Journal of Statistical Planning and Inference 219 (2022) 98–119

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

A likelihood-ratio type test for stochastic block models with bounded degrees

Mingao Yuan [a,*], Yang Feng [b,1], Zuofeng Shang [c,2]

[a] Department of Statistics, North Dakota State University, 1340 Administration Ave, Fargo, ND 58102, United States of America
[b] College of Global Public Health, New York University, 665 Broadway, New York, NY 10003, United States of America
[c] Department of Mathematical Sciences, New Jersey Institute of Technology, 323 Dr M.L.K. Jr. Blvd, Newark, NJ 07102, United States of America

ARTICLE INFO

Article history:
Received 12 February 2020
Received in revised form 6 December 2021
Accepted 7 December 2021
Available online 17 December 2021

AMS 2000 subject classifications:
primary 62G10
secondary 05C80

Keywords:
Stochastic block model
Bounded degrees
Hypothesis testing
Likelihood ratio
Contiguity theory

ABSTRACT

A fundamental problem in network data analysis is to test Erdös–Rényi model $\mathcal{G}\left(n, \frac{a+b}{2n}\right)$ versus a bisection stochastic block model $\mathcal{G}\left(n, \frac{a}{n}, \frac{b}{n}\right)$, where $a, b > 0$ are constants that represent the expected degrees of the graphs and n denotes the number of nodes. This problem serves as the foundation of many other problems such as testing-based methods for determining the number of communities (Bickel and Sarkar, 2016; Lei, 2016) and community detection (Montanari and Sen, 2016). Existing work has been focusing on growing-degree regime $a, b \to \infty$ (Bickel and Sarkar, 2016; Lei, 2016; Montanari and Sen, 2016; Banerjee and Ma, 2017; Banerjee, 2018; Gao and Lafferty, 2017a,b) while leaving the bounded-degree regime untreated. In this paper, we propose a likelihood-ratio (LR) type procedure based on regularization to test stochastic block models with bounded degrees. We derive the limit distributions as power Poisson laws under both null and alternative hypotheses, based on which the limit power of the test is carefully analyzed. We also examine a Monte-Carlo method that partly resolves the computational cost issue. The proposed procedures are examined by both simulated and real-world data. The proof depends on a contiguity theory developed by Janson (1995).

1. Introduction

In recent years, stochastic block model (SBM) has attracted increasing attention in statistics and machine learning. It provides the researchers a ground to study many important problems that arise in network data such as community detection or clustering (Amini et al., 2013; Amini and Levina, 2018; Neeman and Netrapalli, 2014; Sarkar and Bickel, 2015; Bickel and Chen, 2009; Zhao et al., 2012), goodness-of-fit of SBMs (Bickel and Sarkar, 2016; Lei, 2016; Montanari and Sen, 2016; Banerjee and Ma, 2017; Banerjee, 2018; Gao and Lafferty, 2017a,b) or various phase transition phenomena (Mossel et al., 2015, 2017; Abbe and Sandon, 2018). See Abbe (2018) for a comprehensive review about recent development in this field. A key assumption in most of the literature is that the expected degree of every node tends to infinity along with the number of nodes n. For instance, in community detection (Bickel and Chen, 2009; Zhao et al., 2012), such a condition is needed for proving weak consistency of the detection methods; to prove strong consistency, the expected

* Corresponding author.
E-mail addresses: mingao.yuan@ndsu.edu (M. Yuan), yang.feng@nyu.edu (Y. Feng), zuofeng.shang@njit.edu (Z. Shang).
1 Supported by NSF CAREER, United States Grant DMS-1554804.
2 Supported by NSF, United States DMS-1764280.

https://doi.org/10.1016/j.jspi.2021.12.005
0378-3758/

degree is further assumed to grow faster than $\log n$. For goodness-of-fit test, the growing-degree condition is needed to derive various asymptotic distributions for the test statistics (Bickel and Sarkar, 2016; Lei, 2016; Banerjee and Ma, 2017; Banerjee, 2018; Gao and Lafferty, 2017a,b).

