# Second-order Approximation of Exponential Random Graph Models

Wen-Yi Ding and Xiao Fang

Soochow University and The Chinese University of Hong Kong

Abstract: Exponential random graph models (ERGMs) are flexible probability models allowing edge dependency. However, it is known that, to a first-order approximation, many ERGMs behave like Erdös–Rényi random graphs, where edges are independent. In this paper, to distinguish ERGMs from Erdös–Rényi random graphs, we consider second-order approximations of ERGMs using two-stars and triangles. We prove that the second-order approximation indeed achieves second-order accuracy in the triangle-free case. The new approximation is formally obtained by Hoeffding decomposition and rigorously justified using Stein's method.

AMS 2020 subject classification: 60B10, 05C80

**Keywords and phrases:** Erdös–Rényi graph, exponential random graphs, Hoeffding decomposition, second-order approximation, Stein's method.

## 1 Introduction

### 1.1 Background

Exponential random graph models (ERGMs) are frequently used as parametric statistical models in network analysis; see, for example, Wasserman and Faust (1994). We refer to Bhamidi et al. (2011) and Chatterjee and Diaconis (2013) for the history and the following formulation of ERGM. Let  $\mathcal{G}_n$  be the space of all simple graphs<sup>1</sup> on n labeled vertices. ERGM assigns probability

$$p_{\beta}(G) = \frac{1}{Z(\beta)} \exp\left\{n^2 \sum_{i=1}^k \beta_i t(H_i, G)\right\}$$
(1.1)

to each  $G \in \mathcal{G}_n$ , where  $\beta = (\beta_1, \dots, \beta_k)$  is a vector of real parameters,  $H_1, \dots, H_k$  are (typically small) simple graphs without isolated vertices, |hom(H, G)| denotes the number of homomorphisms of H into G (i.e., edge-preserving injective maps  $V(H) \to V(G)$ , where V(H) and V(G) are the vertex sets of H and G respectively),  $|\cdot|$  denotes cardinality when applied to a set,

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

<sup>&</sup>lt;sup>1</sup>In this paper, simple graphs mean undirected graphs without self-loops or multiple edges

denotes the homomorphism density, and  $Z(\beta)$  is a normalizing constant. The  $n^2$  and  $n^{|V(H)|}$  factors in (1.1) ensure a nontrivial large n limit. In this paper, we always take  $H_1$  to be an edge and assume  $\beta_2, \ldots, \beta_k$  are positive ( $\beta_1$  can be negative), although our formal approximation also applies to negative  $\beta$ 's. Note that if k = 1, then (1.1) is the Erdös–Rényi model G(n, p) where every edge is present with probability  $p = e^{2\beta_1}(1 + e^{2\beta_1})^{-1}$ , independent of each other. If  $k \geq 2$ , (1.1) "encourages" the presence of the corresponding subgraphs. For example, if k = 2 and  $H_2$  is a rectangle, then (1.1) becomes

$$p_{\beta_1,\beta_2}(G) = \frac{1}{Z(\beta_1,\beta_2)} \exp\left\{2\beta_1 E + \frac{8\beta_2}{n^2} \Box\right\},$$
 (1.2)

where E and  $\square$  denote the number of edges and rectangles in the graph G, respectively, and the constants 2 and 8 are the number of automorphisms of an edge and a rectangle, respectively.

Because of the non-linear nature of (1.1), ERGMs are notoriously more difficult to analyze than classical exponential families of distributions. The ground-breaking works by Bhamidi et al. (2011) and Chatterjee and Diaconis (2013) reveal the following fact: In a certain parameter region (called the *subcritical* region), (1.1) behaves like an Erdös–Rényi random graph G(n,p) in a suitable sense. Here,  $0 is determined by the parameters <math>\beta$  in (1.1). In fact, the subcritical region can fill the whole space except for a set of parameters with Lebesgue measure zero; see, for example, the two-star ERGM studied in Mukherjee and Xu (2023). Moreover, ERGMs with different parameters can lead to the same p; see Chatterjee and Diaconis (2013, Fig. 2) for the case of the triangle ERGM.

## 1.2 Second-order expansion

Because ERGMs may be indistinguishable from G(n, p) to a first-order approximation, we consider the second-order approximation in this paper. For this purpose, we first explain a new way of obtaining p in the first-order approximation using Hoeffding decomposition (Hoeffding (1948)). Taking (1.2) as an example, we rewrite it as

$$p_{\beta_1,\beta_2}(G) = \frac{1}{Z(\beta_1,\beta_2)} \exp\left\{ \left[ 2\beta_1 + 8\beta_2 p^3 \right] E + \frac{8\beta_2}{n^2} \left[ \Box - n^2 p^3 E \right] \right\}. \tag{1.3}$$

Suppose p satisfies

$$2\Phi(p) := 2\beta_1 + 8\beta_2 p^3 = \log \frac{p}{1-p}.$$
(1.4)

Then, up to a constant factor independent of G, the probability of G is proportional to

$$p_{\beta_1,\beta_2}(G) \propto \exp\left\{\frac{8\beta_2}{n^2} \left[\Box - n^2 p^3 E\right]\right\} p^E (1-p)^{n-E}.$$
 (1.5)

What appears inside the brackets [...] in (1.5) is the number of rectangles  $\square$  subtracted by its approximate leading term in the Hoeffding decomposition<sup>2</sup> under the Erdös–Rényi model

 $<sup>^2</sup>$ We remark that the term  $n^2p^3E$  should be  $(n-2)(n-3)p^3E$  in the Hoeffding decomposition. However, the difference is negligible for our purpose. This remark applies to all the approximate Hoeffding decompositions considered in this paper.

