

ON THE PHASE TRANSITION CURVE IN A DIRECTED EXPONENTIAL RANDOM GRAPH MODEL

DAVID ARISTOFF,* *Colorado State University*

LINGJIONG ZHU,** *Florida State University*

Abstract

We consider a family of directed exponential random graph models parametrized by edges and outward stars. Much of the important statistical content of such models is given by the normalization constant of the models, and, in particular, an appropriately scaled limit of the normalization, which is called the *free energy*. We derive precise asymptotics for the normalization constant for *finite* graphs. We use this to derive a formula for the free energy. The limit is analytic everywhere except along a curve corresponding to a first-order phase transition. We examine unusual behavior of the model along the phase transition curve.

Keywords: Dense random graph; exponential random graph; graph limit; entropy; phase transition

2010 Mathematics Subject Classification: Primary 05C80; 82B26

1. Introduction

Probabilistic ensembles with one or more adjustable parameters are often used to model complex networks, including social networks, biological networks, and the internet; see [9], [10], [16], and [19]. One of the standard complex network models are the *exponential random graph models* (ERGMs), originally studied by Besag [4]. We refer the reader to [24], [26], and [28] for the history and a review of recent developments.

Much of the statistical content of such models can be described by the probability normalization. An appropriately scaled limit of the normalization, which is called the *free energy*, is useful for understanding properties of large graphs sampled from ERGMs. The term free energy comes from an analogous quantity in statistical physics; see Section 1.3. In that setting, the free energy is used to draw phase diagrams corresponding (for example) to the familiar solid, liquid, and vapor phases of matter [11]. In the random graph setting, it has recently been used to understand asymptotic behavior of ERGMs, including singular behavior, e.g. ill-posedness of parameter fitting problems [6].

The study of the free energy in ERGMs dates to Park and Newman [21], [22], who used mean-field and other nonrigorous approximations. For early research, see also the references in [13]. The first rigorous study appeared in Chatterjee and Diaconis [6], who used a large deviation approach [7]. Radin and Yin [23] used the work of Chatterjee and Diaconis to

Received 22 December 2016; revision received 26 October 2017.

* Postal address: Department of Mathematics, Colorado State University, 1874 Campus Delivery, Fort Collins, CO-80523, USA. Email address: aristoff@math.colostate.edu

** Postal address: Department of Mathematics, Florida State University, 1017 Academic Way, Tallahassee, FL-32306, USA. Email address: zhu@math.fsu.edu

formalize the notion of phases for ERGMs, explicitly computing phase diagrams for a family of two-parameter models. A similar three-parameter family was studied by Yin [30].

We consider a family of directed exponential random graphs parametrized by edges and outward directed p -stars. Versions of this model in which edges and outward directed p -stars are held fixed, instead of controlled by a parameter, were studied in [2]. Such models are standard and important in the literature of social networks; see, e.g. [14] and [17] and the references therein. Many complex networks have directed structures. Examples include email networks and social networks. In email networks, just because user Bob's email address appears in user Alice's address book does not necessarily mean that the reverse is also true, although it often is; see, e.g. [20]. In social networks, when two users Alice and Bob interact as peers, one expects that messages will be exchanged between them in both directions. However, if user Alice sends messages to user Bob, who is a celebrity or news source, it is likely that user Bob will not send messages in return; see, e.g. [8]. In this paper we consider the statistics of edges and outward directed p -stars for two main reasons. First, this is the model that is more analytically tractable than the more general directed ERGMs, and the main results we obtain in this paper rely on the special properties of this model. Second, our directed ERGM with edge and outward directed p -stars has a nice microeconomic interpretation, which can be viewed as the long-run equilibrium of a game in which the players are maximizing their own utilities. We will describe this in more detail in Section 1.2.

For directed graphs, the results of Chatterjee and Diaconis [6] and Radin and Yin [23] do not directly apply, as the large deviation techniques they used have been developed only for undirected graphs. Instead of adapting these techniques to the directed case, we use more direct methods which lead to more precise asymptotics. In particular, we are able to completely characterize the phase behavior of our models. The phase diagram we obtain, with a first-order phase transition and a critical point, is nearly identical to that found in the undirected case [23]. But because our asymptotics are more precise, we can go beyond results of the type in [23] by studying the phase transition curve itself.

The limitation of our model is that it concerns the edge and outward directed p -stars, rather than the more general statistics in the more general directed ERGMs. Our method and results cannot be applied to the undirected ERGM either. It would certainly be very interesting mathematically to pursue the similar results as obtained in this paper for the more general directed ERGMs or undirected ERGMs, although it might be a huge challenge and an open problem in terms of mathematics. We study our particular model because of its analytical tractability and exact solvability. At least for this special model, we can study and understand the precise asymptotics rigorously and completely.

1.1. The model

We consider the following ERGM. Fix $p \geq 2$. For a directed graph X on n nodes and real parameters β_1, β_2 , define

$$\mathbb{P}_{n, \beta_1, \beta_2}(X) \sim \exp\left(\beta_1 E(X) + \frac{p! \beta_2}{n^{p-1}} S(X)\right), \quad (1)$$

where $E(X) := \sum_{i,j} X_{ij}$ and $S(X) := (1/p!) \sum_{i, j_1, \dots, j_p} X_{ij_1} X_{ij_2} \cdots X_{ij_p}$ are, respectively, the number of directed edges and outward directed p -stars in X , and ' \sim ' denotes equality up to a normalization constant. We consider only directed graphs without duplicate edges, though loops will be allowed.

The following reformulation of (1) will be useful. A simple directed graph X on n nodes is given by its adjacency 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 node i to node j ; otherwise, $X_{ij} = 0$. Note that we allow $X_{ii} = 0$ or 1, corresponding to the absence or presence, respectively, of a loop at node i . Define

$$e(X) = n^{-2} \sum_{1 \leq i, j \leq n} X_{ij}, \quad s(X) = n^{-p-1} \sum_{1 \leq i, j_1, j_2, \dots, j_p \leq n} X_{ij_1} X_{ij_2} \cdots X_{ij_p}.$$

Observe that

$$e(X) = n^{-2} E(X), \quad s(X) = p! n^{-p-1} S(X).$$

(Note that we have allowed loops to contribute to p -stars; this is a minor point because the number of loops is of lower order than the number of directed edges.) We think of $e(X)$ and $s(X)$ as homomorphism densities. That is, $e(X)$ is the probability that a random function from a directed edge into X is a homomorphism, i.e. an edge-preserving map between the vertex sets. Similarly, $s(X)$ is the probability that a random function from an outward directed p -star into X is a homomorphism. See [6] for more details. With this notation, we rewrite (1) as

$$\mathbb{P}_{n, \beta_1, \beta_2}(X) = Z_n(\beta_1, \beta_2)^{-1} \exp[n^2(\beta_1 e(X) + \beta_2 s(X))], \tag{2}$$

with $Z_n(\beta_1, \beta_2)$ the normalization constant. We will study

$$\psi_n(\beta_1, \beta_2) = n^{-2} \log Z_n(\beta_1, \beta_2)$$

as well as its limit

$$\psi(\beta_1, \beta_2) := \lim_{n \rightarrow \infty} \psi_n(\beta_1, \beta_2).$$

We refer to $\psi(\beta_1, \beta_2)$ as the *free energy*. It is important for understanding the structure and statistical properties of the model. In particular, first-order partial derivatives of ψ with respect to β_1 and β_2 correspond to the limiting edge and star densities in the model. Similarly, second-order partial derivatives correspond to the limiting edge and star variances. Consequently, singularities of $\psi(\beta_1, \beta_2)$ correspond to singular behavior in the model as the parameters vary. For instance, a singularity in a first-order partial derivative of ψ corresponds to a jump discontinuity of the limiting edge/star densities as β_1, β_2 vary across the singularity.

We find below that the first derivative of ψ is singular along a certain curve in the (β_1, β_2) plane. This curve has an endpoint, at which the second derivative of ψ is singular. A similar singularity has been found in the undirected version of the model; see [23]. Our results are novel because, in contrast with the undirected case, we are able to obtain sharp asymptotics for ψ_n and its partial derivatives at *finite* n . This allows us to make precise statements about the nature of the singularity in the model. In particular, we can describe the scaling of edge and star variances along the singularity. We explore this in detail in Sections 1.3 and 1.4 by using an analogy with the grand canonical ensemble in statistical physics, an exponential family similar to (2). See also [23] for a similar discussion.

1.2. Network formation

In this section we consider a microeconomic model of network formation that will be seen to converge in the long-run to the directed ERGM model that we propose. Similar network formation models in economics literature which convergence to the equilibrium of ERGMs can be found in, e.g. [3], [5], and [17].

Consider n players and if there is a link from player i to player j , we have $X_{ij} = 1$ and 0 otherwise. For simplicity, we assume that $X_{ii} = 0$ always. The link from player i to player j can be interpreted as an email message, or the invitation to an event in a social network. Note

that an email message may not get replied to, and an invitation may not get accepted. Thus, it fits into the directed network setting. For player i , we define his/her utility function as

$$u_i(X) = \sum_{j=1}^n \beta_1 X_{ij} + \sum_{1 \leq j_1, j_2, \dots, j_p} \frac{\beta_2}{n^{p-1}} X_{ij_1} \cdots X_{ij_p}.$$

In this setting, player i will have an incentive β_1 to form a link to any other player j and will have an incentive β_2/n^{p-1} to form links simultaneously to p players j_1, \dots, j_p . We suppose that β_1 is positive and β_2 is negative. That means there is a positive incentive to invite/message your friends, but there is a negative incentive to invite/message p friends or more simultaneously. That can be explained as the incentive to be sociable but not overly sociable.

Over a long period of time, we assume that at each (discrete) time, a player i updates X_{ij} , $j \neq i$, to maximize his/her utility, and before player i updates the links, he/she receives an idiosyncratic shock to his/her preferences that the econometrician cannot observe. The shocks are assumed to be independent and identically distributed (i.i.d.) logistic shocks among players and across time, which is a standard assumption in economics and statistics; see, e.g. [27]. Under these assumptions, the network formation process evolves according to a Markov chain which is irreducible and aperiodic and, hence, as time goes to ∞ it converges to an equilibrium with a stationary distribution. In the absence of random shocks, the network formation process will converge to a Nash network as time goes to ∞ , where a Nash network is a network in which a player has no profitable deviations from his/her current linking strategy. The random shock models unobservables that could influence the utility of additional links; see, e.g. [17]. As time evolves, the network will converge to an equilibrium (see, e.g. [5] and [17], and the derivations follow the same arguments as in Appendix A of [17]) in which the stationary probability of observing a particular network configuration X is given precisely by

$$Z_n(\beta_1, \beta_2)^{-1} \exp \left[\beta_1 \sum_{i,j} X_{ij} + \frac{\beta_2}{n^{p-1}} \sum_{i,j_1,\dots,j_p} X_{ij_1} \cdots X_{ij_p} \right],$$

which is the directed ERGM model we defined in (2), where the exponent

$$\beta_1 \sum_{i,j} X_{ij} + \frac{\beta_2}{n^{p-1}} \sum_{i,j_1,\dots,j_p} X_{ij_1} \cdots X_{ij_p}$$

is known as the potential function in the economics literature, which is the combined utility of n players.

For econometricians, it is crucial to understand how to estimate the parameters β_1, β_2 from real-world data. Note that the normalizing constant $Z_n(\beta_1, \beta_2)$ and, hence, $\psi_n(\beta_1, \beta_2)$ depend on the parameters β_1, β_2 to be estimated, which brings a challenge to the maximum likelihood estimation method. The Markov chain Monte Carlo (MCMC) method has been proposed for the estimation; see, e.g. [25]. But the MCMC method becomes computationally expensive for large networks of size n . The alternative approach developed in recent years is the variational inference, by directly computing and analyzing the constant $\psi_n(\beta_1, \beta_2)$ as $n \rightarrow \infty$, see, e.g. [18], which will be the focus for the rest of this paper.

1.3. The grand canonical ensemble and phase transitions

To explain how our results fit into the phase diagram framework of [23], we compare our model (2) with the grand canonical ensemble from statistical physics, which describes the

statistical properties of matter in thermal equilibrium [12]. We consider the grand canonical ensemble defined by, for $Y \subset [-n/2, n/2]^d$ with $d = 2$ or 3 ,

$$\mathbb{P}_{n,\beta,\mu}(Y) = Z_n(\beta, \mu)^{-1} \exp(n^d[\beta\mu\mathcal{N}(Y) - \beta\mathcal{E}(Y)]),$$

where μ is the chemical potential, and $\beta = 1/(k_B T)$ with T denoting temperature and k_B Boltzmann's constant. Here $Z_n(\beta, \mu)$ is the normalization constant. Each element of Y represents a particle, with $\mathcal{N}(Y) = |Y|/n^d$ the density of Y , and $\mathcal{E}(Y)$ the energy per volume of Y . A standard fact in statistical physics is that average physical properties of the model can be obtained from

$$\psi_n(\beta, \mu) := n^{-d} \log Z_n(\beta, \mu).$$

In particular, the average and variance of $\mathcal{N}(Y)$ and $\mathcal{E}(Y)$, or, more generally, all of their moments, can be obtained by differentiating $\psi_n(\beta, \mu)$ with respect to β or μ . Usually, n is very large and it is appropriate to consider

$$\psi(\beta, \mu) := \lim_{n \rightarrow \infty} \psi_n(\beta, \mu),$$

which exists under appropriate conditions on \mathcal{E} . One utility of this limit is that

$$\lim_{n \rightarrow \infty} \frac{\partial^{i+j}}{\partial \beta^i \partial \mu^j} \psi_n(\beta, \mu) = \frac{\partial^{i+j}}{\partial \beta^i \partial \mu^j} \lim_{n \rightarrow \infty} \psi_n(\beta, \mu) = \frac{\partial^{i+j}}{\partial \beta^i \partial \mu^j} \psi(\beta, \mu) \quad (3)$$

whenever i, j are such that the derivative on the right-hand side exists [29]. This means that, in the limit $n \rightarrow \infty$, moments of $\mathcal{N}(Y)$ and $\mathcal{E}(Y)$ can be computed directly from $\psi(\beta, \mu)$, provided the appropriate partial derivatives of $\psi(\beta, \mu)$ exist.