Many real-world network data sets are highly sparse. For instance, the LinkedIn network, the real-world coauthorship networks, power transmission networks and web link networks all have small average degrees (see Leskovec et al. (2008), Strogatz (2001)). Therefore, it is reasonable to assume bounded degrees in such networks. There is a breakthrough recently made by Mossel et al. (2015, 2017) and Abbe and Sandon (2018) about the possibility of successfully detecting the community structures when the expected degree of SBM is bounded. Specifically, the signal-to-noise ratio (SNR) of the multi-community SBM is used in these work as a phase transition parameter to indicate the possibility of successful detection. Motivated by such a groundbreaking result, it is natural to ask whether one can propose successful testing methods for SBMs with bounded degrees. Progress in this field may help researchers better understand the roles played by the expected degrees of SBMs in hypothesis testing, as well as provide a substantially broader scope of network models in which a successful test is possible.

In this paper, we address this problem in the bisection SBM scenario. We propose a likelihood-ratio (LR) type test statistic to distinguish an Erdös-Rényi model versus a bisection SBM whose expected degrees are finite constants, and investigate its asymptotic properties. In what follows, we describe the models and our contributions more explicitly.

1.1. Models and our contributions

Let us provide a brief review for Erdös-Rényi model and bisection SBM. Throughout the whole paper, assume that $a > b > 0$ are *fixed and known* constants unless otherwise indicated. For $n \in \mathbb{N}$, let $\mathcal{G}\left(n, \frac{a}{n}, \frac{b}{n}\right)$ denote the bisection stochastic block model of random $\pm$-labeled graphs in which each vertex $u \in [n] := \{1, 2, \ldots, n\}$ is assigned, independently and uniformly at random, a label $\sigma_u \in \{\pm\}$, and then each possible edge (u, v) is included with probability a/n if $\sigma_u = \sigma_v$ and with probability b/n if $\sigma_u \neq \sigma_v$. Let $A = [A_{uv}]_{u,v=1}^n \in \{0, 1\}^{n\times n}$ denote the observed symmetric adjacency matrix in which $A_{uu} = 0$ for all $1 \le u \le n$, and for $1 \le u < v \le n$, $A_{uv} = 1$ indicates the inclusion of edge (u, v) and $A_{uv} = 0$ otherwise. Conditional on $\sigma = (\sigma_1, \ldots, \sigma_n)$, the variables A_{uv}, $1 \le u < v \le n$, are assumed to be independent which follow

$$P(A_{uv} = 1|\sigma) = p_{uv}(\sigma) \quad \text{and} \quad P(A_{uv} = 0|\sigma) = q_{uv}(\sigma), \tag{1}$$

where

$$p_{uv}(\sigma) = \begin{cases} \frac{a}{n}, & \sigma_u = \sigma_v \\ \frac{b}{n}, & \sigma_u \neq \sigma_v \end{cases}, q_{uv}(\sigma) = 1 - p_{uv}(\sigma).$$

The Erdös-Rényi model $\mathcal{G}\left(n, \frac{a+b}{2n}\right)$ has the same average degree as $\mathcal{G}\left(n, \frac{a}{n}, \frac{b}{n}\right)$. It is interesting to decide which model an observed graph is generated from. Specifically, we are interested in the following hypothesis testing problem

$$H_0 : A \sim \mathcal{G}\left(n, \frac{a+b}{2n}\right) \quad \text{vs.} \quad H_1 : A \sim \mathcal{G}\left(n, \frac{a}{n}, \frac{b}{n}\right). \tag{2}$$

To be more specific, we want to test whether the nodes on an observed random graph belong to the same community, or they belong to two equal-sized communities.