G(n,p). Therefore, in a suitable sense (cf. (1.11)), we can hope the ERGM in (1.2) to be close to G(n,p) whose probability distribution is

$$p_1(G) := p_{1,\beta_1,\beta_2}(G) \propto p^E (1-p)^{n-E}$$
(1.6)

for p satisfying (1.4). The equation (1.4) is exactly what appears in Bhamidi et al. (2011) and Chatterjee and Diaconis (2013), although they used different methods to derive it, namely, Glauber dynamics (Glauber (1963)) and non-linear large deviations (Chatterjee and Varadhan (2011)), respectively. The subcritical region in this case contains parameters  $\beta_1$ ,  $\beta_2$  such that there is a unique solution  $p \in (0,1)$  to (1.4) and  $\varphi'(p) < 1$ , where

$$\varphi(a) := \frac{e^{2\Phi(a)}}{e^{2\Phi(a)} + 1}, \quad \Phi(a) := \beta_1 + 4\beta_2 a^3. \tag{1.7}$$

Next, we use the above idea to formally derive a second-order approximation for (1.2). See Section 2 for the derivation for the general case. Taking into account the next-order term in the approximate Hoeffding decomposition, we have

$$p_{\beta_1,\beta_2}(G) \propto \exp\left\{\frac{8\beta_2}{n^2} \left[\Box - np^2 \left(V - 2npE\right) - n^2 p^3 E\right]\right\} \times \exp\left\{\frac{8\beta_2 p^2}{n} \left[V - 2npE\right]\right\} p^E (1-p)^{n-E},$$

$$(1.8)$$

where V denotes the number of two-stars (subgraphs with three vertices and two edges connecting them) in G. Because two-term decomposition further reduces variance, it is natural to use the two-star ERGM

$$p_2(G) := p_{2,\beta_1,\beta_2}(G) \propto \exp\left\{\frac{8\beta_2 p^2}{n} \left[V - 2npE\right]\right\} p^E (1-p)^{n-E}$$
 (1.9)

as a second-order approximation to (1.2). Two-star ERGMs are highly tractable; see Mukherjee and Xu (2023) and the earlier work by Park and Newman (2004).

Note that Hoeffding decomposition is done under the Erdös–Rényi model; it was initially unclear whether the remaining term of the decomposition in (1.8) still has a smaller-order variance under the two-star ERGM (1.9). Fortunately, we can justify (1.9) indeed achieves second-order accuracy as explained next. See more details in Section 3.

### 1.3 Justification and discussion

We identify a simple graph on n labeled vertices  $\{1, \ldots, n\}$  with an element  $x = (x_{ij})_{1 \leqslant i < j \leqslant n} \in \{0, 1\}^{\mathcal{I}}$ , where  $\mathcal{I} := \{(i, j) : 1 \leqslant i < j \leqslant n\}$  and  $x_{ij} = 1$  if and only if there is an edge between vertices i and j. In this way, the ERGM (1.2) induces a random element  $X \in \{0, 1\}^{\mathcal{I}}$ . Similarly, (1.6) and (1.9) induce a random element Z (first-order approximation) and Y (second-order approximation), respectively, in  $\{0, 1\}^{\mathcal{I}}$ . Let  $h : \{0, 1\}^{\mathcal{I}} \to \mathbb{R}$  be a test function. For  $x \in \{0, 1\}^{\mathcal{I}}$  and  $s \in \mathcal{I}$ , define  $x^{(s,1)}$  to have 1 in the sth coordinate and otherwise the same as x, and define  $x^{(s,0)}$  similarly except there is a 0 in the sth coordinate. Define

$$\Delta_s h(x) := h(x^{(s,1)}) - h(x^{(s,0)}), \quad \|\Delta h\| := \sup_{x \in \{0,1\}^{\mathcal{I}}, s \in \mathcal{I}} |\Delta_s h(x)|. \tag{1.10}$$

Applying the general result by Reinert and Ross (2019, Theorem 1.13) to (1.2) in the sub-critical region, we have

$$|\mathbb{E}h(X) - \mathbb{E}h(Z)| \leqslant C||\Delta h||n^{3/2},\tag{1.11}$$

where the constant C depends only on  $\beta_1, \beta_2$ . The upper bound in (1.11) is, in fact, sharp as will be shown in Proposition 3.1. Applying the general result in Section 3, we have, in a subset of the subcritical region,

$$|\mathbb{E}h(X) - \mathbb{E}h(Y)| \leqslant C||\Delta h||n. \tag{1.12}$$

Comparing (1.11) with (1.12), we see that the second-order approximation (1.9) indeed achieves second-order accuracy.

The proof of (1.12) relies on recent advances in understanding ERGMs. According to Mukherjee and Xu (2023), it appears that the two-star ERGM only changes marginal edge probabilities from p to p+O(1/n) (cf. (3.11)). The  $1/\sqrt{n}$  order effect is moving edges to form two-stars (cf. the proof of Proposition 3.1). Therefore, the Hoeffding decomposition under G(n,p) is still suitable for our purpose. We then extend the method of Reinert and Ross (2019) and Bresler and Nagaraj (2019) to second-order approximations. While the firstorder bound needed in Reinert and Ross (2019) is easier to obtain using the independence structure of G(n,p), our second-order bound requires working under  $p_2$  such as in (1.9). We make use of the results of Ganguly and Nam (2021) and Sambale and Sinulis (2020) to obtain the desired bound. Due to our dependence on the works of Mukherjee and Xu (2023) and Sambale and Sinulis (2020), we restrict ourselves to triangle-free ERGMs and a subset of the subcritical region in Theorem 3.1 in Section 3. While the extension to the triangle case may be merely a technical problem, we feel that the naive extension of the above expansion to third or higher-order approximations may not be valid. This is because changing marginal edge probabilities by an amount of order 1/n already incurs too much error for the third-order expansion.

As a related problem, according to Bresler and Nagaraj (2018, Theorem 15), there is a statistical test based on one sample of a random graph to distinguish ERGM and the Erdös–Rényi model. Based on a naive consideration comparing variances, it seems that no test can distinguish ERGMs with their second-order approximations with high probability using a constant number of i.i.d. samples. We leave a more careful study in this direction to future work.

While this work may have potential applications in network analysis, we do not pursue them within the scope of this paper. For instance, in the field of information dissemination, owing to distinct purposes and environmental conditions, we may aim to encourage specific shapes and scales within the fundamental grid units of information dissemination. The results in this paper suggest that forming two-stars and triangles could be the initial step in achieving these goals.

### 1.4 Organization

The rest of the paper is organized as follows. In Section 2, we derive the formal second-order approximation to general ERGMs in the subcritical region. In Section 3, we rigorously justify the approximation in a subset of the subcritical region for triangle-free ERGMs.

# 2 Formal expansion

In this section, we derive the formal second-order approximation to general ERGMs in the subcritical region.