The limit $\psi(\beta, \mu)$ is key to understanding phases of matter. In particular, when \mathcal{E} is suitably chosen, $\psi(\beta, \mu)$ is analytic except along two curves with an endpoint. These curves correspond to the physical transitions between solid, liquid, and vapor phases, colloquially known as freezing, melting, boiling and sublimating. The endpoint of these curves is called the critical point [11]. See Figure 1(a). These transitions are *first order*, meaning the first derivative of the free energy has a jump discontinuity across the transition curve. On the phase transition curve there are coexisting phases of high and low density, and, thus, a nonvanishing variance of $\mathcal{N}(Y)$ and $\mathcal{E}(Y)$ in the limit $n \rightarrow \infty$.

Unfortunately, rigorous analysis of $\psi(\beta, \mu)$ is difficult. Though the statements in the previous paragraph are widely believed and supported by numerical experiments, proofs are possible only in very special cases [15]. On the other hand, analysis of the ERGM free energy $\psi(\beta_1, \beta_2)$ is relatively tractable. Indeed, we show that $\psi(\beta_1, \beta_2)$ exhibits behavior very similar to what is conjectured for the grand canonical free energy $\psi(\beta, \mu)$.

1.4. Singularities of the ERGM free energy

We show that $\psi(\beta_1, \beta_2)$ is analytic except along a certain curve, which we call the *phase transition curve*. The curve has an endpoint, which we call the *critical point*. We prove that on the phase transition curve, but away from the critical point, the first-order partial derivatives of $\psi(\beta_1, \beta_2)$ have a jump discontinuity. Moreover, at the critical point, the first-order partial derivatives of $\psi(\beta_1, \beta_2)$ are continuous but the second-order derivatives diverge. Precisely the same behavior is believed to occur on the liquid–vapor transition curve in the grand canonical ensemble. See Figure 1.

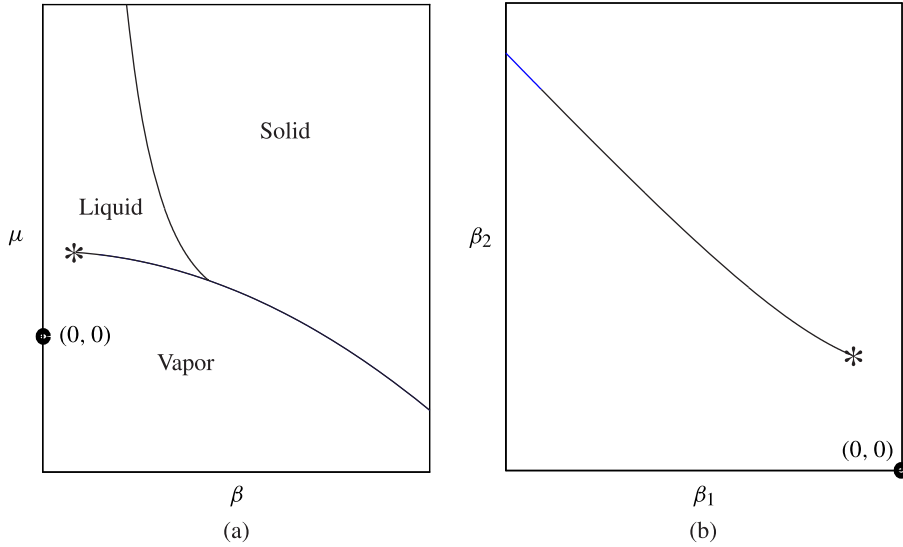


FIGURE 1: Simple phase diagrams in (a) the grand canonical ensemble, and (b) the ERGM model. The critical point is labeled with a ‘*’.

To understand these singularities better, consider the following. Just as in (3),

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\partial}{\partial \beta_i} \psi_n(\beta_1, \beta_2) &= \frac{\partial}{\partial \beta_i} \psi(\beta_1, \beta_2), & i \in \{1, 2\}, \\ \lim_{n \rightarrow \infty} \frac{\partial^2}{\partial \beta_i \partial \beta_j} \psi_n(\beta_1, \beta_2) &= \frac{\partial^2}{\partial \beta_i \partial \beta_j} \psi(\beta_1, \beta_2), & i, j \in \{1, 2\}, \end{aligned} \tag{4}$$

if the derivatives on the right-hand side exist. The commuting of limits follows from very general arguments of Yang and Lee [29]. Though their proof is written in the statistical mechanics setting of Section 1.3, their arguments go through without any difficulties in our case. See also [23] for remarks on this issue in the undirected graph setting.

Next, from simple computations,

$$\begin{aligned} \frac{\partial}{\partial \beta_1} \psi_n(\beta_1, \beta_2) &= \mathbb{E}_n[e(X)], & \frac{\partial}{\partial \beta_2} \psi_n(\beta_1, \beta_2) &= \mathbb{E}_n[s(X)], \\ \frac{\partial^2}{\partial \beta_1^2} \psi_n(\beta_1, \beta_2) &= n^2 \text{var}_n(e(X)), & \frac{\partial^2}{\partial \beta_2^2} \psi_n(\beta_1, \beta_2) &= n^2 \text{var}_n(s(X)), \\ \frac{\partial}{\partial \beta_1 \partial \beta_2} \psi_n(\beta_1, \beta_2) &= \frac{\partial}{\partial \beta_2 \partial \beta_1} \psi_n(\beta_1, \beta_2) &= n^2 \text{cov}_n(e(X), s(X)). \end{aligned} \tag{5}$$

Thus, a jump discontinuity in $\partial \psi(\beta_1, \beta_2) / \partial \beta_1$ (respectively, $\partial \psi(\beta_1, \beta_2) / \partial \beta_2$) along the transition curve implies a jump in the average value of $e(X)$ (respectively, $s(X)$) across the curve in the limit $n \rightarrow \infty$. Similarly, at the critical point, divergence of $\partial^2 \psi(\beta_1, \beta_2) / \partial \beta_1^2$ (respectively, $\partial^2 \psi(\beta_1, \beta_2) / \partial \beta_2^2$) implies that the variance of $e(X)$ (respectively, $s(X)$) decays more slowly than n^{-2} . Away from the transition curve, all partial derivatives of $\psi(\beta_1, \beta_2)$ of all orders exist and are finite, so, in particular, the variance of $e(X)$ and $s(X)$ decays at least as fast as n^{-2} .

More detailed statements would require an analysis of $\psi_n(\beta_1, \beta_2)$ for finite n ; this is much more difficult to study than the limit $\psi(\beta_1, \beta_2)$. See [6] and [23] for an analysis of the free energy in the undirected version of the model.

In the context of social networks (Section 1.2), the quantities computed in (5) reveal the statistics of the network at its long-run equilibrium. For instance, $\mathbb{E}_n[e(X)]$ denotes the average number of outward links from an average player. The existence of phase transitions tells us that as the direct benefit, β_1 and β_2 , that a player receives in his/her utility vary smoothly, many important statistics of the network, and, hence, the network structure at its long-run equilibrium, may vary nonsmoothly.

Our results rely on new and precise asymptotics for $\psi_n(\beta_1, \beta_2)$ and its partial derivatives. We use our asymptotics to calculate $\psi(\beta_1, \beta_2)$ and the scaling of the variance/covariance of $e(X)$ and $s(X)$ on the phase transition curve. The formula for $\psi(\beta_1, \beta_2)$ resembles the analogous free energy in the undirected version of the model. However, the behavior of our model along the phase transition curve is qualitatively different from the undirected case; see the discussion after Theorem 6 below.

To the best of the authors' knowledge, the study of scaling on the transition curve has not been carried out before, in either the directed or undirected versions of the model. (This is because the scaling cannot be computed from $\psi(\beta_1, \beta_2)$ alone, as (4) does not hold along the phase transition curve.) We find that the variances of $e(X)$ and $s(X)$ vanish on the transition curve as $n \rightarrow \infty$, while the edge probability between fixed nodes is a Bernoulli random variable whose parameter is a convex combination of the expected values of $e(X)$ just above and below the curve. Combining these results, we show below that large graphs do not resemble Erdős–Rényi random graphs with a binary distributed parameter; that is, the first-order phase transition does not correspond to phase coexistence in the usual sense. This is unexpected in light of the statistical physics analogy above. See the discussion below Theorem 7.

The remainder of this paper is organized as follows. Our main results are stated in Section 2. The results are obtained by estimates, stated in Section 3, which allow for a precise computation of $\psi_n(\beta_1, \beta_2)$ and derivatives thereof. All proofs are in Section 4.

2. Notation and results

Our main results rely on the following trick. Observe that we can rewrite

$$e(X) = n^{-2} \sum_{i=1}^n \left(\sum_{j=1}^n X_{ij} \right), \quad s(X) = n^{-p-1} \sum_{i=1}^n \left(\sum_{j=1}^n X_{ij} \right)^p,$$

where $\sum_{j=1}^n X_{ij}$, $i = 1, \dots, n$, are independent random variables. Thus, we can calculate $Z_n(\beta_1, \beta_2)$, and, hence, also $\psi_n(\beta_1, \beta_2)$, as 'one-dimensional' objects. See Proposition 1 below for details. Our analysis will depend heavily on the function

$$\ell(x) := \beta_1 x + \beta_2 x^p - x \log x - (1 - x) \log(1 - x).$$

It is easy to see that ℓ is analytic in $(0, 1)$ and continuous on $[0, 1]$. Note that ℓ is essentially identical to the function of the same name studied in [23]: after multiplying β_1 and β_2 by 2, the functions differ only by a constant. This allows us to use results from [23] relating to ℓ .

Our first result is the following formula for the free energy.

Theorem 1. *For any β_1, β_2 as $n \rightarrow \infty$, we have*

$$\psi_n(\beta_1, \beta_2) = \max_{x \in [0,1]} \ell(x) + O(n^{-1} \log n).$$

In particular, letting $n \rightarrow \infty$, we obtain

$$\psi(\beta_1, \beta_2) = \max_{x \in [0,1]} \ell(x).$$

Essentially the same formula holds in the undirected case [23]. Thus, the phase diagram for our model is the same as in the undirected version, after multiplying β_1 and β_2 by 2. The following result from [23] can be used to characterize the curve along which $\psi(\beta_1, \beta_2)$ is singular.

Theorem 2. (Radin and Yin [23].) *There is a certain curve in the (β_1, β_2) -plane with the endpoint*

$$(\beta_1^c, \beta_2^c) = \left(\log(p-1) - \frac{p}{p-1}, \frac{p^{p-1}}{(p-1)^p} \right),$$

such that off the curve and at the endpoint, ℓ has a unique global maximizer $x^* \in (0, 1)$, while on the curve away from the endpoint, ℓ has two global maximizers, x_1^* and x_2^* , with $0 < x_1^* < (p-1)/p < x_2^* < 1$.

The curve in Theorem 2 will be called the *phase transition curve* and written $\beta_2 = q(\beta_1)$. The endpoint will be called the *critical point*. Though our free energy $\psi(\beta_1, \beta_2)$ is essentially the same as its undirected counterpart, the behavior of our model on the phase transition curve is qualitatively different from the undirected version, as we will see below.

It is not possible to write an explicit equation for the phase transition curve in general; see [23] for a graph obtained numerically. However, in [23] it was shown that $q(\beta_1)$ is continuous and decreasing in β_1 , with $\lim_{\beta_1 \rightarrow -\infty} |q(\beta_1) + \beta_1| = 0$. We have the following more precise result.

Theorem 3. (i) *It holds that $q(\beta_1)$ is differentiable for $\beta_1 < \beta_1^c$ with*

$$q'(\beta_1) = -\frac{x_1^* - x_2^*}{(x_1^*)^p - (x_2^*)^p} < 0.$$

In particular,

$$\lim_{\beta_1 \rightarrow \beta_1^c} q'(\beta_1) = -\frac{p^{p-2}}{(p-1)^{p-1}} \quad \text{and} \quad \lim_{\beta_1 \rightarrow -\infty} q'(\beta_1) = -1.$$

(ii) *It holds that $q(\beta_1)$ is convex in β_1 .*

When $p = 2$, along the line $\beta_1 + \beta_2 = 0$ the function ℓ is symmetric around $\frac{1}{2}$. It follows that $x_1^* + x_2^* = 1$ along this line, so Theorem 3 implies that $\ell(\beta_1) = -\beta_1$. See Figure 2(a).

In the following theorems we obtain the scaling of the variance and covariance of $e(X)$ and $s(X)$. See [23] for computations of these quantities off the phase transition curve in the undirected graph case. Here we compute the scaling of the variance and covariance at all (β_1, β_2) , including on the phase transition curve and at the critical point. On the transition curve, we use precise asymptotics for partial derivatives of $\psi_n(\beta_1, \beta_2)$ to obtain the scalings. We emphasize that such asymptotics cannot be obtained directly from large deviations techniques of the type used in [6] and [23].