Let $\kappa = \frac{(a-b)^2}{2(a+b)}$ denote the signal-to-noise ratio (SNR) associated with $\mathcal{G}\left(n, \frac{a}{n}, \frac{b}{n}\right)$. It was conjectured by Decelle, Krzkala, Moore and Zdeborová (Decelle et al., 2011) that successful community detection is possible when $\kappa \ge 1$, and impossible when $\kappa < 1$. This conjecture was recently proved by Mossel, Neeman and Sly (Mossel et al., 2015) through Janson's continuity theory (Janson, 1995). In the meantime, their result indicates that *no test can be successful when* $\kappa < 1$ (see Mossel et al. (2015) and Montanari and Sen (2016)), and so we primarily focus on the high SNR scenario $\kappa \ge 1$. Classic likelihood-ratio (LR) tests for (2) are not valid since the probability measures associated with H_0 and H_1 are asymptotically orthogonal as discovered by Mossel et al. (2015). The result of Mossel et al. (2015) also implies that counting the cycles of length $\log^{1/4} n$ leads to an asymptotically valid test; see their Theorem 4. However, such test is unrealistic since n should be at least e^{81} to make the length at least 3 (the smallest length of a valid cycle). In Section 2, we propose a regularized LR-type test for (2) to address these limitations. Our test does not suffer from the orthogonality issue of LR and is applicable for moderately large n. Our test involves a regularization parameter that can reduce the variability of the classic LR test so that it becomes valid. Based on a contiguity theory for random regular graphs developed by Janson (1995), we derive the asymptotic distributions as power Poisson laws under both H_0 and H_1, which turn out to be infinite products of power Poisson variables (see Section 2.1). Based on power Poisson laws, we rigorously analyze the asymptotic power of our test. In Section 2.2, we show that the test is powerful provided that κ approaches infinity, and the limit power is not sensitive to the choice of regularization parameter. Our test is practically useful in that the parameters a, b can be consistently estimated when $\kappa > 1$, and so the regularization parameter can be empirically selected. Our procedure is based on averaged likelihood-ratios whose computational cost scales exponentially with n. This computational issue is partly resolved in Section 2.3 via Monte Carlo approximations, with the number of experiments suggested to guarantee the success of such approximations. Simulation examples are provided in Section 3.1 to demonstrate the finite sample performance of our methods. In particular, our method achieves desirable size and power, while the methods designed for denser graphs appear to be less powerful.

1.2. Related references

The problem of testing (2) has been recently considered by Bickel and Sarkar (2016), Lei (2016), Montanari and Sen (2016), Banerjee and Ma (2017), Banerjee (2018) and Gao and Lafferty (2017a,b) but only in the growing-degree regime, i.e., $a, b \to \infty$. Specifically, Bickel and Sarkar (2016), Montanari and Sen (2016) and Lei (2016) proposed spectral algorithms; Banerjee and Ma (2017), Banerjee (2018) proposed linear spectral statistics and LR test relating to signed cycles; Gao and Lafferty (2017a,b) proposed algorithms based subgraph counts. In particular, the LR test by Banerjee and Ma (2017) was proposed under low SNR which may not be directly applicable here. The growing-degree condition is necessary to guarantee the validity of all these methods which also result in different asymptotic laws than ours. As far as we know, an effective testing procedure that distinguishes SBMs with bounded degrees is still missing. As a side remark, the power Poisson law is unique in sparse network models with bounded degrees as demonstrated in Janson (1995). In the end, we mention a few papers addressing different models or testing problems than ours: Fosdick and Hoff (2015) proposed a test for examining dependence between network factors and nodal-level attributes; Maugis et al. (2017) proposed a variant of multivariate t-test for model diagnosis based on a collection of network samples.

2. LR-type test and asymptotic properties

The classic LR test requires the calculations of the marginal probability distributions of A_{uv}'s under both H_0 and H_1. By straightforward calculations, it can be shown that, under H_1, the marginal distribution of A is

$$P_1(A) = \sum_{\sigma\in\{\pm\}^n} P(A|\sigma)P(\sigma) = 2^{-n} \sum_{\sigma\in\{\pm\}^n} \prod_{u<v} p_{uv}(\sigma)^{A_{uv}} q_{uv}(\sigma)^{1-A_{uv}};$$

and under H_0, the marginal distribution of A is

$$P_0(A) = \prod_{u<v} p_0^{A_{uv}} q_0^{1-A_{uv}},$$

where $p_0 = 1 - q_0 = \frac{a+b}{2n}$. The classic LR test for (2) is then given as follows:

$$Y_n = \frac{P_1(A)}{P_0(A)} = 2^{-n} \sum_{\sigma\in\{\pm\}^n} \prod_{u<v} \left(\frac{p_{uv}(\sigma)}{p_0}\right)^{A_{uv}} \left(\frac{q_{uv}(\sigma)}{q_0}\right)^{1-A_{uv}}, \tag{3}$$

where $p_{uv}(\sigma)$ and $q_{uv}(\sigma)$ are defined in (1). However, Mossel et al. (2015) show that $P_0(\cdot)$ and $P_1(\cdot)$ are asymptotically orthogonal when $\kappa \geq 1$. So with positive probability, Y_n is asymptotically degenerate to either 0 or ∞. Here we provide a more heuristic understanding for such degenerateness phenomenon. Note that the probability ratio $\frac{p_{uv}(\sigma)}{p_0}$ is equal to either $\frac{2a}{a+b}$ or $\frac{2b}{a+b}$, depending on whether u, v belong to the same community. When $\kappa \geq 1$, i.e., $a - b$ is large compared with $a + b$, the two probability ratios considerably differ from each other which brings too much uncertainty into Y_n.