Consider the ERGM (1.1) in the space  $\mathcal{G}_n$  of all simple graphs on n labeled vertices. Suppose  $H_1$  is an edge,  $\beta_2, \ldots, \beta_k$  are all positive parameters. Define (cf. (1.7))

$$\Phi(a) := \sum_{i=1}^{k} \beta_i |E(H_i)| a^{|E(H_i)|-1}, \quad \varphi(a) := \frac{e^{2\Phi(a)}}{e^{2\Phi(a)} + 1}, \tag{2.1}$$

where  $|E(H_i)|$  is the number of edges in the subgraph  $H_l$ . The so-called subcritical region (cf. Bhamidi et al. (2011) and Chatterjee and Diaconis (2013)) contains all the parameters  $\beta = (\beta_1, \ldots, \beta_k)$  such that there is a unique solution  $p := p(\beta)$  to the equation  $\varphi(a) = a$  in (0,1) and  $\varphi'(p) < 1$ . We always use p to denote the unique solution in the rest of the paper. It can be verified that p satisfies

$$2\Phi(p) = \log(\frac{p}{1-p}). \tag{2.2}$$

To formally derive the second-order approximation to the ERGM (1.1), we will apply the Hoeffding decomposition to subgraph counting functions of the Erdös–Rényi graph G(n, p). We refer to Janson and Nowicki (1991, Example 2 (cont.) and Lemma 1 on page 344) for details. Let  $\{Z_{ij}: 1 \leq i < j \leq n\}$  be i.i.d. Bernoulli(p) random variables representing the presence of  $N := \binom{n}{2}$  edges. Define  $Z_{ji} := Z_{ij}$  for  $1 \leq i < j \leq n$ . For m distinct indices  $s_1, \ldots, s_m \in \mathcal{I} := \{(i,j): 1 \leq i < j \leq n\}$ , we have the following orthogonal decomposition (sum of uncorrelated mean-zero terms)

$$Z_{s_1} \dots Z_{s_m} - \mathbb{E}Z_{s_1} \dots Z_{s_m} = \sum_{l=1}^m p^{m-l} \sum_{1 \le i_1 < \dots i_l \le m} (Z_{s_{i_1}} - p) \dots (Z_{s_{i_l}} - p).$$
 (2.3)

In fact, the equation can be checked by the binomial theorem and the uncorrelatedness of summands follows from the independence of the Z's.

Let E, V and  $\triangle$  denote the number of edges, two-stars and triangles, respectively, in the graph G. If G is random, then E, V and  $\triangle$  are regarded as random variables. We subtract the leading terms in their (approximate) Hoeffding decomposition under G(n, p) and define

$$\widetilde{E} := E - \mathbb{E}_p E = \sum_{1 \le i < j \le n} (Z_{ij} - p), \tag{2.4}$$

$$\widetilde{V} := V - \mathbb{E}_p V - 2np\widetilde{E} \approx \sum_{1 \leq i < j < k \leq n} \left[ (Z_{ij} - p)(Z_{ik} - p) + (Z_{ji} - p)(Z_{jk} - p) + (Z_{ki} - p)(Z_{kj} - p) \right],$$

(2.5)

$$\widetilde{\triangle} := \triangle - \mathbb{E}_p \triangle - p\widetilde{V} - np^2 \widetilde{E} \approx \sum_{1 \le i < j < k \le n} (Z_{ij} - p)(Z_{ik} - p)(Z_{jk} - p), \tag{2.6}$$

where  $\mathbb{E}_p$  denotes the expectation with respect to the Erdös-Rényi random graph G(n, p). Since we are interested in large n behavior, for simplicity of notation, we again used an approximate Hoeffding decomposition by neglecting some smaller order terms (namely  $n-l \approx$ 

n for bounded constants l). From (2.4)–(2.6), the number of homomorphisms of an edge to G is

$$2E = \mathbb{E}_p(2E) + 2\widetilde{E};$$

the number of homomorphisms of a two-star to G is

$$2V = \mathbb{E}_p(2V) + 2\widetilde{V} + 4np\widetilde{E};$$

the number of homomorphisms of a triangle to G is

$$6\triangle = \mathbb{E}_p(6\triangle) + 6\widetilde{\triangle} + 6p\widetilde{V} + 6np^2\widetilde{E}.$$

For a general subgraph  $H_i$  with  $v_i \ge 2$  vertices,  $e_i \ge 1$  edges,  $s_i \ge 0$  two-stars and  $t_i \ge 0$  triangles, the leading terms (involving  $\widetilde{\Delta}, \widetilde{V}, \widetilde{E}$ ) in the approximate Hoeffding decomposition of |hom(H, G)| are

$$\mathbb{E}_{p}|\text{hom}(H_{i},G)| + n^{v_{i}-3}6t_{i}p^{e_{i}-3}\widetilde{\triangle} + n^{v_{i}-3}2s_{i}p^{e_{i}-2}\widetilde{V} + n^{v_{i}-2}2e_{i}p^{e_{i}-1}\widetilde{E}. \tag{2.7}$$

In fact, from (2.3), it can be checked by combinatorics that the full Hoeffding decomposition for  $|\text{hom}(H_i, G)|$ , where  $H_i$  is a graph with  $v_i$  vertices and  $e_i$  edges, is

$$|\hom(H_i, G)| = \mathbb{E}_p|\hom(H_i, G)| + \sum_{G_{ij} \subset H_i} (n - v_{ij}) \dots (n - v_i + 1) p^{e_i - e_{ij}} |\hom(G_{ij}, H_i)| \# (G_{ij}, G),$$
(2.8)

where the sum is over all distinct, nonempty, simple subgraphs  $G_{ij}$  of  $H_i$  without isolated vertices,  $v_{ij}$  and  $e_{ij}$  are the number of vertices and edges of  $G_{ij}$ , respectively,  $\#(G_{ij}, G)$  denotes the number of (not necessarily induced) copies of  $G_{ij}$  in G except that each edge indicator  $Z_s$ ,  $s \in \mathcal{I}$ , is replaced by  $(Z_s - p)$  (cf. the right-hand sides of (2.4)–(2.6) for the cases of edges, two-stars and triangles).

