. (2010). *Networks: An Introduction.* Oxford University Press, Oxford.
[24] Pikhurko, O. and A. Razborov. (2012). Asymptotic structure of graphs with the minimum number of triangles. arXiv:1203.4393.
[25] Radin, C. and M. Yin. (2013). Phase transitions in exponential random graphs. *Annals of Applied Probability.* **23**, 2458-2471.
[26] Radin, C. and L. Sadun. (2013). Phase transitions in a complex network. *J. Phys. A: Math. Theor.* **46**, 305002.
[27] Radin, C. and L. Sadun. (2013). Singularities in the entropy of asymptotically large simple graphs. Preprint.
[28] Radin, C., Ren, K. and L. Sadun. (2014). The asymptotics of large constrained graphs. *J. Phys. A: Math. Theor.* **47**, 175001.
[29] Razborov, A. A. (2008) On the minimal density of triangles in graphs. *Combin. Probab. Comput.* **17**, 603-618.
[30] Rinaldo, A., Fienberg, S. and Y. Zhou. (2009). On the geometry of discrete exponential families with application to exponential random graph models. *Electron. J. Stat.* **3**, 446-484.
[31] Snijders, T. A. B., Pattison, P., Robins, G. L. and M. Handcock. (2006). New specifications for exponential random graph models. *Sociological Methodology.* **36**, 99-153.
[32] Wasserman, S. and K. Faust. (2010). *Social Network Analysis: Methods and Applications. Structural Analysis in the Social Sciences*, 2nd ed. Cambridge Univ. Press, Cambridge.
[33] Ye, J. and J. Liu. (2012). Sparse methods for biomedical data. *SIGKDD Explor.* **14**, 4-15.
[34] Yin, M. (2013). Critical phenomena in exponential random graphs. *J. Stat. Phys.* **153**, 1008-1021.
[35] Yin, M., Rinaldo, A. and S. Fadnavis. (2013). Asymptotic quantization of exponential random graphs. Preprint.
[36] Zhu, L. (2014). Asymptotic structure of constrained exponential random graph models. Preprint.

Department of Mathematics
University of Denver
2280 S. Vine St.
Denver, CO-80208
United States of America
*E-mail address*: mei.yin@du.edu

School of Mathematics
University of Minnesota-Twin Cities
206 Church Street S.E.
Minneapolis, MN-55455
United States of America
*E-mail address*: zhul@umn.edu