We propose a regularized LR test, called as ε-LR test, to resolve the degenerateness issue. The idea is quite natural: incorporate a regularization parameter ε into Y_n to reduce its uncertainty. Our ε-LR test is defined as follows. Let $\kappa_\varepsilon = \frac{(a_\varepsilon - b_\varepsilon)^2}{2(a+b)}$, where $a_\varepsilon = a - \varepsilon$, $b_\varepsilon = b + \varepsilon$. For any ε satisfying

$$0 < \varepsilon < \frac{a-b}{2} \text{ and } \kappa_\varepsilon < 1, \tag{4}$$

define

$$Y_n^\varepsilon = 2^{-n} \sum_{\sigma\in\{\pm\}^n} \prod_{u<v} \left(\frac{p_{uv}^\varepsilon(\sigma)}{p_0}\right)^{A_{uv}} \left(\frac{q_{uv}^\varepsilon(\sigma)}{q_0}\right)^{1-A_{uv}}, \tag{5}$$

where

$$p_{uv}^\varepsilon(\sigma) = \begin{cases} \frac{a_\varepsilon}{n}, & \sigma_u = \sigma_v \\ \frac{b_\varepsilon}{n}, & \sigma_u \neq \sigma_v \end{cases}, \quad q_{uv}^\varepsilon(\sigma) = 1 - p_{uv}^\varepsilon(\sigma).$$

In other words, we replace $p_{uv}(\sigma)$ and $q_{uv}(\sigma)$ in (3) by their counterparts $p_{uv}^\varepsilon(\sigma)$ and $q_{uv}^\varepsilon(\sigma)$. The new probability ratio $\frac{p_{uv}^\varepsilon(\sigma)}{p_0}$ is equal to either $\frac{2a_\varepsilon}{a+b}$ or $\frac{2b_\varepsilon}{a+b}$, which are closer to each other due to regularization. Such a trick will be proven to effectively reduce the variability of the classic LR test. Asymptotic distributions and power analysis of Y_n^ε are provided in subsequent Sections 2.1 and 2.2.

Remark 2.1. A more naive approach is to reject H_0 if $Y_n > c$ with $c > 0$ a predetermined constant. However, the choice of c is a challenging issue. In particular, due to the degenerateness of Y_n, it is hard to determine the (asymptotic) probability of rejection given any value of c, which poses challenges in analyzing size and power of the test. Instead, our ε-LR test has valid asymptotic distributions which avoids the above issues.

2.1. Power Poisson laws

Let us first present a power Poisson law for Y_n^ε under H_0.

Theorem 2.1. *If $\kappa \geq 1$ and ε satisfies (4), then under H_0, $Y_n^\varepsilon \xrightarrow{d} W_0^\varepsilon$ as $n \to \infty$, where*

$$W_0^\varepsilon = \prod_{m=3}^{\infty} \left(1+\delta_m^\varepsilon\right)^{Z_m^0} \exp\left(-\lambda_m \delta_m^\varepsilon\right), \quad Z_m^0 \overset{ind}{\sim} \text{Poisson}\left(\lambda_m\right).$$

Here, $\lambda_m = \frac{1}{2m}\left(\frac{a+b}{2}\right)^m$ and $\delta_m^\varepsilon = \left(\frac{a_\varepsilon - b_\varepsilon}{a+b}\right)^m$.

Theorem 2.1 shows that, under H_0, Y_n^ε converges in distribution to an infinite product of power Poisson variables. Its proof is based on a contiguity theory for regular random graphs developed by Janson (1995). Power Poisson law is unique in sparse network with bounded degree, e.g., the number of subgraphs, the number of perfect matchings and the number of edge colorings all follow such a law (see Janson (1995)). This decidedly differs from the growing-degree regime. For instance, when the average degree is growing along with n, Bickel and Sarkar (2016) proposed a spectral algorithm that follows Tracy–Widom law; Banerjee and Ma (2017), Banerjee (2018) examined the classic LR statistics under $\kappa < 1$ and linear spectral statistics relating to signed cycles that follow power Gaussian law; Gao and Lafferty (2017a,b) proposed subgraph-based algorithms that follow Gaussian distributions.