From (2.7), the ERGM (1.1) can be rewritten as

$$p_{\beta}(G) = \frac{1}{Z(\beta)} \exp\left\{n^{2} \sum_{i=1}^{k} \beta_{i} t(H_{i}, G)\right\}$$

$$= \frac{1}{Z(\beta)} \exp\left\{Const. + Remain.\right\}$$

$$\times \exp\left\{\left(\sum_{i=1}^{k} \beta_{i} t_{i} p^{e_{i}-3}\right) \frac{6\widetilde{\triangle}}{n} + \left(\sum_{i=1}^{k} \beta_{i} s_{i} p^{e_{i}-2}\right) \frac{2\widetilde{V}}{n}\right\} \exp\left\{\left(\sum_{i=1}^{k} \beta_{i} e_{i} p^{e_{i}-1}\right) 2\widetilde{E}\right\},$$

$$(2.9)$$

where Const. is a constant not depending on G, Remain, are the remaining terms after subtracting the three leading terms in the approximate Hoeffding decomposition. Because p satisfies (2.2), the last factor in (2.9) is proportional to  $p^E(1-p)^{n-E}$ . Therefore, it is not hard to believe that G(n, p) whose probability distribution is

$$p_1(G) := p_{1,\beta}(G) \propto p^E (1-p)^{n-E}$$
(2.10)

for p satisfying (2.2) provides a first-order approximation to the ERGM (1.1). It is known that the next-order terms in the Hoeffding decomposition are represented by subgraphs with

three vertices; this can be checked by direct computation; see also Janson and Nowicki (1991, Lemma 2) for a more general result allowing additional random variables to be associated with vertices. Therefore, based on (2.9), it is natural to propose the following second-order approximation to ERGM.

**Definition 2.1.** For the ERGM (1.1) in the subcritical region and p satisfying (2.2), the second-order approximation is given by a random graph whose probability distribution is

$$p_2(G) := p_{2,\beta}(G) \propto \exp\Big\{\Big(\sum_{i=1}^k \beta_i t_i p^{e_i - 3}\Big) \frac{6\widetilde{\triangle}}{n} + \Big(\sum_{i=1}^k \beta_i s_i p^{e_i - 2}\Big) \frac{2\widetilde{V}}{n}\Big\} p^E (1 - p)^{n - E}$$
 (2.11)

for  $G \in \mathcal{G}_n$ .

# 3 Justification of the formal expansion

Recall that we identify a simple graph on n labeled vertices  $\{1, \ldots, n\}$  with an element  $x = (x_{ij})_{1 \leq i < j \leq n} \in \{0, 1\}^{\mathcal{I}}$ , where  $\mathcal{I} := \mathcal{I}_n := \{(i, j) : 1 \leq i < j \leq n\}$  and  $x_{ij} = 1$  if and only if there is an edge between vertices i and j. In this way, the ERGM (1.1) induces a random element  $X \in \{0, 1\}^{\mathcal{I}}$ . Similarly, (2.10) and (2.11) induces a random element Z (first-order approximation) and Y (second-order approximation), respectively, in  $\{0, 1\}^{\mathcal{I}}$ . Let  $h: \{0, 1\}^{\mathcal{I}} \to \mathbb{R}$  be a test function. Recall the definition of  $\|\Delta h\|$  in (1.10).

Using Stein's method, Reinert and Ross (2019) justified the first-order approximation of the ERGM (1.1) by the Erdös–Rényi graph  $p_1$  in (2.10) in the subcritical region by proving

$$|\mathbb{E}h(X) - \mathbb{E}h(Z)| \leqslant C||\Delta h||n^{3/2},\tag{3.1}$$

where  $C := C(\beta, H)$  is a constant depending only on the parameters  $\beta_1, \ldots, \beta_k$  and subgraphs  $H_1, \ldots, H_k$  in the definition of the ERGM. The bound (3.1) turns out to be sharp as shown in the next proposition.

**Proposition 3.1.** When X is the rectangle ERGM (1.2) in the subcritical region and Z is the Erdös-Rényi graph G(n,p) with p satisfying (1.4), there exist some parameter values  $\beta_1, \beta_2$  satisfying  $\Phi'(1) < 2$ , where  $\Phi(a)$  was defined in (1.7), and a sequence of functions  $h_n : \{0,1\}^{\mathcal{I}_n} \to \mathbb{R}$  such that

$$\liminf_{n\to\infty} \frac{|\mathbb{E}h_n(X) - \mathbb{E}h_n(Z)|}{\|\Delta h_n\|n^{3/2}} > 0.$$

We prove Proposition 3.1 at the end of this section. The next theorem is our main result, which shows that the proposed second-order approximation in Definition 2.1 indeed achieves second-order accuracy for triangle-free ERGMs in a subset of the subcritical region (where  $\beta_2, \ldots, \beta_k$  are sufficiently small positive constants).

**Theorem 3.1.** For the ERGM (1.1), assume  $\Phi'(1) < 2$  (known as the Dobrushin's uniqueness region), where  $\Phi(a)$  was defined in (2.1). Assume in addition that  $s_i = 0$  for all  $2 \le i \le k$  (all the  $H_i$ 's are triangle-free). Then, we have

$$|\mathbb{E}h(X) - \mathbb{E}h(Y)| \leqslant C||\Delta h||n, \tag{3.2}$$

where  $C := C(\beta, H)$  is a constant depending only on the parameters  $\beta_1, \ldots, \beta_k$  and the subgraphs  $H_1, \ldots, H_k$  in the definition of the ERGM.

Proof of Theorem 3.1. In this proof, let  $C := C(\beta, H)$  denote positive constants depending only on the parameters  $\beta_1, \ldots, \beta_k$  and the subgraphs  $H_1, \ldots, H_k$  in the definition of the ERGM and may differ from line to line.

Recall from (2.1) that

$$\Phi(a) = \sum_{i=1}^{k} \beta_i e_i a^{e_i - 1}, \quad e_i = |E(H_i)|.$$

We have

$$\Phi'(a) = \sum_{i=2}^{k} \beta_i e_i(e_i - 1) a^{e_i - 2}, \quad \Phi'(1) = \sum_{i=2}^{k} \beta_i e_i(e_i - 1).$$

Let  $\Phi_2(a)$  be the corresponding function to the second-order approximation (2.11) with  $t_i = 0$  for all i because of the triangle-free condition. Then,

$$\Phi_2'(1) = \sum_{i=2}^k \beta_i 2s_i p^{e_i - 2}.$$

From  $e_i(e_i - 1) \ge 2s_i$  and the condition that  $\Phi'(1) < 2$ , we have

$$\Phi_2'(1) < 2. \tag{3.3}$$

It can be checked that the condition  $\Phi'(1) < 2$  is stronger than the subcritical condition; see Ganguly and Nam (2021, Remark 2.3). Therefore, both the ERGM (1.1) and (2.11) are in the subcritical region under the condition of Theorem 3.1.