Theorem 4. *Off the phase transition curve,*

$$\lim_{n \rightarrow \infty} \frac{\partial^2}{\partial \beta_1^2} \psi_n(\beta_1, \beta_2) = \frac{\partial^2}{\partial \beta_1^2} \lim_{n \rightarrow \infty} \psi_n(\beta_1, \beta_2) = \frac{1}{|\ell''(x^*)|}.$$

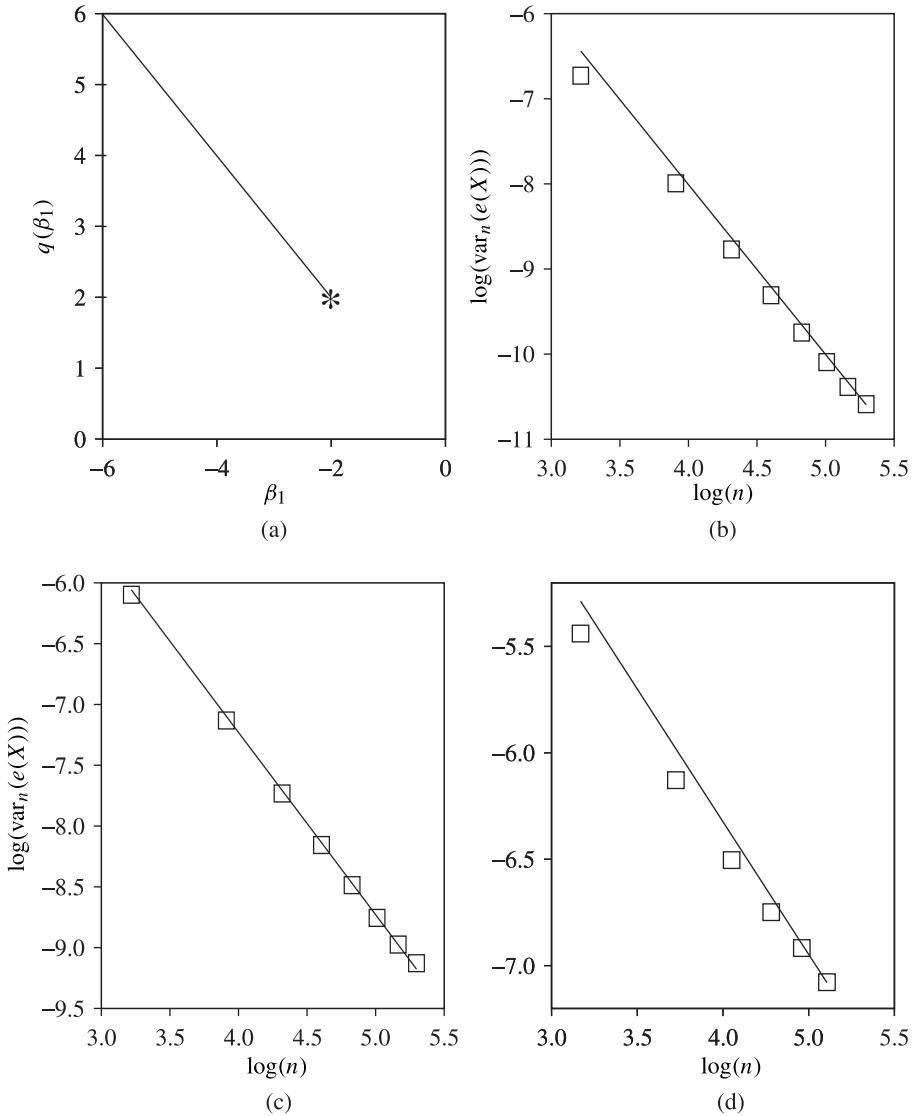


FIGURE 2: (a) The graph of the phase transition curve $\beta_2 = q(\beta_1)$ when $p = 2$, with the critical point labeled by $*$. (b)–(d) Scaling of the variance of $e(X)$, (b) off the phase transition curve, (c) on the phase transition curve away from the critical point, and (d) at the critical point. For (b)–(d) we use $p = 2$ and (β_1, β_2) values of $(-\frac{3}{2}, \frac{3}{2})$, $(-2, 2)$ and $(-\frac{5}{2}, \frac{5}{2})$, respectively. The straight lines are obtained from the scaling in Theorem 4, and the squares are obtained by Monte Carlo simulation.

On the phase transition curve except at the critical point,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \frac{\partial^2}{\partial \beta_1^2} \psi_n(\beta_1, \beta_2) = (x_1^* - x_2^*)^2 \sqrt{x_1^*(1 - x_1^*)|\ell''(x_1^*)|} \sqrt{x_2^*(1 - x_2^*)|\ell''(x_2^*)|} \times \left(\sqrt{x_1^*(1 - x_1^*)|\ell''(x_1^*)|} + \sqrt{x_2^*(1 - x_2^*)|\ell''(x_2^*)|} \right)^{-2}.$$

At the critical point,

$$\lim_{n \rightarrow \infty} \frac{1}{n^{1/2}} \frac{\partial^2}{\partial \beta_1^2} \psi_n(\beta_1, \beta_2) = \frac{\Gamma(3/4)}{\Gamma(1/4)} \frac{2\sqrt{6}(p-1)}{p^{5/2}}.$$

Theorem 5. *Off the phase transition curve,*

$$\lim_{n \rightarrow \infty} \frac{\partial^2}{\partial \beta_2^2} \psi_n(\beta_1, \beta_2) = \frac{\partial^2}{\partial \beta_2^2} \lim_{n \rightarrow \infty} \psi_n(\beta_1, \beta_2) = \frac{p^2(x^*)^{2p-2}}{|\ell''(x^*)|}.$$

On the transition curve except at the critical point,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \frac{\partial^2}{\partial \beta_2^2} \psi_n(\beta_1, \beta_2) &= ((x_1^*)^p - (x_2^*)^p)^2 \sqrt{x_1^*(1-x_1^*)|\ell''(x_1^*)|} \sqrt{x_2^*(1-x_2^*)|\ell''(x_2^*)|} \\ &\times \left(\sqrt{x_1^*(1-x_1^*)|\ell''(x_1^*)|} + \sqrt{x_2^*(1-x_2^*)|\ell''(x_2^*)|} \right)^{-2}. \end{aligned}$$

At the critical point,

$$\lim_{n \rightarrow \infty} \frac{1}{n^{1/2}} \frac{\partial^2}{\partial \beta_2^2} \psi_n(\beta_1, \beta_2) = \frac{2\sqrt{6}\Gamma(3/4)}{\Gamma(1/4)} \frac{(p-1)^{2p-1}}{p^{2p-3/2}}.$$

Theorem 6. *Off the phase transition curve,*

$$\lim_{n \rightarrow \infty} \frac{\partial^2}{\partial \beta_1 \partial \beta_2} \psi_n(\beta_1, \beta_2) = \frac{\partial^2}{\partial \beta_1 \partial \beta_2} \lim_{n \rightarrow \infty} \psi_n(\beta_1, \beta_2) = \frac{p(x^*)^{p-1}}{|\ell''(x^*)|}.$$

On the transition curve except at the critical point,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \frac{\partial^2}{\partial \beta_1 \partial \beta_2} \psi_n(\beta_1, \beta_2) &= ((x_1^*)^p - (x_2^*)^p)(x_1^* - x_2^*) \sqrt{x_1^*(1-x_1^*)|\ell''(x_1^*)|} \sqrt{x_2^*(1-x_2^*)|\ell''(x_2^*)|} \\ &\times \left(\sqrt{x_1^*(1-x_1^*)|\ell''(x_1^*)|} + \sqrt{x_2^*(1-x_2^*)|\ell''(x_2^*)|} \right)^{-2}. \end{aligned}$$

At the critical point,

$$\lim_{n \rightarrow \infty} \frac{1}{n^{1/2}} \frac{\partial^2}{\partial \beta_1 \partial \beta_2} \psi_n(\beta_1, \beta_2) = \frac{2\sqrt{6}\Gamma(3/4)}{\Gamma(1/4)} \frac{(p-1)^p}{p^{p+1/2}}.$$

See Figure 2 for a comparison of these results with Monte Carlo simulation. By Theorems 4 and 5, and (5), the variances of $e(X)$ and $s(X)$ are of order n^{-2} off the transition curve, of order n^{-1} on the phase transition curve away from the critical point, and of order $n^{-3/2}$ at the critical point. In particular, the variances of $e(X)$ and $s(X)$ vanish on the transition curve as $n \rightarrow \infty$.

Theorem 7. *Off the phase transition curve and at the critical point,*

$$\lim_{n \rightarrow \infty} \mathbb{P}_n(X_{12} = 1) = x^*.$$

On the phase transition curve except at the critical point,

$$\lim_{n \rightarrow \infty} \mathbb{P}_n(X_{12} = 1) = \alpha x_1^* + (1-\alpha)x_2^*,$$

where

$$\alpha := \frac{\sqrt{x_2^*(1-x_2^*)|\ell''(x_2^*)|}}{\sqrt{x_1^*(1-x_1^*)|\ell''(x_1^*)|} + \sqrt{x_2^*(1-x_2^*)|\ell''(x_2^*)|}}. \tag{6}$$

To put our results in perspective, we compare with the undirected case. Since the behavior of the directed and undirected models off the phase transition curve is similar, our discussion focuses on the models on the phase transition curve.

In the undirected case (see Theorem 3.4 of [23]), on the phase transition curve away from the critical point, for large n a typical graph looks like a sample from $G(n, p^*)$, where p^* is a distribution on the two global maximizers $x_1^* < x_2^*$ of ℓ . This distribution was not specified in [23]. However, it is easy to check that the proof of Theorem 7 goes through for undirected graphs. In particular, since $\alpha \in (0, 1)$ and $x_1^* < x_2^*$, the distribution of p^* is nontrivial (i.e. not deterministic), which indicates phase coexistence. Since, for large n , a graph looks like a sample from either $G(n, x_1^*)$ or $G(n, x_2^*)$, both with positive probability, the variances of $e(X)$ and $s(X)$ do not vanish. This is expected along first-order phase transitions, as discussed in Section 1.4.

The situation in our directed graph model is qualitatively different. Let $\vec{G}(n, p)$ be the directed Erdős–Rényi graph on n nodes, in which there is a directed edge between each ordered pair of nodes with probability p . Along the phase transition curve, since the variances of $e(X)$ and $s(X)$ vanish as $n \rightarrow \infty$, a typical large graph in our model does *not* behave like $\vec{G}(n, p^*)$ with p^* sampled from a nontrivial distribution on x_1^* and x_2^* . Thus, there is no phase coexistence in the sense described above. We do not have a more precise result about graph structure along the transition curve, but we suspect the following is true. For large n , on the phase transition curve away from the critical point, a typical graph is ‘bipodal’: there is a node set of size approximately equal to αn in which each node has an outward edge to any other node with probability approximately equal to x_1^* , and another node set of size approximately equal to $(1 - \alpha)n$ in which each node has an outward edge to any other node with probability approximately equal to x_2^* . Indeed, a similar result was obtained for a closely related model [2]. See also [1].

Finally, we obtain the asymptotics of the joint distribution of two directed edges, e.g. X_{12} and X_{34} , or X_{12} and X_{13} .

Theorem 8. (i) *Off the phase transition curve and at the critical point,*

$$\lim_{n \rightarrow \infty} \mathbb{P}_n(X_{12} = 1, X_{34} = 1) = (x^*)^2.$$

On the phase transition curve except at the critical point,

$$\lim_{n \rightarrow \infty} \mathbb{P}_n(X_{12} = 1, X_{34} = 1) = (\alpha x_1^* + (1 - \alpha)x_2^*)^2,$$

where α is defined in (6).

(ii) *Off the phase transition curve and at the critical point,*

$$\lim_{n \rightarrow \infty} \mathbb{P}_n(X_{12} = 1, X_{13} = 1) = (x^*)^2.$$

On the phase transition curve except at the critical point,

$$\lim_{n \rightarrow \infty} \mathbb{P}_n(X_{12} = 1, X_{13} = 1) = \alpha(x_1^*)^2 + (1 - \alpha)(x_2^*)^2,$$

where α is defined in (6).

Remark 1. (i) The result in Theorem 8(i) holds for any $\lim_{n \rightarrow \infty} \mathbb{P}_n(X_{1i} = 1, X_{jk} = 1)$ as long as $j \neq 1$ and $i \neq 1, k \neq j$.

(ii) The formulas for $\lim_{n \rightarrow \infty} \mathbb{P}_n(X_{12} = 1, X_{34} = 0)$, $\lim_{n \rightarrow \infty} \mathbb{P}_n(X_{12} = 1, X_{13} = 0)$, and so on follow directly from Theorem 7 and Theorem 8.

(iii) From the results in Theorem 8, we see that two distinct edges are asymptotically independent off the phase transition curve and at the critical point, and if two directed edges do not share the same root, then they are asymptotically independent even at the critical point.

3. Key estimates

First we have the following formula for the normalization $Z_n(\beta_1, \beta_2)$.

Proposition 1. Let W be a binomial random variable with parameters n and $\frac{1}{2}$:

$$\mathbb{P}(W = i) = 2^{-n} \binom{n}{i}.$$

Then

$$Z_n(\beta_1, \beta_2) = 2^{n^2} \left(\mathbb{E} \left[\exp \left(\beta_1 W + \frac{\beta_2}{n^{p-1}} W^p \right) \right] \right)^n.$$

Next we approximate the expectation in Proposition 1 in terms of an integral.

Proposition 2. Let W be a binomial random variable with parameters n and $\frac{1}{2}$. Then, for any $r < 1$,

$$\begin{aligned} & \mathbb{E} \left[W^k \exp \left(\beta_1 W + \frac{\beta_2}{n^{p-1}} W^p \right) \right] \\ &= \begin{cases} (1 + O(n^{1/2-r})) n^k 2^{-n} \sqrt{\frac{n}{2\pi}} \int_0^1 \sqrt{\frac{x^{2k}}{x(1-x)}} e^{n\ell(x)} dx, & (\beta_1, \beta_2) \neq (\beta_1^c, \beta_2^c), \\ (1 + O(n^{1/4-r})) n^k 2^{-n} \sqrt{\frac{n}{2\pi}} \int_0^1 \sqrt{\frac{x^{2k}}{x(1-x)}} e^{n\ell(x)} dx, & (\beta_1, \beta_2) = (\beta_1^c, \beta_2^c). \end{cases} \end{aligned}$$

Finally, we provide a technical lemma for computing the integral in Proposition 2.

Proposition 3. Let f be an analytic function in $(0, 1)$ with Taylor expansion at $c \in (0, 1)$ given by

$$f(x) = d_0(c) + d_1(c)(x - c) + d_2(c)(x - c)^2 + \dots + d_j(c) := \frac{f^{(j)}(c)}{j!}.$$

For $c \in (0, 1)$, define

$$b_k(c) = \frac{\ell^k(c)}{k!}, \quad \alpha_k(c) = \Gamma\left(\frac{k}{2}\right) |b_2(c)|^{-k/2}, \quad \gamma_k(c) = \frac{1}{2} \Gamma\left(\frac{k}{4}\right) |b_4(c)|^{-k/4}.$$

Assume that $f(x) e^{n\ell(x)} \in L^1[0, 1]$ for each n . Then as $n \rightarrow \infty$, we have the following.