#### Step 1. The general bound.

Recall the one-to-one correspondence between a graph G and its edge representation vector x. Recall also that X and Y are the edge representation of the random graphs (1.1) and (2.11), respectively. Straightforward modification of Reinert and Ross (2019, Theorems 1.1) and using their Lemma 3.2 for their value of  $\rho$  in Dobrushin's uniqueness region lead to the following bound

$$|\mathbb{E}h(X) - \mathbb{E}h(Y)| \leq C||\Delta h|| \sum_{s \in \mathcal{I}} \mathbb{E}|\Delta_s L(Y) - \Delta_s \widetilde{L}(Y)|,$$

where the sum is over  $\mathcal{I} = \{(i,j) : 1 \leq i < j \leq n\}$  and for each s, the differencing operator  $\Delta_s$  was defined in (1.10) and (recall (1.1), the last line of (2.9), (2.11) and  $t_i = 0$  for all i in the triangle-free case)

$$L(x) = n^2 \sum_{i=1}^{k} \beta_i \frac{|\text{hom}(H_i, G)|}{n^{v_i}},$$

$$\widetilde{L}(x) = \left(\sum_{i=1}^{k} \beta_i s_i p^{e_i - 2}\right) \frac{2\widetilde{V}}{n} + \left(\sum_{i=1}^{k} \beta_i e_i p^{e_i - 1}\right) 2\widetilde{E}.$$

To prove Theorem 3.1, it suffices to show, for any edge s connecting two vertices a and b,

$$\mathbb{E}|\Delta_s(L(Y) - \widetilde{L}(Y))| \leqslant \frac{C}{n}.$$
(3.4)

In the following proof, we fix the edge s.

### Step 2. Hoeffding decomposition under $p_2$ .

From their expressions above, we have

$$L(x) - \widetilde{L}(x) = \sum_{i=1}^{k} \frac{\beta_i}{n} \left\{ \frac{|\text{hom}(H_i, G)|}{n^{v_i - 3}} - s_i p^{e_i - 2} 2\widetilde{V} - n e_i p^{e_i - 1} 2\widetilde{E} \right\} =: \sum_{i=1}^{k} \frac{\beta_i}{n} R_i(x).$$

To prove (3.4), it suffices to prove, for any  $1 \le i \le k$ ,

$$\mathbb{E}|\Delta_s R_i(Y)| \leqslant C. \tag{3.5}$$

Case 1. If  $H_i$  is an edge, then  $R_i(x) = Const.$ , where Const. denote constants which are independent of x and hence will vanish after taking the differencing operator  $\Delta_s$ .

Case 2. If  $H_i$  is a two-star, then, from (2.5),

$$R_i(x) = 2V - 2\widetilde{V} - 4np\widetilde{E} = Const.$$

Case 3. Under the triangle-tree condition, the only remaining possibilities for  $H_i$  are simple graphs with  $v_i \geqslant 4$ . For Case 3, from (2.8), we have

$$R_{i}(x) = Const. + \sum_{\substack{G_{ij} \subset H_{i}: \\ v_{ij} \geqslant 4}} \frac{(n - v_{ij}) \cdots (n - v_{i} + 1) p^{e_{i} - e_{ij}} |\text{hom}(G_{ij}, H_{i})|}{n^{v_{i} - 3}} \# \widetilde{(G_{ij}, G)}$$

$$+ \left[ \frac{(n - 3) \cdots (n - v_{i} + 1)}{n^{v_{i} - 3}} - 1 \right] p^{e_{i} - 2} (2s_{i}) \widetilde{V}$$

$$+ \left[ \frac{(n - 2) \cdots (n - v_{i} + 1)}{n^{v_{i} - 3}} - n \right] p^{e_{i} - 1} (2e_{i}) \widetilde{E}$$

$$=: Const. + R_{i1}(x) + R_{i2}(x) + R_{i3}(x),$$

where, as in (2.8), the sum is over all distinct subgraphs  $G_{ij}$  of  $H_i$  without isolated vertices,  $v_{ij}$  and  $e_{ij}$  are the number of vertices and edges of  $G_{ij}$ , respectively. Note that under the triangle-free condition of the theorem, the terms  $R_{i2}(x)$  and  $R_{i3}(x)$  above correspond to the cases  $v_{ij} = 3$  and  $v_{ij} = 2$ , respectively.

Because  $\Delta_s \widetilde{E} = 1$  and  $\left| \frac{(n-2)\cdots(n-v_i+1)}{n^{v_i-3}} - n \right| \leq C$ , we have  $\mathbb{E}|R_{i3}(Y)| \leq C$ . Because  $|\Delta_s \widetilde{V}| \leq Cn$  and  $\left| \frac{(n-3)\cdots(n-v_i+1)}{n^{v_i-3}} - 1 \right| \leq C/n$ , we have  $\mathbb{E}|R_{i2}(Y)| \leq C$ . In fact,  $\mathbb{E}|R_{i2}(Y)|=o(1)$ , but we don't need it here.

To prove (3.5), it remains to prove

$$\mathbb{E}|\Delta_s R_{i1}(Y)| \leqslant C. \tag{3.6}$$

Therefore, we fix  $G_{ij} \subset H_i$  with  $v_{ij} \geqslant 4$ . From the definitions of  $\#(G_{ij}, G)$  below (2.8) and  $\Delta_s$ , we have

$$\Delta_s \# \widetilde{(G_{ij}, G)} = \# \widetilde{(G_{ij}^{(s)}, G)}, \tag{3.7}$$

which is the number of copies of  $G_{ij}$  in G containing the edge s, except that each remaining edge indicator  $Z_r$  is replaced by  $(Z_r - p)$ .

As in Sambale and Sinulis (2020), as building blocks of the Hoeffding decomposition under  $p_2$ , we define

$$f_{d,A} = \sum_{I \in \mathcal{I}^d} A_I \sum_{P \in \mathcal{P}(I)} (-1)^{M(P)} \left\{ \prod_{\substack{J \in P \\ |J| = 1}} (Y_J - \mathbb{E}Y_J) \right\} \prod_{\substack{J \in P \\ |J| > 1}} \left\{ \mathbb{E} \prod_{l \in J} (Y_l - \mathbb{E}Y_l) \right\}, \tag{3.8}$$

where  $d \ge 1$  is an integer,  $\mathcal{I} = \{(i, j) : 1 \le i < j \le n\}$ , A is a d-tensor with vanishing diagonal (i.e.,  $A_I > 0$  only if all the d elements in I are distinct),

$$\mathcal{P}(I) = \{ S \subset 2^I : S \text{ is a partition of } I \},$$

M(P) is the number of subsets with more than one element in the partition P and for a singleton set  $J = \{l\}$ ,  $Y_J := Y_l$ . For our purpose, it is enough to consider those values of d bounded by the maximum number of edges of subgraphs  $H_1, \ldots, H_k$  in the following.

From Sambale and Sinulis (2020, Theorem 3.7)<sup>3</sup>, under the condition (3.3),

$$\mathbb{E}(f_{d,A})^2 \leqslant C \|A\|_2^2,\tag{3.9}$$

where  $||A||_2$  is the Euclidean norm of the tensor A when viewed as a vector.

Next, we express (3.7) in terms of (3.8). We need the following facts. From Ganguly and Nam (2021, Eq.(34)), for any fixed  $m \ge 1$  and distinct edges  $l_1, \ldots, l_m$ , we have, in the subcritical region,

$$|\mathbb{E}(Y_{l_1}|Y_{l_2},\dots,Y_{l_m}) - \mathbb{E}Y_{l_1}| \leqslant \frac{C}{n}.$$
(3.10)

From Mukherjee and Xu (2023, Lemma 3.3(d)), we have, in the subcritical region,

$$|\mathbb{E}\bar{\phi} - t| \leqslant \sqrt{\mathbb{E}(\bar{\phi} - t)^2} \leqslant C/n,$$

where t = 2p - 1 + O(1/n) (see their Lemma 1.2(a)) and  $\mathbb{E}\bar{\phi} = \mathbb{E}(\phi_1 + \dots + \phi_n)/n = 2\mathbb{E}Y_l - 1$  for any  $l \in \mathcal{I}$  (using their Eqs. (8) and (10), the symmetry and a connection between the total degree and the total number of edges of a graph). Therefore, for any  $l \in \mathcal{I}$ ,

$$|\mathbb{E}Y_l - p| \leqslant \frac{C}{n}.\tag{3.11}$$

Decomposing according to the union L of singleton sets in the partition P in (3.8), we write

$$f_{d,A} = \sum_{I \in \mathcal{I}^d} A_I \sum_{L \subset I} C_L \left\{ \prod_{l \in L} (Y_l - \mathbb{E}Y_l) \right\},$$

where  $C_L = 1$  if L = I and, from (3.10),  $C_L = O(1/n)$  if  $L \neq I$  (in the case that L is the empty set, the product  $\prod_{l \in L}$  is understood to be equal to 1). From (3.11), we further rewrite  $f_{d,A}$  as

$$f_{d,A} = \sum_{I \in \mathcal{I}^d} A_I \sum_{L \subset I} C_L' \left\{ \prod_{l \in L} (Y_l - p) \right\},$$
 (3.12)

<sup>&</sup>lt;sup>3</sup>They assumed that the subgraphs  $H_1, \ldots, H_k$  in the definition of ERGM are all connected at the beginning of their paper. However, as far as we checked their proof, this requirement is not needed.

where  $C'_L = 1$  if L = I and  $C'_L = O(1/n)$  if  $L \neq I$ . From (3.7), for  $d = e_{ij} - 1$  (total number of edges of  $G_{ij}$  subtracted by one fixed edge s) and some  $\{A_I : I \in \mathcal{I}^d\}$  with uniformly bounded entries,

$$\Delta_s \# \widetilde{(G_{ij}, G)} = \# \widetilde{(G_{ij}^{(s)}, G)}$$

$$= \sum_{I \in \mathcal{I}^d} A_I \prod_{l \in I} (Y_l - p)$$

$$= f_{d,A} - \sum_{I \in \mathcal{I}^d} A_I \sum_{\substack{L \subset I: \\ L \neq I}} C'_L \left\{ \prod_{l \in L} (Y_l - p) \right\}. \tag{3.13}$$

#### Step 3. Final bound.

Using the the expression (3.13), we now turn to proving (3.6) for each  $G_{ij} \subset H_i$  with  $v_{ij} \ge 4$ .

First of all,

$$\mathbb{E}\left|\frac{(n-v_{ij})\cdots(v-v_{i}+1)p^{e_{i}-e_{ij}}|\text{hom}(G_{ij},H_{i})|}{n^{v_{i}-3}}f_{d,A}\right| \leqslant \frac{Cn^{v_{i}-v_{ij}}}{n^{v_{i}-3}}\sqrt{n^{v_{ij}-2}} \leqslant C, \quad (3.14)$$

where the first inequality is from (3.9) and the fact that the number of non-zero entries of  $\{A_I : I \in \mathcal{I}^d\}$  involved is of the order  $O(n^{v_{ij}-2})$  (the number of choices of the remaining  $v_{ij}-2$  vertices of  $G_{ij}$  after fixing the two vertices a, b connected by s) and the second inequality is because  $v_{ij} \geqslant 4$ .

Next, for the remaining terms in (3.13), repeated using the derivation for (3.13) and expressing  $\prod_{l\in L}(Y_l-p)$  using  $f_{d,A}$  with  $d\leqslant |L|$ , we obtain the upper bound (recall  $C'_L=O(1/n)$ )

$$\mathbb{E}\left|\frac{(n-v_{ij})\cdots(v-v_{i}+1)p^{e_{i}-e_{ij}}|\text{hom}(G_{ij},H_{i})|}{n^{v_{i}-3}}\sum_{I\in\mathcal{I}^{d}}A_{I}\sum_{\substack{L\subset I:\\L\neq I}}C'_{L}\left\{\prod_{l\in L}(Y_{l}-p)\right\}\right|$$

$$\leqslant \frac{C}{n^{v_{ij}-3}}\frac{1}{n}\sum_{0\leqslant v\leqslant v_{ij}-2}n^{v}\sqrt{n^{v_{ij}-2-v}},$$

where the sum is over all possible number v of isolated vertices of  $G_{ij}$  after removing some edges (but keeping at least one edge to map to s),  $n^v$  is the order of the number of choices of these v isolated vertices, and  $n^{v_{ij}-2-v}$  comes from an argument using (3.9) as for the first inequality in (3.14). For each  $0 \le v \le v_{ij} - 2$ , we have

$$\frac{C}{n^{v_{ij}-3}} \frac{1}{n} n^v \sqrt{n^{v_{ij}-2-v}} \leqslant C n^{-v_{ij}/2+v/2+1} \leqslant C.$$

Therefore, (3.6) is proved and the theorem follows.