(i) Off the phase transition curve,

$$\int_0^1 f(x) e^{n\ell(x)} dx = e^{n\ell(c)} [n^{-1/2} d_0 \alpha_1 + n^{-3/2} \Lambda + O(n^{-5/2})],$$

where

$$\Lambda := d_2\alpha_3 + d_1b_3\alpha_5 + d_0b_4\alpha_5 + \frac{1}{2}d_0b_3^2\alpha_7$$

with $c = x^*$ the unique maximizer of ℓ , and $d_j = d_j(c)$, $b_j = b_j(c)$, $\alpha_j = \alpha_j(c)$.

(ii) On the phase transition curve except at the critical point,

$$\int_0^1 f(x) e^{n\ell(x)} dx = e^{n\ell(c)} [n^{-1/2}(d_0(c_1)\alpha_1(c_1) + d_0(c_2)\alpha_1(c_2)) + O(n^{-3/2})],$$

where c_1 and c_2 are the maximizers of ℓ .

(iii) At the critical point,

$$\int_0^1 f(x) e^{n\ell(x)} dx = e^{n\ell(c)} [n^{-1/4}d_0\gamma_1 + n^{-3/4}\Theta + O(n^{-5/4})],$$

where

$$\Theta := d_2\gamma_3 + d_1b_5\gamma_7 + d_0b_6\gamma_7 + \frac{1}{2}d_0b_5^2\gamma_{11}$$

with $c = x^*$ the unique maximizer of ℓ , and $d_j = d_j(c)$, $b_j = b_j(c)$, $\gamma_j = \gamma_j(c)$.

Note that this strategy allows for a relatively precise computation of $Z_n(\beta_1, \beta_2)$. Unfortunately, arbitrary precision cannot be achieved, due to the error inherent in the sum to integral approximation of Proposition 2.

4. Proofs

Before turning to the proofs of the theorems of Section 2, we will prove the estimates from Section 3. The following result will be needed in almost all of our proofs.

Proposition 4. *Off the phase transition curve,*

$$\ell'(x^*) = 0, \quad \ell''(x^*) < 0.$$

On the phase transition curve except at the critical point,

$$\ell'(x_1^*) = \ell'(x_2^*) = 0, \quad \ell''(x_1^*) < 0, \quad \ell''(x_2^*) < 0.$$

At the critical point,

$$\ell'(x^*) = \ell''(x^*) = \ell'''(x^*) = 0, \quad \ell^{(4)}(x^*) = \frac{-p^5}{(p-1)^2} < 0.$$

Proof. It is straightforward to compute that

$$\begin{aligned} \ell'(x) &= \beta_1 + p\beta_2x^{p-1} - \log\left(\frac{x}{1-x}\right), \\ \ell''(x) &= p(p-1)\beta_2x^{p-2} - \frac{1}{x} - \frac{1}{1-x}, \\ \ell'''(x) &= p(p-1)(p-2)\beta_2x^{p-3} + \frac{1}{x^2} - \frac{1}{(1-x)^2}, \\ \ell^{(4)}(x) &= p(p-1)(p-2)(p-3)\beta_2x^{p-4} - \frac{2}{x^3} - \frac{2}{(1-x)^3}. \end{aligned}$$

Since $\lim_{x \rightarrow 0^+} \ell'(x) = +\infty$ and $\lim_{x \rightarrow 1^-} \ell'(x) = -\infty$, the maximum is achieved at a local maximum, and we have $\ell'(x^*) = 0$.

We first show that $\ell''(x^*) < 0$ off the critical point (where x^* denotes either x_1^* or x_2^* if we are on the phase transition curve). Following the proof of Proposition 3.2 of [23], we first analyze the properties of $\ell''(x)$. We can write $\ell''(x)$ as

$$\ell''(x) = x^{p-2} p(p-1) \left[\beta_2 - \frac{1}{p(p-1)x^{p-1}(1-x)} \right].$$

Consider the function

$$m(x) := \frac{1}{p(p-1)x^{p-1}(1-x)}.$$

It is easy to observe that $m(x) \geq p^{p-1}/(p-1)^p$ and the equality holds if and only if $x = (p-1)/p$.

- (i) If $\beta_2 < p^{p-1}/(p-1)^p$, $\ell''(x) < 0$ on $[0, 1]$ and, in particular, $\ell''(x^*) < 0$.
- (ii) If $\beta_2 > p^{p-1}/(p-1)^p$, there exist $0 < x_1 < (p-1)/p < x_2 < 1$ so that $\ell''(x) < 0$ on $0 < x < x_1$, $\ell''(x) > 0$ on $x_1 < x < x_2$, and $\ell''(x) < 0$ on $x_2 < x < 1$. Moreover, $\ell''(x_1) = \ell''(x_2) = 0$. If $\ell'(x_1) \geq 0$, $\ell(x)$ has a unique local and, hence, global maximizer $x^* > x_2$; if $\ell'(x_2) \leq 0$, $\ell(x)$ has a unique local and, hence, global maximizer $x^* < x_1$. Finally, if $\ell'(x_1) < 0 < \ell'(x_2)$ then $\ell(x)$ has two local maximizers x_1^* and x_2^* so that $x_1^* < x_1 < (p-1)/p < x_2 < x_2^*$. Since ℓ'' vanishes only at x_1 and x_2 , we have proved that $\ell''(x^*) < 0$.
- (iii) If $\beta_2 = p^{p-1}/(p-1)^p$, $\ell''(x) \leq 0$ on $[0, 1]$ and $\ell''(x) = 0$ if and only if $x = (p-1)/p$ by the properties of $m(x)$. Therefore, $\ell''(x^*) = 0$ if and only if $x^* = (p-1)/p$. Since $\ell'(x^*) = 0$, $x^* = (p-1)/p$ if and only if

$$\beta_1 = -p \frac{p^{p-1}}{(p-1)^p} \left(\frac{p-1}{p} \right)^{p-1} + \log \left(\frac{(p-1)/p}{1 - (p-1)/p} \right) = \beta_1^c.$$

Hence, $\ell''(x^*) < 0$ off the critical point and $\ell''(x^*) = 0$ at the critical point.

Furthermore, at the critical point $(\beta_1, \beta_2) = (\beta_1^c, \beta_2^c)$, we can compute

$$\ell'''(x^*) = p(p-1)(p-2) \frac{p^{p-1}}{(p-1)^p} \frac{(p-1)^{p-3}}{p^{p-3}} + \frac{p^2}{(p-1)^2} - p^2 = 0.$$

Moreover,

$$\ell^{(4)}(x^*) = p(p-1)(p-2)(p-3) \frac{p^{p-1}}{(p-1)^p} \frac{(p-1)^{p-4}}{p^{p-4}} - \frac{2p^3}{(p-1)^3} - 2p^3 = \frac{-p^5}{(p-1)^2} < 0.$$

This completes the proof. □

The next three proofs are for the results in Section 3.

Proof of Proposition 1. Let $Y = (Y_{ij})_{1 \leq i, j \leq n}$ be an $n \times n$ matrix of i.i.d. Bernoulli random variables:

$$\mathbb{P}(Y_{ij} = 0) = \frac{1}{2} = \mathbb{P}(Y_{ij} = 1).$$

For $i = 1, \dots, n$, define

$$W_i = \sum_{j=1}^n Y_{ij}.$$

Then

$$\begin{aligned}
 Z_n(\beta_1, \beta_2) &= 2^{n^2} \mathbb{E}[\exp(n^2(\beta_1 e(Y) + \beta_2 s(Y)))] \\
 &= 2^{n^2} \mathbb{E}\left[\exp\left(\sum_{i=1}^n \beta_1 W_i + \frac{\beta_2}{n^{p-1}} W_i^p\right)\right] \\
 &= 2^{n^2} \mathbb{E}\left[\prod_{i=1}^n \exp\left(\beta_1 W_i + \frac{\beta_2}{n^{p-1}} W_i^p\right)\right] \\
 &= 2^{n^2} \prod_{i=1}^n \mathbb{E}\left[\exp\left(\beta_1 W_i + \frac{\beta_2}{n^{p-1}} W_i^p\right)\right] \\
 &= 2^{n^2} \left(\mathbb{E}\left[\exp\left(\beta_1 W + \frac{\beta_2}{n^{p-1}} W^p\right)\right]\right)^n. \quad \square
 \end{aligned}$$

Proof of Proposition 2. We will prove only the $k = 0$ case, as the other cases are easy extensions. Observe that

$$\mathbb{E}\left[\exp\left(\beta_1 W + \frac{\beta_2}{n^{p-1}} W^p\right)\right] = 2^{-n} \sum_{i=1}^n \binom{n}{i} \exp\left(\beta_1 i + \frac{\beta_2}{n^{p-1}} i^p\right).$$

Using the fact that, for all $n \geq 1$,

$$n \log n - n + \frac{1}{2} \log n \leq \log n! \leq n \log n - n + \frac{1}{2} \log n + 1,$$

we obtain

$$\binom{n}{i} \leq \exp\left(n\left[-\frac{i}{n} \log \frac{i}{n} - \left(1 - \frac{i}{n}\right) \log\left(1 - \frac{i}{n}\right) + \frac{1}{2n} \log \frac{n}{i(n-i)} + \frac{1}{n}\right]\right). \quad (7)$$

Define

$$A_n = \left\{i \in \{1, \dots, n\} : \frac{i}{n} \in (\varepsilon, 1 - \varepsilon)\right\},$$

where $\varepsilon > 0$ will be specified momentarily. From (7), for any $\varepsilon \in (0, 1)$, we have

$$\begin{aligned}
 \max_{i \in \{1, \dots, n\} \setminus A_n} \binom{n}{i} \exp\left(\beta_1 i + \frac{\beta_2}{n^{p-1}} i^p\right) &\leq e \cdot \left(1 - \frac{1}{n}\right)^{-1/2} \sup_{x \in [0, 1] \setminus (\varepsilon, 1 - \varepsilon)} e^{n\ell(x)} \\
 &\leq 3 \sup_{x \in [0, 1] \setminus (\varepsilon, 1 - \varepsilon)} e^{n\ell(x)}. \quad (8)
 \end{aligned}$$

Since $\ell'(x) \rightarrow \infty$ as $x \rightarrow 0$ and $\ell'(x) \rightarrow -\infty$ as $x \rightarrow 1$, the optimizer x_* is in $(0, 1)$ and we may choose $\varepsilon > 0$ such that, for some $\delta > 0$,

$$\sup_{x \in [0, 1] \setminus (\varepsilon, 1 - \varepsilon)} \ell(x) < \ell(x_*) - \delta.$$

Thus, $\sup_{x \in [0, 1] \setminus (\varepsilon, 1 - \varepsilon)} e^{n\ell(x)} \leq e^{n(\ell(x_*) - \delta)}$, and using this with (8) yields

$$\sum_{i \in \{1, \dots, n\} \setminus A_n} \binom{n}{i} \exp\left(\beta_1 i + \frac{\beta_2}{n^{p-1}} i^p\right) = O(e^{n(\ell(x_*) - \delta)}).$$

For $i \in A_n$, Stirling’s formula allows us to write

$$\binom{n}{i} = (1 + O(n^{-1})) \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{i(n-i)}} \exp\left(n\left[-\frac{i}{n} \log \frac{i}{n} - \left(1 - \frac{i}{n}\right) \log\left(1 - \frac{i}{n}\right)\right]\right).$$

The last two displays yield

$$\begin{aligned} & \mathbb{E}\left[\exp\left(\beta_1 W + \frac{\beta_2}{n^{p-1}} W^p\right)\right] \\ &= 2^{-n} \left(\sum_{i=1}^n \binom{n}{i} \exp\left(\beta_1 i + \frac{\beta_2}{n^{p-1}} i^p\right)\right) \\ &= 2^{-n} \left(O(e^{n(\ell(x^*)-\delta)}) + \sum_{i \in A_n} \binom{n}{i} \exp\left(\beta_1 i + \frac{\beta_2}{n^{p-1}} i^p\right)\right) \\ &= 2^{-n} \left(O(e^{n(\ell(x^*)-\delta)}) + (1 + O(n^{-1})) \frac{1}{\sqrt{2\pi n}} \sum_{i \in A_n} \sqrt{\frac{1}{(i/n)(1-i/n)}} e^{n\ell(i/n)}\right). \end{aligned} \tag{9}$$

We will approximate the sum in (9) by an integral. Consider first the case off the transition curve. Thus, there is a unique maximizer x^* of ℓ , and $\ell'(x^*) = 0, \ell''(x^*) < 0$. Let $q \in (\frac{1}{3}, \frac{1}{2})$ and define

$$B_n = \left\{i \in \{1, \dots, n\} : \frac{i}{n} \in (x^* - n^{-q}, x^* + n^{-q})\right\}.$$

For any $j \in A_n$, note that

$$\begin{aligned} & \left|\frac{1}{n} \sqrt{\frac{1}{(j/n)(1-j/n)}} e^{n\ell(j/n)} - \int_{j/n}^{j/n+1/n} \sqrt{\frac{1}{x(1-x)}} e^{n\ell(x)} dx\right| \\ & \leq \frac{1}{n} \max_{x,y \in [j/n, j/n+1/n]} \left|\sqrt{\frac{1}{x(1-x)}} e^{n\ell(x)} - \sqrt{\frac{1}{y(1-y)}} e^{n\ell(y)}\right| \\ & \leq \frac{1}{n} \max_{x \in [j/n, j/n+1/n]} \sqrt{\frac{1}{x(1-x)}} \max_{x,y \in [j/n, j/n+1/n]} |e^{n\ell(x)} - e^{n\ell(y)}| \\ & \quad + \frac{1}{n} e^{n\ell(x^*)} \max_{x,y \in [j/n, j/n+1/n]} \left|\sqrt{\frac{1}{x(1-x)}} - \sqrt{\frac{1}{y(1-y)}}\right| \\ & = O(n^{-1}) \max_{x,y \in [j/n, j/n+1/n]} |e^{n\ell(x)} - e^{n\ell(y)}| + O(n^{-2}) e^{n\ell(x^*)}. \end{aligned} \tag{10}$$

Fix $j \in A_n$ and let $x, y \in [j/n, j/n + 1/n]$. Note that, for all x ,

$$|e^x - 1| \leq e^{|x|} - 1.$$

We use this, the fact that $\ell''(x^*) < 0$, and the mean value theorem to write

$$\begin{aligned} |e^{n\ell(x)} - e^{n\ell(y)}| &= e^{n\ell(x^*)} e^{n(\ell(y)-\ell(x^*))} |e^{n(\ell(x)-\ell(y))} - 1| \\ &= e^{n\ell(x^*)} \exp\left(n \frac{\ell''(x^*)}{2} (y-x^*)^2 + n \frac{\ell'''(\xi)}{6} (y-x^*)^3\right) \\ & \quad \times \left|\exp\left(n\ell'(y)(x-y) + \frac{n\ell''(v)}{2} (x-y)^2\right) - 1\right| \end{aligned}$$

$$\begin{aligned}
 &= e^{n\ell(x^*)} \exp\left(n \frac{\ell''(x^*)}{2} (y - x^*)^2 + n \frac{\ell'''(\xi)}{6} (y - x^*)^3\right) \\
 &\quad \times \left| \exp\left(n\ell''(\zeta)(y - x^*)(x - y) + \frac{n\ell''(\nu)}{2} (x - y)^2\right) - 1 \right| \\
 &\leq e^{n\ell(x^*)} \exp\left(n \frac{-|\ell''(x^*)|}{2} (y - x^*)^2 + n \frac{\ell'''(\xi)}{6} (y - x^*)^3\right) \\
 &\quad \times \left(\exp\left(n|\ell''(\zeta)||y - x^*||x - y| + \frac{n|\ell''(\nu)|}{2} (x - y)^2\right) - 1 \right), \tag{11}
 \end{aligned}$$

where ξ and ζ are between y and x^* , and ν is between y and x . Observe that

$$\begin{aligned}
 &\exp\left(n \frac{-|\ell''(x^*)|}{2} (y - x^*)^2 + n \frac{\ell'''(\xi)}{6} (y - x^*)^3\right) \\
 &= \begin{cases} O\left(\exp\left(-\frac{|\ell''(x^*)|}{2} n^{1-2q}\right)\right) (1 + O(n)), & j \notin B_n, \\ 1 + O(n^{1-3q}), & j \in B_n, \end{cases}
 \end{aligned}$$

and that

$$\exp\left(n|\ell''(\zeta)||y - x^*||x - y| + \frac{n|\ell''(\nu)|}{2} (x - y)^2\right) - 1 = \begin{cases} O(1), & j \notin B_n, \\ O(n^{-q}), & j \in B_n. \end{cases}$$

Let $t = 1 - 2q > 0$ and $\omega \in (0, |\ell''(x^*)|/2)$. From the last three displays, we see that

$$\max_{x, y \in [j/n, (j+1)/n]} |e^{n\ell(x)} - e^{n\ell(y)}| = \begin{cases} e^{n\ell(x^*)} O(\exp(-\omega n^t)), & j \notin B_n, \\ e^{n\ell(x^*)} O(n^{-q}), & j \in B_n, \end{cases}$$

and so, from (10),

$$\begin{aligned}
 &\left| \frac{1}{n} \sqrt{\frac{1}{(j/n)(1 - j/n)}} e^{n\ell(j/n)} - \int_{j/n}^{(j+1)/n} \sqrt{\frac{1}{x(1 - x)}} e^{n\ell(x)} dx \right| \\
 &= \begin{cases} e^{n\ell(x^*)} O(\exp(-\omega n^t)), & j \notin B_n, \\ e^{n\ell(x^*)} O(n^{-1-q}), & j \in B_n. \end{cases} \tag{12}
 \end{aligned}$$

Observe that

$$|B_n| = O(n^{1-q}), \quad |A_n \setminus B_n| = O(n). \tag{13}$$

Now, from (12), for any $r < 1$,

$$\begin{aligned}
 &\left| \frac{1}{n} \sum_{i \in A_n} \sqrt{\frac{1}{(i/n)(1 - i/n)}} e^{n\ell(i/n)} - \int_0^1 \sqrt{\frac{1}{x(1 - x)}} e^{n\ell(x)} dx \right| \\
 &\leq e^{n\ell(x^*)} (|B_n| O(n^{-1-q}) + |A_n \setminus B_n| O(\exp(-\omega n^t))) \\
 &\quad + \int_{[0,1] \setminus [1/n, 1-1/n]} \sqrt{\frac{1}{x(1 - x)}} e^{n\ell(x)} dx \\
 &\leq e^{n\ell(x^*)} (O(n^{-2q}) + O(n \exp(-\omega n^t))) + O(e^{n\ell(x^*)-\delta}) \\
 &\leq e^{n\ell(x^*)} O(n^{-r}). \tag{14}
 \end{aligned}$$

Now, by (14) and Proposition 3,

$$\left| \frac{1}{n} \sum_{i \in A_n} \sqrt{\frac{1}{(i/n)(1-i/n)}} e^{n\ell(i/n)} - \int_0^1 \sqrt{\frac{1}{x(1-x)}} e^{n\ell(x)} dx \right| \left(\int_0^1 \sqrt{\frac{1}{x(1-x)}} e^{n\ell(x)} dx \right)^{-1} = O(n^{1/2-r}).$$

Thus,

$$\frac{1}{n} \sum_{i \in A_n} \sqrt{\frac{1}{(i/n)(1-i/n)}} e^{n\ell(i/n)} = (1 + O(n^{1/2-r})) \int_0^1 \sqrt{\frac{1}{x(1-x)}} e^{n\ell(x)} dx.$$

Now, from (9), we conclude that

$$\mathbb{E} \left[\exp \left(\beta_1 W + \frac{\beta_2}{n^{p-1}} W^p \right) \right] = (1 + O(n^{1/2-r})) 2^{-n} \sqrt{\frac{n}{2\pi}} \int_0^1 \sqrt{\frac{1}{x(1-x)}} e^{n\ell(x)} dx.$$

Next, consider (β_1, β_2) on the transition curve away from the critical point. By Theorem 2, there are two maximizers of ℓ , say x_1^* and x_2^* . Defining

$$B_n = \left\{ i \in \{1, \dots, n\} : \frac{i}{n} \in (x_1^* - n^{-q}, x_1^* + n^{-q}) \cup (x_2^* - n^{-q}, x_2^* + n^{-q}) \right\},$$

it is not difficult to see that the arguments above can be repeated to obtain the same result.

Finally, consider the case at the critical point. Here, (10) still holds, but (11) needs to be modified, as follows. By Proposition 4, we have $\ell'(x^*) = \ell''(x^*) = \ell'''(x^*) = 0$ and $\ell^{(4)}(x^*) < 0$, so by the mean value theorem, we have

$$\begin{aligned} & |e^{n\ell(x)} - e^{n\ell(y)}| \\ &= e^{n\ell(x^*)} e^{n(\ell(y)-\ell(x^*))} |e^{n(\ell(x)-\ell(y))} - 1| \\ &= e^{n\ell(x^*)} \exp \left(n \frac{\ell^{(4)}(x^*)}{4!} (y-x^*)^4 + n \frac{\ell^{(5)}(\xi)}{5!} (y-x^*)^5 \right) \\ &\quad \times \left| \exp \left(n\ell'(y)(x-y) + \frac{n\ell''(v)}{2} (x-y)^2 \right) - 1 \right| \\ &= e^{n\ell(x^*)} \exp \left(n \frac{\ell^{(4)}(x^*)}{4!} (y-x^*)^4 + n \frac{\ell^{(5)}(\xi)}{5!} (y-x^*)^5 \right) \\ &\quad \times \left| \exp \left(n\ell^{(4)}(\zeta)(v-x^*)(u-x^*)(y-x^*)(x-y) + \frac{n\ell''(v)}{2} (x-y)^2 \right) - 1 \right| \\ &\leq e^{n\ell(x^*)} \exp \left(n \frac{-|\ell^{(4)}(x^*)|}{4!} (y-x^*)^4 + n \frac{\ell^{(5)}(\xi)}{5!} (y-x^*)^5 \right) \\ &\quad \times \left(\exp \left(n|\ell^{(4)}(\zeta)||v-x^*||u-x^*||y-x^*||x-y| + \frac{n|\ell''(v)|}{2} (x-y)^2 \right) - 1 \right), \end{aligned}$$

where u, v, ξ , and ζ are between y and x^* , and v is between y and x . Let B_n be defined as in

the analysis of the case off the critical curve. Let $q \in (\frac{1}{5}, \frac{1}{4})$ and note that

$$\begin{aligned} & \exp\left(n \frac{-|\ell^{(4)}(x^*)|}{4!} (y - x^*)^4 + n \frac{\ell^{(5)}(\xi)}{5!} (y - x^*)^5\right) \\ &= \begin{cases} O\left(\exp\left(\frac{-|\ell^{(4)}(x^*)|}{4!} n^{1-4q}\right)\right) (1 + O(n)), & j \notin B_n, \\ 1 + O(n^{1-5q}), & j \in B_n, \end{cases} \end{aligned}$$

and that

$$\begin{aligned} & \exp\left(n|\ell^{(4)}(\zeta)||v - x^*||u - x^*||y - x^*||x - y| + \frac{n|\ell''(v)|}{2}(x - y)^2\right) - 1 \\ &= \begin{cases} O(1), & j \notin B_n, \\ O(n^{-3q}), & j \in B_n. \end{cases} \end{aligned}$$

Let $\omega \in (0, |\ell^{(4)}(x^*)|/4!)$ and $t = 1 - 4q > 0$. From the last three displays, we see that

$$\max_{x, y \in [j/n, j/n+1/n]} |e^{n\ell(x)} - e^{n\ell(y)}| = \begin{cases} e^{n\ell(x^*)} O(\exp(-\omega n^t)), & j \notin B_n, \\ e^{n\ell(x^*)} O(n^{-3q}), & j \in B_n. \end{cases}$$

So, from (10),

$$\begin{aligned} & \left| \frac{1}{n} \sqrt{\frac{1}{(j/n)(1 - j/n)}} e^{n\ell(j/n)} - \int_{j/n}^{j/n+1/n} \sqrt{\frac{1}{x(1 - x)}} e^{n\ell(x)} dx \right| \\ &= \begin{cases} e^{n\ell(x^*)} O(\exp(-\omega n^t)), & j \notin B_n, \\ e^{n\ell(x^*)} O(n^{-1-3q}), & j \in B_n. \end{cases} \end{aligned} \tag{15}$$

Using (13) and (15), for any $r < 1$,

$$\begin{aligned} & \left| \frac{1}{n} \sum_{i \in A_n} \sqrt{\frac{1}{(i/n)(1 - i/n)}} e^{n\ell(i/n)} - \int_0^1 \sqrt{\frac{1}{x(1 - x)}} e^{n\ell(x)} dx \right| \\ & \leq e^{n\ell(x^*)} (|B_n| O(n^{-1-3q}) + |A_n \setminus B_n| O(\exp(-\omega n^t))) \\ & \quad + \int_{[0,1] \setminus [\varepsilon+1/n, 1-\varepsilon-1/n]} \sqrt{\frac{1}{x(1 - x)}} e^{n\ell(x)} dx \\ & \leq e^{n\ell(x^*)} (O(n^{-4q}) + O(\exp(-\omega n^t))) + O(e^{n\ell(x^*)-\delta}) \\ & \leq e^{n\ell(x^*)} O(n^{-r}). \end{aligned} \tag{16}$$

Now, by (16) and Proposition 3,

$$\begin{aligned} & \left| \frac{1}{n} \sum_{i \in A_n} \sqrt{\frac{1}{(i/n)(1 - i/n)}} e^{n\ell(i/n)} - \int_0^1 \sqrt{\frac{1}{x(1 - x)}} e^{n\ell(x)} dx \right| \left(\int_0^1 \sqrt{\frac{1}{x(1 - x)}} e^{n\ell(x)} dx \right)^{-1} \\ &= O(n^{1/4-r}). \end{aligned}$$

Thus,

$$\frac{1}{n} \sum_{i \in A_n} \sqrt{\frac{1}{(i/n)(1-i/n)}} e^{n\ell(i/n)} = (1 + O(n^{1/4-r})) \int_0^1 \sqrt{\frac{1}{x(1-x)}} e^{n\ell(x)} dx.$$

Now, from (9), we conclude that

$$\mathbb{E} \left[\exp \left(\beta_1 W + \frac{\beta_2}{n^{p-1}} W^p \right) \right] = (1 + O(n^{1/4-r})) 2^{-n} \sqrt{\frac{n}{2\pi}} \int_0^1 \sqrt{\frac{1}{x(1-x)}} e^{n\ell(x)} dx.$$

This completes the proof. □

Proof of Proposition 3. We will prove only (i) and (iii), as (ii) is standard.