Proof of Proposition 3.1. We identify a graph G on n vertices with a  $\binom{n}{2}$ -dimensional vector x as at the beginning of this section. Similarly, a random graph corresponds to a random vector in the same way. Let X follow the rectangle ERGM (1.2) with probability

$$p_{\beta_1,\beta_2}(G) \propto \exp\{\frac{8\beta_2}{n^2}\Box + 2\beta_1 E\},$$

Z follow the Erdös–Rényi model G(n, p) with p in (1.4), and Y follow the two-star ERGM (1.9)

$$p_2(G) \propto \exp\{\frac{2\widetilde{\beta}_2}{n}V + 2\widetilde{\beta}_1 E\},$$

where, by a rewriting of (1.9),

$$\widetilde{\beta}_2 = 4\beta_2 p^2, \ \widetilde{\beta}_1 = -8\beta_2 p^3 + \frac{1}{2} \log(\frac{p}{1-p}).$$

For each n, choose the test function  $h_n(x)$  to be

$$h_{n}(x) := \mathbb{E} \sum_{i=1}^{n} \left\{ \left( \frac{2d_{i}(G)}{n-1} - 2p + \frac{W_{i}}{\sqrt{n\widetilde{\beta}_{2}/2}} \right)^{2} 1_{\left\{ \left( \frac{2d_{i}(G)}{n-1} - 2p + \frac{W_{i}}{\sqrt{n\widetilde{\beta}_{2}/2}} \right)^{2} \leq \frac{M}{n} \right\}} + \frac{M}{n} 1_{\left\{ \left( \frac{2d_{i}(G)}{n-1} - 2p + \frac{W_{i}}{\sqrt{n\widetilde{\beta}_{2}/2}} \right)^{2} > \frac{M}{n} \right\}} \right\},$$

$$(3.15)$$

where  $d_i(G)$  is the degree of the vertex i in the graph  $G, W_1, \ldots, W_n$  are i.i.d. standard normal random variables independent of everything else, the expectation is with respect to the W's,  $1_{\{\ldots\}}$  is the indicator function, and M is a sufficiently large constant to be chosen below (for (3.19) and (3.24)). From Theorem 3.1, for  $\beta_1, \beta_2$  satisfying  $\Phi'(1) < 2$ , we have  $|\mathbb{E}h_n(X) - \mathbb{E}h_n(Y)| \leq C||\Delta h_n||n$ . Therefore, to prove the proposition, it suffices to show

$$\liminf_{n \to \infty} \frac{|\mathbb{E}h_n(Y) - \mathbb{E}h_n(Z)|}{\|\Delta h_n\|_{n^{3/2}}} > 0$$
(3.16)

for some parameter values  $\beta_1, \beta_2$  satisfying  $\Phi'(1) < 2$ . Because changing an edge only changes  $d_i(G)$  by 1 for two out of n vertices and because of the trunction in the definition of  $h_n(x)$ , we have  $\|\Delta h_n\| \leq C/n^{3/2}$ . Therefore, to prove (3.16), it suffices to show

$$\liminf_{n \to \infty} |\mathbb{E}h_n(Y) - \mathbb{E}h_n(Z)| > 0$$
(3.17)

for some parameter values  $\beta_1, \beta_2$  satisfying  $\Phi'(1) < 2$ .

We first consider the simpler term  $\mathbb{E}h_n(Z)$ . Because Z follows G(n,p), the degree  $d_i(Z)$  of each vertex i follows the binomial distribution Bin(n-1,p). Recall  $W_i \sim N(0,1)$  are independent normal variables independent of everything else. We have

$$\limsup_{n \to \infty} \mathbb{E}h_n(Z) \leqslant \limsup_{n \to \infty} \mathbb{E}\sum_{i=1}^n \left(\frac{2d_i(Z)}{n-1} - 2p + \frac{W_i}{\sqrt{n\widetilde{\beta}_2/2}}\right)^2 = 4p(1-p) + \frac{2}{\widetilde{\beta}_2}.$$
 (3.18)

Next, we consider  $\mathbb{E}h_n(Y)$ . In fact, Mukherjee and Xu (2023, Lemma 2.2(a)) implies the following

**Claim.** For sufficiently large M,  $\liminf_{n\to\infty} \mathbb{E}h_n(Y)$  is sufficiently close to  $1/\tilde{a}_1$ , where

$$\widetilde{a}_1 = \widetilde{\theta} - \widetilde{\theta}^2 (1 - \widetilde{t}^2), \ \widetilde{\theta} = \frac{\widetilde{\beta}_2}{2}, \ \widetilde{t} = 2p - 1.$$

Take, for example,  $\beta_2 = 0.16$  (which satisfies  $\Phi'(1) < 2$ ),  $\beta_1 = -0.08$ , p = 0.5, then (1.4) is satisfied. We can compute  $1/\tilde{a}_1 = 13.58696 > 13.5 = 4p(1-p) + 2/\tilde{\beta}_2$ . Then (3.17) follows from (3.18) and the above claim. Therefore, the proposition follows.

We are left to prove the above claim.

**Proof of Claim.** As in Mukherjee and Xu (2023), define random variables

$$\phi_i := \frac{2d_i(Y)}{n-1} - 1 + \frac{W_i}{\sqrt{(n-1)\theta}}, \ 1 \leqslant i \leqslant n, \quad \theta := \frac{\widetilde{\beta}_2(n-1)}{2n}.$$

From Mukherjee and Xu (2023, page 6 in the supplementary material),  $\phi := (\phi_1, \dots, \phi_n)$  has joint probability density function  $f_n(\phi)$  and

$$\frac{(n-1)\lambda_2}{2} \sum_{i=1}^n (\phi_i - t)^2 \leqslant -\log f_n(\phi) \leqslant \frac{(n-1)\lambda_1}{2} \sum_{i=1}^n (\phi_i - t)^2,$$

where  $\lambda_1, \lambda_2$  are two positive constants depending only on  $\beta_1, \beta_2$  and t = 2p - 1 + O(1/n) (see Mukherjee and Xu (2023, Lemma 1.2(a))). From a straightforward modification of the proof of Mukherjee and Xu (2023, Lemma 3.3(b)), we obtain, for a sufficiently large constant M,

$$\mathbb{E}\sum_{i=1}^{n} (\phi_i - t)^2 1_{\{\sum_{i=1}^{n} (\phi_i - t)^2 > M\}} \to 0 \text{ as } n \to \infty.$$
 (3.19)