(i) Note that, for $b > 0$ and $k \in \mathbb{N}$,

$$\int_{-\infty}^{\infty} x^k e^{-bx^2} dx = \begin{cases} 0, & k \text{ odd,} \\ \Gamma\left(\frac{k+1}{2}\right) b^{-(k+1)/2} & k \text{ even.} \end{cases}$$

So, for any $\delta > 0$,

$$\begin{aligned} \int_{-\delta}^{\delta} u^k e^{-nbu^2} du &= n^{-(k+1)/2} \int_{-\delta n^{1/2}}^{\delta n^{1/2}} x^k e^{-bx^2} dx \\ &= n^{-(k+1)/2} \left(O(e^{-bn}) + \int_{-\infty}^{\infty} x^k e^{-bx^2} dx \right) \\ &= \begin{cases} 0, & k \text{ odd,} \\ \Gamma\left(\frac{k+1}{2}\right) (nb)^{-(k+1)/2} + O(e^{-bn}), & k \text{ even.} \end{cases} \end{aligned}$$

Now let $c = x^*$ and $u = x - c$, and set $0 < \delta < \min\{c, 1 - c\}$. We use Taylor expansions of x^k and $\ell(x)$ at c , and of e^x at 0, along with Proposition 4, to compute

$$\begin{aligned} &\int_{c-\delta}^{c+\delta} f(x) e^{n\ell(x)} dx \\ &= \int_{-\delta}^{\delta} [d_0 + d_1 u + \dots] e^{n(b_0 + b_1 u + b_2 u^2 + \dots)} du \\ &= e^{n\ell(c)} \int_{-\delta}^{\delta} [d_0 + d_1 u + \dots] e^{nb_2 u^2 + nb_3 u^3 + \dots} du \\ &= e^{n\ell(c)} \int_{-\delta}^{\delta} [d_0 + d_1 u + \dots] \left[1 + (nb_3 u^3 + \dots) + \frac{1}{2} (nb_3 u^3 + \dots)^2 + \dots \right] e^{nb_2 u^2} du \\ &= e^{n\ell(c)} [n^{-1/2} d_0 \alpha_1 + n^{-3/2} \Lambda + O(n^{-5/2})], \end{aligned}$$

where the last step is obtained by collecting terms of the same order, and the interchange of sum and integral is justified by the dominated convergence theorem. Since $x^* = c$ is the unique global maximizer of ℓ , we conclude that, for some $\varepsilon > 0$,

$$\int_0^1 f(x) e^{n\ell(x)} dx = \int_{c-\delta}^{c+\delta} f(x) e^{n\ell(x)} dx + O(e^{n\ell(c)-\varepsilon}).$$

It follows that

$$\int_0^1 f(x) e^{n\ell(x)} dx = e^{n\ell(c)} [n^{-1/2} d_0 \alpha_1 + n^{-3/2} \Lambda + O(n^{-5/2})].$$

(iii) Note that, for $b > 0$ and $k \in \mathbb{N}$,

$$\int_{-\infty}^{\infty} x^k e^{-bx^4} dx = \begin{cases} 0, & k \text{ odd,} \\ \frac{1}{2} \Gamma\left(\frac{k+1}{4}\right) b^{-(k+1)/4}, & k \text{ even.} \end{cases}$$

So, for any $\delta > 0$,

$$\begin{aligned} \int_{-\delta}^{\delta} u^k e^{-nbu^4} du &= n^{-(k+1)/4} \int_{-\delta n^{1/4}}^{\delta n^{1/4}} x^k e^{-bx^4} dx \\ &= n^{-(k+1)/4} \left(O(e^{-bn}) + \int_{-\infty}^{\infty} x^k e^{-bx^4} dx \right) \\ &= \begin{cases} 0, & k \text{ odd,} \\ \frac{1}{2} \Gamma\left(\frac{k+1}{4}\right) (nb)^{-(k+1)/4} + O(e^{-bn}), & k \text{ even.} \end{cases} \end{aligned}$$

As before, we let $c = x^*$, $u = x - c$, set $0 < \delta = \min\{c, 1 - c\}$, and use Taylor expansions of x^k and $\ell(x)$ at c , and e^x at 0, along with Proposition 4, to write

$$\begin{aligned} &\int_{c-\delta}^{c+\delta} f(x) e^{n\ell(x)} dx \\ &= \int_{c-\delta}^{c+\delta} [d_0 + d_1 u + \dots] e^{n(b_0 + b_1 u + b_2 u^2 + \dots)} du \\ &= e^{n\ell(c)} \int_{c-\delta}^{c+\delta} [d_0 + d_1 u + \dots] e^{nb_4 u^4 + nb_5 u^5 + \dots} du \\ &= e^{n\ell(c)} \int_{c-\delta}^{c+\delta} [d_0 + d_1 u + \dots] [1 + (nb_5 u^5 + \dots) + \frac{1}{2}(nb_5 u^5 + \dots)^2 + \dots] e^{nb_4 u^4} du \\ &= e^{n\ell(c)} [n^{-1/4} d_0 \gamma_1 + n^{-3/4} \Theta + O(n^{-5/4})], \end{aligned}$$

where again the last step is obtained by collecting terms of the same order, and the interchange of sum and integral is justified by the dominated convergence theorem. As before, since $x^* = c$ is the unique global maximizer of ℓ , we conclude that

$$\int_0^1 f(x) e^{n\ell(x)} dx = e^{n\ell(c)} [n^{-1/4} d_0 \gamma_1 + n^{-3/4} \Theta + O(n^{-5/4})]. \quad \square$$

The remainder of the proofs are for the results in Section 2.

Proof of Theorem 1. By Propositions 2 and 3, we have

$$\begin{aligned} \psi_n(\beta_1, \beta_2) &= n^{-2} \log Z_n(\beta_1, \beta_2) \\ &= \log 2 + n^{-1} \log \mathbb{E} \left[\exp \left(\beta_1 W + \frac{\beta_2}{n^{p-1}} W^p \right) \right] \end{aligned}$$

$$\begin{aligned}
 &= O(n^{-1} \log n) + \frac{1}{n} \log \int_0^1 \sqrt{\frac{1}{x(1-x)}} e^{n\ell(x)} dx \\
 &= O(n^{-1} \log n) + \ell(x^*),
 \end{aligned}
 \tag{17}$$

as required. □

Proof of Theorem 3. (i) Along the phase transition curve, we have

$$\beta_1 + pq(\beta_1)(x_1^*)^{p-1} - \log\left(\frac{x_1^*}{1-x_1^*}\right) = 0,
 \tag{18}$$

$$\beta_1 + pq(\beta_1)(x_2^*)^{p-1} - \log\left(\frac{x_2^*}{1-x_2^*}\right) = 0,
 \tag{19}$$

$$\begin{aligned}
 &\beta_1 x_1^* + q(\beta_1)(x_1^*)^p - x_1^* \log x_1^* - (1-x_1^*) \log(1-x_1^*) \\
 &= \beta_1 x_2^* + q(\beta_1)(x_2^*)^p - x_2^* \log x_2^* - (1-x_2^*) \log(1-x_2^*).
 \end{aligned}$$

Let $x_1^* < x_2^*$ be the two local maximizers of ℓ in the V-shaped region [23] that contains the phase transition curve except the critical point. By Proposition 4, $\ell''(x_1^*)$ and $\ell''(x_2^*)$ are nonzero away from the critical point. The implicit function theorem implies that x_1^* and x_2^* are analytic functions of both β_1 and β_2 . Differentiating (4) with respect to β_1 and using (18) and (19), we see that

$$x_1^* + q'(\beta_1)(x_1^*)^p = x_2^* + q'(\beta_1)(x_2^*)^p,$$

which implies that

$$q'(\beta_1) = -\frac{x_1^* - x_2^*}{(x_1^*)^p - (x_2^*)^p}.
 \tag{20}$$

As $\beta_1 \rightarrow \beta_1^c$, $x_2^* - x_1^* \rightarrow 0$ and both x_2^* and x_1^* converge to the common maximizer $x_c^* = (p-1)/p$. Therefore,

$$\lim_{\beta_1 \rightarrow \beta_1^c} q'(\beta_1) = -\frac{1}{p(x_c^*)^{p-1}} = -\frac{p^{p-2}}{(p-1)^{p-1}}.$$

Since $x_1^* \rightarrow 0$ and $x_2^* \rightarrow 1$ as $\beta_1 \rightarrow -\infty$, we obtain $\lim_{\beta_1 \rightarrow -\infty} q'(\beta_1) = -1$.

(ii) Differentiating $q'(\beta_1)$ with respect to β_1 , we obtain

$$\begin{aligned}
 q''(\beta_1) &= -\frac{1}{((x_1^*)^p - (x_2^*)^p)^2} [(1-p)(x_1^*)^p + p(x_1^*)^{p-1}x_2^* - (x_2^*)^p] \frac{\partial x_1^*}{\partial \beta_1} \\
 &\quad - \frac{1}{((x_1^*)^p - (x_2^*)^p)^2} [(1-p)(x_2^*)^p + p(x_2^*)^{p-1}x_1^* - (x_1^*)^p] \frac{\partial x_2^*}{\partial \beta_1}.
 \end{aligned}
 \tag{21}$$

Differentiating (18) and (19) with respect to β_1 , we obtain

$$1 + pq'(\beta_1)(x_1^*)^{p-1} + \left[pq(\beta_1)(p-1)(x_1^*)^{p-2} - \frac{1}{x_1^*(1-x_1^*)} \right] \frac{\partial x_1^*}{\partial \beta_1} = 0,
 \tag{22}$$

$$1 + pq'(\beta_1)(x_2^*)^{p-1} + \left[pq(\beta_1)(p-1)(x_2^*)^{p-2} - \frac{1}{x_2^*(1-x_2^*)} \right] \frac{\partial x_2^*}{\partial \beta_1} = 0.
 \tag{23}$$

Note that, from (20),

$$\begin{aligned}
 1 + pq'(\beta_1)(x_1^*)^{p-1} &= 1 - p \frac{x_1^* - x_2^*}{(x_1^*)^p - (x_2^*)^p} (x_1^*)^{p-1} \\
 &= 1 - \frac{p(x_1^*)^{p-1}}{(x_1^*)^{p-1} + (x_1^*)^{p-2}x_2^* + \dots + x_1^*(x_2^*)^{p-2} + (x_2^*)^{p-1}} \\
 &> 1 - \frac{p(x_1^*)^{p-1}}{(x_1^*)^{p-1} + (x_1^*)^{p-2}x_1^* + \dots + x_1^*(x_1^*)^{p-2} + (x_1^*)^{p-1}} \\
 &= 0,
 \end{aligned}
 \tag{24}$$

and, analogously,

$$1 + pq'(\beta_1)(x_2^*)^{p-1} = 1 - p \frac{x_1^* - x_2^*}{(x_1^*)^p - (x_2^*)^p} (x_2^*)^{p-1} < 0.
 \tag{25}$$

Moreover, in Proposition 4, we showed that

$$\ell''(x_1^*) = pq(\beta_1)(p - 1)(x_1^*)^{p-2} - \frac{1}{x_1^*(1 - x_1^*)} < 0,
 \tag{26}$$

$$\ell''(x_2^*) = pq(\beta_1)(p - 1)(x_2^*)^{p-2} - \frac{1}{x_2^*(1 - x_2^*)} < 0.
 \tag{27}$$

Therefore, from (22)–(27), we conclude that $\partial x_1^*/\partial \beta_1 > 0$ and $\partial x_2^*/\partial \beta_1 < 0$. Finally, by noting that in (21),

$$(1 - p)(x_1^*)^p + p(x_1^*)^{p-1}x_2^* - (x_2^*)^p < 0, \quad (1 - p)(x_2^*)^p + p(x_2^*)^{p-1}x_1^* - (x_1^*)^p > 0,$$

we conclude that $q''(\beta_1) > 0$. □

In the proofs below, let $d_m^{(n)}$ be defined as in Proposition 3 for the function

$$f(x) = \frac{x^n}{\sqrt{x(1 - x)}}.$$

Proof of Theorem 4. Off the phase transition curve, the result follows immediately from Theorem 1 and [23]. Thus, we prove only the last two displays in Theorem 4.

From the second line of (17), we have

$$\begin{aligned}
 \frac{\partial^2}{\partial \beta_1^2} \psi_n(\beta_1, \beta_2) &= n^{-1} \left\{ \frac{\mathbb{E}[W^2 \exp(\beta_1 W + (\beta_2/n^{p-1})W^p)]}{\mathbb{E}[\exp(\beta_1 W + (\beta_2/n^{p-1})W^p)]} \right. \\
 &\quad \left. - \left(\frac{\mathbb{E}[W \exp(\beta_1 W + (\beta_2/n^{p-1})W^p)]}{\mathbb{E}[\exp(\beta_1 W + (\beta_2/n^{p-1})W^p)]} \right)^2 \right\}.
 \end{aligned}
 \tag{28}$$

We use Proposition 2 and Proposition 3 to estimate each of the terms in (28).

We first consider the case on the transition curve excluding the critical point. By Theorem 2, there are two global maximizers $x_1^* < x_2^*$ of ℓ . We write $\ell(x_1^*) = \ell(x_2^*) = \ell(x^*)$. By Proposition 3 and Proposition 2, for any $r < 1$, we have

$$\begin{aligned}
 &\mathbb{E} \left[W^k \exp \left(\beta_1 W + \frac{\beta_2}{n^{p-1}} W^p \right) \right] \\
 &= [1 + O(n^{1/2-r})] \frac{n^k 2^{-n} \sqrt{n}}{\sqrt{2\pi}} \int_0^1 \sqrt{\frac{x^{2k}}{x(1-x)}} e^{n\ell(x)} dx
 \end{aligned}$$

$$\begin{aligned}
 &= [1 + O(n^{1/2-r})] \frac{n^k 2^{-n} \sqrt{n} e^{n\ell(x^*)}}{\sqrt{2\pi} \sqrt{n}} \\
 &\quad \times \left[\frac{\sqrt{(x_1^*)^{2k}/x_1^*(1-x_1^*)}}{\sqrt{2\pi \ell''(x_1^*)}} + \frac{\sqrt{(x_2^*)^{2k}/x_2^*(1-x_2^*)}}{\sqrt{2\pi \ell''(x_2^*)}} + O(n^{-1}) \right] \\
 &= \frac{n^k 2^{-n} e^{n\ell(x^*)}}{2\pi} \left[\frac{(x_1^*)^k}{\sqrt{x_1^*(1-x_1^*)\ell''(x_1^*)}} + \frac{(x_2^*)^k}{\sqrt{x_2^*(1-x_2^*)\ell''(x_2^*)}} + O(n^{1/2-r}) \right]. \tag{29}
 \end{aligned}$$

Hence,

$$\begin{aligned}
 &\frac{\partial^2}{\partial \beta_1^2} \psi_n(\beta_1, \beta_2) \\
 &= n^{-1} n^2 \left\{ \left[\frac{(x_1^*)^2}{\sqrt{x_1^*(1-x_1^*)|\ell''(x_1^*)|}} + \frac{(x_2^*)^2}{\sqrt{x_2^*(1-x_2^*)|\ell''(x_2^*)|}} \right] \right. \\
 &\quad \times \left. \left[\frac{1}{\sqrt{x_1^*(1-x_1^*)|\ell''(x_1^*)|}} + \frac{1}{\sqrt{x_2^*(1-x_2^*)|\ell''(x_2^*)|}} \right]^{-1} \right\} \\
 &\quad - n^{-1} n^2 \left\{ \left(\frac{x_1^*}{\sqrt{x_1^*(1-x_1^*)|\ell''(x_1^*)|}} + \frac{x_2^*}{\sqrt{x_2^*(1-x_2^*)|\ell''(x_2^*)|}} \right)^2 \right. \\
 &\quad \times \left. \left(\frac{1}{\sqrt{x_1^*(1-x_1^*)|\ell''(x_1^*)|}} + \frac{1}{\sqrt{x_2^*(1-x_2^*)|\ell''(x_2^*)|}} \right)^{-2} \right\} + O(n^{3/2-r}) \\
 &= n \left\{ \frac{(x_1^* - x_2^*)^2}{\sqrt{x_1^*(1-x_1^*)|\ell''(x_1^*)|} \sqrt{x_2^*(1-x_2^*)|\ell''(x_2^*)|}} \right. \\
 &\quad \times \left. \left(\frac{1}{\sqrt{x_1^*(1-x_1^*)|\ell''(x_1^*)|}} + \frac{1}{\sqrt{x_2^*(1-x_2^*)|\ell''(x_2^*)|}} \right)^{-2} \right\} + O(n^{3/2-r}) \\
 &= n \frac{(x_1^* - x_2^*)^2 \sqrt{x_1^*(1-x_1^*)|\ell''(x_1^*)|} \sqrt{x_2^*(1-x_2^*)|\ell''(x_2^*)|}}{(\sqrt{x_1^*(1-x_1^*)|\ell''(x_1^*)|} + \sqrt{x_2^*(1-x_2^*)|\ell''(x_2^*)|})^2} + O(n^{3/2-r}).
 \end{aligned}$$

Next consider the case at the critical point. By Proposition 3 and Proposition 2, for any $r < 1$,

$$\begin{aligned}
 \mathbb{E} \left[W^k \exp \left(\beta_1 W + \frac{\beta_2}{n^{p-1}} W^p \right) \right] &= [1 + O(n^{1/4-r})] \frac{n^k 2^{-n} \sqrt{n}}{\sqrt{2\pi}} \int_0^1 \sqrt{\frac{x^{2k}}{x(1-x)}} e^{n\ell(x)} dx \\
 &= \frac{n^k 2^{-n} \sqrt{n}}{\sqrt{2\pi}} e^{n\ell(x^*)} [n^{-1/4} d_0^{(k)} \gamma_1 + n^{-3/4} \Theta^{(k)} + O(n^{-r})],
 \end{aligned}$$

where

$$\Theta^{(k)} := d_2^{(k)} \gamma_3 + d_1^{(k)} b_5 \gamma_7 + d_0^{(k)} b_6 \gamma_7 + \frac{1}{2} d_0^{(k)} b_5^2 \gamma_{11}, \quad k = 0, 1, 2.$$

Then

$$\begin{aligned}
 d_0^{(0)} &= \frac{1}{\sqrt{x^*(1-x^*)}}, & d_0^{(1)} &= \frac{x^*}{\sqrt{x^*(1-x^*)}}, & d_0^{(2)} &= \frac{(x^*)^2}{\sqrt{x^*(1-x^*)}}, \\
 d_1^{(0)} &= \frac{x^* - 1/2}{(x^*(1-x^*))^{3/2}}, & d_1^{(1)} &= \frac{x^*/2}{(x^*(1-x^*))^{3/2}}, & d_1^{(2)} &= \frac{(3/2)(x^*)^2 - (x^*)^3}{(x^*(1-x^*))^{3/2}}, \\
 d_2^{(0)} &= \frac{2(x^*)^2 - 2x^* + 3/4}{2(x^*(1-x^*))^{5/2}}, & d_2^{(1)} &= \frac{(x^*)^2 - x^*/4}{2(x^*(1-x^*))^{5/2}}, & d_2^{(2)} &= \frac{(3/4)(x^*)^2}{2(x^*(1-x^*))^{5/2}}.
 \end{aligned}$$

It is easy to observe that $d_0^{(2)}d_0^{(0)} = (d_0^{(1)})^2$. By differentiating this identity, we obtain $d_1^{(2)}d_0^{(0)} + d_0^{(2)}d_1^{(0)} = 2d_1^{(1)}d_0^{(1)}$. Therefore, by (28) and (29),

$$\begin{aligned} \frac{\partial^2}{\partial \beta_1^2} \psi_n(\beta_1, \beta_2) &= n^{-1}n^2 \frac{n^{-1/4}d_0^{(2)}\gamma_1 + n^{-3/4}\Theta^{(2)}}{n^{-1/4}d_0^{(0)}\gamma_1 + n^{-3/4}\Theta^{(0)}} - n^{-1}n^2 \frac{(n^{-1/4}d_0^{(1)}\gamma_1 + n^{-3/4}\Theta^{(1)})^2}{(n^{-1/4}d_0^{(0)}\gamma_1 + n^{-3/4}\Theta^{(0)})^2} \\ &\quad + O(n^{5/4-r}) \\ &= n \frac{n^{-1}\gamma_1[d_0^{(2)}\Theta^{(0)} + d_0^{(0)}\Theta^{(2)} - 2d_0^{(1)}\Theta^{(1)}] + O(n^{-3/2})}{n^{-1/2}(d_0^{(0)})^2\gamma_1^2} + O(n^{5/4-r}) \\ &= \frac{n^{1/2}}{(d_0^{(0)})^2\gamma_1} [\gamma_3(d_0^{(2)}d_2^{(0)} + d_0^{(0)}d_2^{(2)} - 2d_0^{(1)}d_2^{(1)})] \\ &\quad + \frac{n^{1/2}}{(d_0^{(0)})^2\gamma_1} [b_5\gamma_7(d_0^{(2)}d_1^{(0)} + d_0^{(0)}d_1^{(2)} - 2d_0^{(1)}d_1^{(1)})] + O(n^{5/4-r}) \\ &= \frac{n^{1/2}\gamma_3}{(d_0^{(0)})^2\gamma_1} (d_0^{(2)}d_2^{(0)} + d_0^{(0)}d_2^{(2)} - 2d_0^{(1)}d_2^{(1)}) + O(n^{5/4-r}) \\ &= n^{1/2} \frac{\gamma_3}{\gamma_1} + O(n^{5/4-r}) \\ &= n^{1/2} \frac{\Gamma(3/4)}{\Gamma(1/4)} \frac{1}{\sqrt{\ell^{(4)}(x^*)/4!}} + O(n^{5/4-r}) \\ &= n^{1/2} \frac{\Gamma(3/4)}{\Gamma(1/4)} \frac{2\sqrt{6}(p-1)}{p^{5/2}} + O(n^{5/4-r}), \end{aligned}$$

where we used Proposition 4 in the last line. □

Proof of Theorem 5. We prove only the last two displays in Theorem 5, since the first display follows immediately from Theorem 1 and the results of [23]. From the second line of (17), we have

$$\begin{aligned} \frac{\partial^2}{\partial \beta_2^2} \psi_n(\beta_1, \beta_2) &= n^{-1} \left\{ \frac{\mathbb{E}[(W^{2p}/n^{2(p-1)}) \exp(\beta_1 W + (\beta_2/n^{p-1}) W^p)]}{\mathbb{E}[\exp(\beta_1 W + (\beta_2/n^{p-1}) W^p)]} \right. \\ &\quad \left. - \left(\frac{\mathbb{E}[(W^p/n^{p-1}) \exp(\beta_1 W + (\beta_2/n^{p-1}) W^p)]}{\mathbb{E}[\exp(\beta_1 W + (\beta_2/n^{p-1}) W^p)]} \right)^2 \right\}. \end{aligned}$$

Consider first the case on the phase transition curve excluding the critical point. Then, similar to the proof of Theorem 4, for any $r < 1$,

$$\begin{aligned} \frac{\partial^2}{\partial \beta_2^2} \psi_n(\beta_1, \beta_2) &= n \frac{((x_1^*)^p - (x_2^*)^p)^2 \sqrt{x_1^*(1-x_1^*)|\ell''(x_1^*)|} \sqrt{x_2^*(1-x_2^*)|\ell''(x_2^*)|}}{(\sqrt{x_1^*(1-x_1^*)|\ell''(x_1^*)|} + \sqrt{x_2^*(1-x_2^*)|\ell''(x_2^*)|})^2} + O(n^{3/2-r}). \end{aligned}$$

Now consider the case at the critical point. We have

$$d_0^{(p)} = \frac{(x^*)^p}{\sqrt{x^*(1-x^*)}}, \quad d_1^{(p)} = \frac{(p-1/2)(x^*)^p - (p-1)(x^*)^{p+1}}{(x^*(1-x^*))^{3/2}},$$

and

$$d_2^{(p)} = \frac{(p^2 - 2p + 3/4)(x^*)^p - (2p^2 - 5p + 2)(x^*)^{p+1} + (p^2 - 3p + 2)(x^*)^{p+2}}{2(x^*(1 - x^*))^{5/2}}.$$

It is easy to observe that $d_0^{(2p)}d_0^{(0)} = (d_0^{(p)})^2$. By differentiating this identity, we obtain

$$d_1^{(2p)}d_0^{(0)} + d_0^{(2p)}d_1^{(0)} = 2d_1^{(p)}d_0^{(p)}.$$

Similar to the proof of Theorem 4, for any $r < 1$,

$$\begin{aligned} & \frac{\partial^2}{\partial \beta_1^2} \psi_n(\beta_1, \beta_2) \\ &= n \frac{(n^{-1/4}d_0^{(2p)}\gamma_1 + n^{-3/4}\Theta^{(2p)})(n^{-1/4}d_0^{(0)}\gamma_1 + n^{-3/4}\Theta^{(0)}) - (n^{-1/4}d_0^{(p)}\gamma_1 + n^{-3/4}\Theta^{(p)})^2}{(n^{-1/4}d_0^{(0)}\gamma_1 + n^{-3/4}\Theta^{(0)})^2} \\ & \quad + O(n^{5/4-r}) \\ &= \frac{n^{1/2}\gamma_3}{(d_0^{(0)})^2\gamma_1} (d_0^{(2p)}d_2^{(0)} + d_0^{(0)}d_2^{(2p)} - 2d_0^{(p)}d_2^{(p)}) + O(n^{5/4-r}) \\ &= p^2(x^*)^{2p-2} \frac{\gamma_3}{\gamma_1} n^{1/2} + O(n^{5/4-r}) \\ &= n^{1/2} p^2 \left(\frac{p-1}{p} \right)^{2p-2} \frac{\Gamma(3/4)}{\Gamma(1/4)} \frac{2\sqrt{6}(p-1)}{p^{5/2}} + O(n^{5/4-r}). \quad \square \end{aligned}$$

Proof of Theorem 6. Again we prove only the last two displays in the theorem. From the second line of (17), we have

$$\begin{aligned} & \frac{\partial^2}{\partial \beta_1 \partial \beta_2} \psi_n(\beta_1, \beta_2) \\ &= n^{-1} \frac{\mathbb{E}[W(W^p/n^{p-1}) \exp(\beta_1 W + (\beta_2/n^{p-1})W^p)]}{\mathbb{E}[\exp(\beta_1 W + (\beta_2/n^{p-1})W^p)]} \\ & \quad - \frac{\mathbb{E}[W \exp(\beta_1 W + (\beta_2/n^{p-1})W^p)] \mathbb{E}[(W^p/n^{p-1}) \exp(\beta_1 W + (\beta_2/n^{p-1})W^p)]}{(\mathbb{E}[\exp(\beta_1 W + (\beta_2/n^{p-1})W^p)])^2}. \end{aligned}$$

Similar to the proof of Theorem 4, on the phase transition curve excluding the critical point, for any $r < 1$,

$$\begin{aligned} & \frac{\partial^2}{\partial \beta_1^2} \psi_n(\beta_1, \beta_2) \\ &= n \frac{((x_1^*)^p - (x_2^*)^p)(x_1^* - x_2^*) \sqrt{x_1^*(1 - x_1^*)} |\ell''(x_1^*)| \sqrt{x_2^*(1 - x_2^*)} |\ell''(x_2^*)|}{(\sqrt{x_1^*(1 - x_1^*)} |\ell''(x_1^*)| + \sqrt{x_2^*(1 - x_2^*)} |\ell''(x_2^*)|)^2} + O(n^{3/2-r}). \end{aligned}$$

Consider now the case at the critical point. It is easy to observe that $d_0^{(p+1)}d_0^{(0)} = (d_0^{(1)})(d_0^{(p)})$. By differentiating this identity, we obtain

$$d_1^{(p+1)}d_0^{(0)} + d_0^{(p+1)}d_1^{(0)} = d_1^{(1)}d_0^{(p)} + d_0^{(1)}d_1^{(p)}.$$