In fact, for any M > 0,

$$\mathbb{E} \sum_{i=1}^{n} (\phi_{i} - t)^{2} 1_{\{\sum_{i=1}^{n} (\phi_{i} - t)^{2} > M\}}$$

$$= \frac{\int_{\mathbb{R}^{n}} e^{-f_{n}(\phi)} \sum_{i=1}^{n} (\phi_{i} - t)^{2} 1_{\{\sum_{i=1}^{n} (\phi_{i} - t)^{2} > M\}} d\phi}{\int_{\mathbb{R}^{n}} e^{-f_{n}(\phi)} d\phi}$$

$$\leq \frac{\int_{\mathbb{R}^{n}} e^{-\frac{(n-1)\lambda_{2}}{2} \sum_{i=1}^{n} (\phi_{i} - t)^{2} \sum_{i=1}^{n} (\phi_{i} - t)^{2} 1_{\{\sum_{i=1}^{n} (\phi_{i} - t)^{2} > M\}} d\phi}{\int_{\mathbb{R}^{n}} e^{-\frac{(n-1)\lambda_{1}}{2} \sum_{i=1}^{n} (\phi_{i} - t)^{2}} d\phi}$$

$$\leq \left(\frac{\lambda_{1}}{\lambda_{2}}\right)^{\frac{n}{2}} \frac{1}{(n-1)\lambda_{2}} \mathbb{E} \chi_{n}^{2} 1_{\{\chi_{n}^{2} \geqslant (n-1)\lambda_{2}M\}},$$

where  $\chi_n^2$  is a chi-square random variable with n degrees of freedom. From the moment generating function of  $\chi_n^2$  ( $\mathbb{E}e^{t\chi_n^2}=1/(1-2t)^{n/2}$  for 0< t< 1/2) and Markov's inequality, we obtain the desired result (3.19) for sufficiently large M.

By Mukherjee and Xu (2023, Lemma 2.2(a)) (their  $a_1 = \tilde{a}_1 + O(1/n)$ ), we have

$$\sqrt{n}\left[\sum_{i=1}^{n}(\phi_i-\bar{\phi})^2-\frac{1}{\widetilde{a}_1}\right]\to N(0,\frac{1}{2\widetilde{a}_1^2})$$
 in distribution,

where  $\bar{\phi} := (\phi_1 + \dots + \phi_n)/n$ . This implies

$$\sqrt{n} \left[ \sum_{i=1}^{n} (\phi_i - t)^2 + n(t - \bar{\phi})^2 - \frac{1}{\tilde{a}_1} \right] \to N(0, \frac{1}{2\tilde{a}_1^2}) \text{ in distribution.}$$
 (3.20)

By Mukherjee and Xu (2023, Lemma 3.3), we have

$$\mathbb{E}[(\bar{\phi} - t)]^2 \leqslant \frac{C}{n^2}.\tag{3.21}$$

From (3.20) and (3.21), we have

$$\sqrt{n} \left[ \sum_{i=1}^{n} (\phi_i - t)^2 - \frac{1}{\widetilde{a}_1} \right] \to N(0, \frac{1}{2\widetilde{a}_1^2}) \text{ in distribution.}$$
 (3.22)

From (3.19) and (3.22), we have

$$\mathbb{E}\sum_{i=1}^{n}(\phi_{i}-t)^{2}\to\frac{1}{\widetilde{a}_{1}}.$$

By symmetry, we have  $n\mathbb{E}(\phi_i - t)^2 \to 1/\tilde{a}_1$ , which implies (recall t = 2p - 1 + O(1/n))

$$n\mathbb{E}\left(\frac{2d_i(Y)}{n-1} - 2p + \frac{W_i}{\sqrt{n\widetilde{\beta}_2/2}}\right)^2 \to \frac{1}{\widetilde{a}_1}.$$
 (3.23)

Moreover, the tightness of the sequence  $\{n(\phi_i-t)^2\}_{n=1}^{\infty}$  (see Mukherjee and Xu (2023, Lemma 3.3(c))) implies the tightness of the sequence  $\{n(\frac{2d_i(Y)}{n-1}-2p+\frac{W_i}{\sqrt{n\tilde{\beta}_2/2}})^2\}_{n=1}^{\infty}$ . Eq.(3.23), together with the tightness and symmetry, leads to the fact that for any  $\varepsilon>0$ , there exists M>0, such that (recall the definition of  $h_n$  in (3.15))

$$\liminf_{n \to \infty} \mathbb{E}h_n(Y) \geqslant \frac{1}{\tilde{a}_1} - \varepsilon, \tag{3.24}$$

proving the claim.

# Acknowledgements

Ding W. Y. was partially supported by the National Social Science Fund of China 20CTQ005. Fang X. was partially supported by Hong Kong RGC GRF 14305821, 14304822, 14303423 and a CUHK direct grant.

## References

Bhamidi, S., Bresler, G., and Sly, A. (2011). Mixing time of exponential random graphs. *The Annals of Applied Probability*, 21(6):2146–2170.

Bresler, G. and Nagaraj, D. (2018). Optimal single sample tests for structured versus unstructured network data. In *Conference On Learning Theory*, pages 1657–1690. PMLR.

Bresler, G. and Nagaraj, D. (2019). Stein's method for stationary distributions of markov chains and application to Ising models. *The Annals of Applied Probability*, 29(5):3230–3265.

- Chatterjee, S. and Diaconis, P. (2013). Estimating and understanding exponential random graph models. *The Annals of Statistics*, 41(5):2428–2461.
- Chatterjee, S. and Varadhan, S. S. (2011). The large deviation principle for the Erdős-Rényi random graph. *European Journal of Combinatorics*, 32(7):1000–1017.
- Ganguly, S. and Nam, K. (2021). Sub-critical exponential random graphs: concentration of measure and some applications. arXiv preprint arXiv:1909.11080v2, to appear in Transactions of the American Mathematical Society.
- Glauber, R. J. (1963). Time-dependent statistics of the ising model. *Journal of Mathematical Physics*, 4(2):294–307.
- Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *The Annals of Mathematical Statistics*, 19(3):293–325.
- Janson, S. and Nowicki, K. (1991). The asymptotic distributions of generalized U-statistics with applications to random graphs. *Probability Theory and Related Fields*, 90:341–375.
- Mukherjee, S. and Xu, Y. (2023). Statistics of the two star ERGM. Bernoulli, 29(1):24–51.
- Park, J. and Newman, M. E. (2004). Statistical mechanics of networks. *Physical Review E*, 70(6):066117.
- Reinert, G. and Ross, N. (2019). Approximating stationary distributions of fast mixing Glauber dynamics, with applications to exponential random graphs. *The Annals of Applied Probability*, 29(5):3201–3229.
- Sambale, H. and Sinulis, A. (2020). Logarithmic Sobolev inequalities for finite spin systems and applications. *Bernoulli*, 26(3):1863–1890.
- Wasserman, S. and Faust, K. (1994). Social network analysis: Methods and applications. Cambridge University Press.