Therefore, similar to the proof of Theorem 4, we obtain, for any $r < 1$,

$$\begin{aligned} & \frac{\partial^2}{\partial \beta_1^2} \psi_n(\beta_1, \beta_2) \\ &= n \frac{(n^{-1/4} d_0^{(p+1)} \gamma_1 + n^{-3/4} \Theta^{(p+1)})(n^{-1/4} d_0^{(0)} \gamma_1 + n^{-3/4} \Theta^{(0)})}{(n^{-1/4} d_0^{(0)} \gamma_1 + n^{-3/4} \Theta^{(0)})^2} \\ & \quad - n \frac{(n^{-1/4} d_0^{(1)} \gamma_1 + n^{-3/4} \Theta^{(1)})(n^{-1/4} d_0^{(p)} \gamma_1 + n^{-3/4} \Theta^{(p)})}{(n^{-1/4} d_0^{(0)} \gamma_1 + n^{-3/4} \Theta^{(0)})^2} + O(n^{5/4-r}) \\ &= n \frac{n^{-1} \gamma_1 [d_0^{(p+1)} \Theta^{(0)} + d_0^{(0)} \Theta^{(p+1)} - d_0^{(1)} \Theta^{(p)} - d_0^{(p)} \Theta^{(1)}] + O(n^{-3/2})}{n^{-1/2} (d_0^{(0)})^2 \gamma_1^2 + O(n^{-1})} + O(n^{5/4-r}) \\ &= \frac{n^{1/2} \gamma_3}{(d_0^{(0)})^2 \gamma_1} (d_0^{(p+1)} d_2^{(0)} + d_0^{(0)} d_2^{(p+1)} - d_0^{(1)} d_2^{(p)} - d_0^{(p)} d_2^{(1)}) + O(n^{5/4-r}) \\ &= p(x^*)^{p-1} \frac{\gamma_3}{\gamma_1} n^{1/2} + O(n^{5/4-r}) \\ &= p \left(\frac{p-1}{p} \right)^{p-1} \frac{\Gamma(3/4)}{\Gamma(1/4)} \frac{2\sqrt{6}(p-1)}{p^{5/2}} n^{1/2} + O(n^{5/4-r}). \quad \square \end{aligned}$$

Proof of Theorem 7. Observe first that $\mathbb{P}_n(X_{12} = 1) = \mathbb{E}_n[X_{12}] = (1/n)\mathbb{E}_n[\sum_{j=1}^n X_{1j}]$. Thus, off the transition curve, we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}_n(X_{12} = 1) &= \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}_n \left[\sum_{j=1}^n X_{1j} \right] \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \frac{\mathbb{E}[W \exp(\beta_1 W + (\beta_2/n^{p-1}) W^p)]}{\mathbb{E}[\exp(\beta_1 W + (\beta_2/n^{p-1}) W^p)]} \\ &= \lim_{n \rightarrow \infty} \frac{(1 + O(n^{1/2-4q})) 2^{-n} \sqrt{n/2\pi} \int_0^1 \sqrt{x^2/x(1-x)} e^{n\ell(x)} dx}{(1 + O(n^{1/2-4q})) 2^{-n} \sqrt{n/2\pi} \int_0^1 \sqrt{1/x(1-x)} e^{n\ell(x)} dx} \\ &= \lim_{n \rightarrow \infty} \frac{\sqrt{2\pi(x^*)^2/x^*(1-x^*)} |\ell''(x^*)| n^{-1/2} e^{n\ell(x^*)}}{\sqrt{2\pi/x^*(1-x^*)} |\ell''(x^*)| n^{-1/2} e^{n\ell(x^*)}} \\ &= x^*. \end{aligned}$$

Similarly, at the critical point,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}_n(X_{12} = 1) &= \lim_{n \rightarrow \infty} \frac{1}{n} \frac{\mathbb{E}[W \exp(\beta_1 W + (\beta_2/n^{p-1}) W^p)]}{\mathbb{E}[\exp(\beta_1 W + (\beta_2/n^{p-1}) W^p)]} \\ &= \lim_{n \rightarrow \infty} \frac{(1 + O(n^{1/4-4q})) 2^{-n} \sqrt{n/2\pi} \int_0^1 \sqrt{x^2/x(1-x)} e^{n\ell(x)} dx}{(1 + O(n^{1/4-4q})) 2^{-n} \sqrt{n/2\pi} \int_0^1 \sqrt{1/x(1-x)} e^{n\ell(x)} dx} \\ &= \lim_{n \rightarrow \infty} \frac{e^{n\ell(x^*)} n^{-1/4} d_0^{(1)} \gamma_1}{e^{n\ell(x^*)} n^{-1/4} d_0^{(0)} \gamma_1} \\ &= x^*. \end{aligned}$$

Finally, on the phase transition curve except at the critical point,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}_n(X_{12} = 1) &= \lim_{n \rightarrow \infty} \frac{1}{n} \frac{\mathbb{E}[W \exp(\beta_1 W + (\beta_2/n^{p-1})W^p)]}{\mathbb{E}[\exp(\beta_1 W + (\beta_2/n^{p-1})W^p)]} \\ &= \lim_{n \rightarrow \infty} \frac{(\sqrt{2\pi(x_1^*)^2/x_1^*(1-x_1^*)|\ell''(x_1^*)|} + \sqrt{2\pi(x_2^*)^2/x_2^*(1-x_2^*)|\ell''(x_2^*)|})n^{-1/2} e^{n\ell(x^*)}}{(\sqrt{2\pi/x_1^*(1-x_1^*)|\ell''(x_1^*)|} + \sqrt{2\pi/x_2^*(1-x_2^*)|\ell''(x_2^*)|})n^{-1/2} e^{n\ell(x^*)}} \\ &= \frac{x_1^* \sqrt{1/x_1^*(1-x_1^*)|\ell''(x_1^*)|} + x_2^* \sqrt{1/x_2^*(1-x_2^*)|\ell''(x_2^*)|}}{\sqrt{1/x_1^*(1-x_1^*)|\ell''(x_1^*)|} + \sqrt{1/x_2^*(1-x_2^*)|\ell''(x_2^*)|}}. \quad \square \end{aligned}$$

Proof of Theorem 8. (i) The result follow from

$$\lim_{n \rightarrow \infty} \mathbb{P}_n(X_{12} = 1, X_{34} = 1) = \lim_{n \rightarrow \infty} \mathbb{P}_n(X_{12} = 1) \lim_{n \rightarrow \infty} \mathbb{P}_n(X_{34} = 1).$$

(ii) Note that $\mathbb{P}_n(X_{12} = X_{13} = 1) = \mathbb{E}_n[X_{12}X_{13}]$, and

$$\mathbb{E}_n[X_{12}X_{13}] = \frac{1}{n^2 - n} \left(\mathbb{E}_n \left[\left(\sum_{j=1}^n X_{1j} \right)^2 \right] - \mathbb{E}_n \left[\sum_{j=1}^n X_{1j} \right] \right).$$

It follows that off the transition curve, we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}_n(X_{12} = X_{13} = 1) &= \lim_{n \rightarrow \infty} \frac{1}{n^2} \mathbb{E}_n \left[\left(\sum_{j=1}^n X_{1j} \right)^2 \right] \\ &= \lim_{n \rightarrow \infty} \frac{1}{n^2} \frac{\mathbb{E}[W^2 \exp(\beta_1 W + (\beta_2/n^{p-1})W^p)]}{\mathbb{E}[\exp(\beta_1 W + (\beta_2/n^{p-1})W^p)]} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n^2} \frac{(n^2 2^{-n} \sqrt{n}/\sqrt{2\pi}) \int_0^1 \sqrt{x^4/x(1-x)} e^{n\ell(x)} dx}{(2^{-n} \sqrt{n}/\sqrt{2\pi}) \int_0^1 \sqrt{1/x(1-x)} e^{n\ell(x)} dx} \\ &= (x^*)^2. \end{aligned}$$

Similarly, at the critical point,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}_n(X_{12} = X_{13} = 1) &= \lim_{n \rightarrow \infty} \frac{1}{n^2} \frac{\mathbb{E}[W^2 \exp(\beta_1 W + (\beta_2/n^{p-1})W^p)]}{\mathbb{E}[\exp(\beta_1 W + (\beta_2/n^{p-1})W^p)]} \\ &= \lim_{n \rightarrow \infty} \frac{\int_0^1 \sqrt{x^4/x(1-x)} e^{n\ell(x)} dx}{\int_0^1 \sqrt{1/x(1-x)} e^{n\ell(x)} dx} \\ &= \frac{d_0^{(2)} \gamma_1}{d_0^{(0)} \gamma_1} \\ &= (x^*)^2. \end{aligned}$$

Finally, on the phase transition curve except at the critical point,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}_n(X_{12} = X_{13} = 1) &= \lim_{n \rightarrow \infty} \frac{1}{n^2} \frac{\mathbb{E}[W^2 \exp(\beta_1 W + (\beta_2/n^{p-1})W^p)]}{\mathbb{E}[\exp(\beta_1 W + (\beta_2/n^{p-1})W^p)]} \\ &= \lim_{n \rightarrow \infty} \frac{\int_0^1 \sqrt{x^4/x(1-x)} e^{n\ell(x)} dx}{\int_0^1 \sqrt{1/x(1-x)} e^{n\ell(x)} dx} \end{aligned}$$

$$\begin{aligned}
&= \frac{(x_1^*)^2/\sqrt{x_1^*(1-x_1^*)\ell''(x_1^*)} + (x_2^*)^2/\sqrt{x_2^*(1-x_2^*)\ell''(x_2^*)}}{1/\sqrt{x_1^*(1-x_1^*)\ell''(x_1^*)} + 1/\sqrt{x_2^*(1-x_2^*)\ell''(x_2^*)}} \\
&= \alpha(x_1^*)^2 + (1-\alpha)(x_2^*)^2. \quad \square
\end{aligned}$$

Acknowledgements

The authors are grateful to the editor and an anonymous referee for helpful comments and suggestions. D. Aristoff and L. Zhu acknowledge support from the National Science Foundation via the awards NSF-DMS-1522398 and NSF-DMS-1613164. The authors are very grateful to Angelo Mele, Dan Pirjol, Charles Radin, Mei Yin, and S. R. S. Varadhan for helpful discussions.

References

- [1] ARISTOFF, D. AND RADIN, C. (2013). Emergent structures in large networks. *J. Appl. Prob.* **50**, 883–888.
- [2] ARISTOFF, D. AND ZHU, L. (2015). Asymptotic structure and singularities in constrained directed graphs. *Stoch. Process. Appl.* **125**, 4154–4177.
- [3] BADEV, A. (2013). Discrete games in endogenous networks: theory and policy. Working paper. University of Pennsylvania.
- [4] BESAG, J. (1975). Statistical analysis of non-lattice data. *J. R. Statist. Soc. D* **24**, 179–195.
- [5] CHANDRASEKHAR, A. G. AND JACKSON, M. O. (2014). Tractable and consistent random graph models. Preprint. Available at <https://arxiv.org/abs/1210.7375>.
- [6] CHATTERJEE, S. AND DIACONIS, P. (2013). Estimating and understanding exponential random graph models. *Ann. Statist.* **41**, 2428–2461.
- [7] CHATTERJEE, S. AND VARADHAN, S. R. S. (2011). The large deviation principle for the Erdős–Rényi random graph. *Europ. J. Combin.* **32**, 1000–1017.
- [8] CHENG, J., ROMERO, D. M., MEEDER, B. AND KLEINBERG, J. (2011). Predicting reciprocity in social networks. In *Proc. IEEE Third International Conference on Social Computing*, IEEE, pp. 49–56.
- [9] FIENBERG, S. E. (2010). Introduction to papers on the modeling and analysis of network data. *Ann. Appl. Statist.* **4**, 1–4.
- [10] FIENBERG, S. E. (2010). Introduction to papers on the modeling and analysis of network data—II. *Ann. Appl. Statist.* **4**, 533–534.
- [11] FISHER, M. E. AND RADIN, C. (2006). Definition of thermodynamic phases and phase transitions. In *Proc. Phase Transitions*. Available at <http://www.aimath.org/pastworkshops/phasetransition.html>.
- [12] GALLAVOTTI, G. (1999). *Statistical Mechanics: A Short Treatise*. Springer, Berlin.
- [13] HÄGGSTRÖM, O. AND JONASSON, J. (1999). Phase transition in the random triangle model. *J. Appl. Prob.* **36**, 1101–1115.
- [14] HOLLAND, P. W. AND LEINHARDT, S. (1981). An exponential family of probability distributions for directed graphs. *J. Amer. Statist. Assoc.* **76**, 33–65.
- [15] LEBOWITZ, J. L., MAZEL, A. E. AND PRESUTTI, E. (1998). Rigorous proof of a liquid-vapor phase transition in a continuum particle system. *Phys. Rev. Lett.* **80**, 4701–4704.
- [16] LOVÁSZ, L. (2009). Very large graphs. In *Current Developments in Mathematics, 2008*, International Press, Somerville, MA, pp. 67–128.
- [17] MELE, A. (2017). A structural model of dense network formation. *Econometrica* **85**, 825–850.
- [18] MELE, A. AND ZHU, L. (2017). Approximate variational estimation for a model of network formation. Preprint. Available at <https://arxiv.org/abs/1702.00308>.
- [19] NEWMAN, M. E. J. (2010). *Networks: An Introduction*. Oxford University Press.
- [20] NEWMAN, M. E. J., FORREST, S. AND BALTHROP, J. (2002). Email networks and the spread of computer viruses. *Phys. Rev. E* **66**, 035101.
- [21] PARK, J. AND NEWMAN, M. E. J. (2004). Solution of the two-star model of a network. *Phys. Rev. E* **70**, 066146.
- [22] PARK, J. AND NEWMAN, M. E. J. (2005). Solution for the properties of a clustered network. *Phys. Rev. E* **72**, 026136.
- [23] RADIN, C. AND YIN, M. (2013). Phase transitions in exponential random graphs. *Ann. Appl. Prob.* **23**, 2458–2471.
- [24] RINALDO, A., FIENBERG, S. AND ZHOU, Y. (2009). On the geometry of discrete exponential families with application to exponential random graph models. *Electron. J. Statist.* **3**, 446–484.
- [25] SNIJDERS, T. A. B. (2002). Markov chain Monte Carlo estimation of exponential random graph models. *J. Social Structure* **3**, 40pp.

- [26] SNIJDERS, T. A. B., PATTISON, P. E., ROBINS, G. L. AND HANDCOCK, M. S. (2006). New specifications for exponential random graph models. *Sociological Methodol.* **36**, 99–153.
- [27] TRAIN, K. E. (2009). *Discrete Choice Methods with Simulation*, 2nd edn. Cambridge University Press.
- [28] WASSERMAN, S. AND FAUST, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press.
- [29] YANG, C. N. AND LEE, T. D. (1952). Statistical theory of equations of state and phase transitions. I. Theory of condensation. *Phys. Rev. (2)* **87**, 404–409.
- [30] YIN, M. (2013). Critical phenomena in exponential random graphs. *J. Statist. Phys.* **153**, 1008–1021.