According to Theorem 2.1, we test (2) at significance level α based on the following rule:

reject H_0 iff $Y_n^\varepsilon \geq w_\alpha^\varepsilon$,

where $w_\alpha^\varepsilon > 0$ satisfies $P(W_0^\varepsilon \leq w_\alpha^\varepsilon) = 1-\alpha$.

The following theorem shows that, under H_1, Y_n^ε asymptotically follows another power Poisson law.

Theorem 2.2. *If $\kappa \geq 1$, ε satisfies (4) and $(a-b)(a_\varepsilon - b_\varepsilon) < \frac{2(a+b)}{3}$, then under H_1, $Y_n^\varepsilon \xrightarrow{d} W_1^\varepsilon$ as $n \to \infty$, where*

$$W_1^\varepsilon = \prod_{m=3}^{\infty} \left(1+\delta_m^\varepsilon\right)^{Z_m^1} \exp\left(-\lambda_m \delta_m^\varepsilon\right), \quad Z_m^1 \overset{ind}{\sim} \text{Poisson}\left(\lambda_m(1+\delta_m)\right).$$

Here, λ_m and δ_m^ε are the same as in Theorem 2.1 and $\delta_m = \left(\frac{a-b}{a+b}\right)^m$.

We notice that W_1^ε differs from W_0^ε only in the Poisson powers, i.e., Z_m^1 has larger means than Z_m^0. Intuitively, the power of Y_n^ε should increase when such differences become substantial.

Based on Theorems 2.1 and 2.2, we can derive the asymptotic power of Y_n^ε as stated in the corollary below. The power is an unexplicit function of (a, b, ε).

Corollary 2.3. *If $\kappa \geq 1$, ε satisfies (4) and $(a-b)(a_\varepsilon - b_\varepsilon) < \frac{2(a+b)}{3}$, then as $n \to \infty$, the power of Y_n^ε satisfies $P(\text{reject}H_0|\text{under}H_1) \to P(a,b,\varepsilon)$, where $P(a,b,\varepsilon) := P(W_1^\varepsilon \geq w_\alpha^\varepsilon)$.*

Remark 2.2. The value of ε can be empirically selected. Specifically, choose ε to satisfy (4) and $(a-b)(a_\varepsilon - b_\varepsilon) < \frac{2(a+b)}{3}$ with a, b therein replaced by their consistent estimators. Existence of such consistent estimators is guaranteed by Mossel et al. (2015) when $\kappa > 1$.

2.2. Power analysis

Corollary 2.3 derives an asymptotic power $P(a,b,\varepsilon)$ for Y_n^ε. In this section, we further examine this power and demonstrate whether and when it can approach one. It is challenging to directly analyze $P(a,b,\varepsilon)$ for fixed a, b due to the lack of explicit expression. Instead, we will consider the relatively easier growing-degree regime ($a+b \to \infty$) and discuss its connection to existing work. Theorem 2.4 provides an explicit expression for the limit of $P(a,b,\varepsilon)$. Let $\Phi(\cdot)$ denote the cumulative distribution function of standard normal variable and $z_{1-\alpha}$ denote its $1-\alpha$ quantile, i.e., $\Phi(z_{1-\alpha}) = 1-\alpha$.

Theorem 2.4. *If $\kappa \geq 1$ and $\varepsilon \in \left(0, \frac{a-b}{2}\right)$ satisfies, when $a+b \to \infty$, $\frac{(a_\varepsilon - b_\varepsilon)^2}{2(a+b)} \to k_1$ and $\frac{(a-b)(a_\varepsilon - b_\varepsilon)}{2(a+b)} \to k_2$ for constants $k_1, k_2 \in (0,1)$, then $P(a,b,\varepsilon) \to \Phi\left(\frac{\sigma_2^2}{\sigma_1} - z_{1-\alpha}\right)$ as $a+b \to \infty$, where $\sigma_l^2 = \sum_{m=3}^{\infty} \frac{1}{2m} k_l^m = -\frac{1}{2}\left(\log(1-k_l) + k_l + \frac{1}{2}k_l^2\right)$, $l = 1, 2$.*

We remark that the limit power $\Phi\left(\frac{\sigma_2^2}{\sigma_1} - z_{1-\alpha}\right)$ approaches one if $\kappa \to \infty$ (regardless of the choice of ε). To see this, note that

$$\frac{\sigma_2^2}{\sigma_1} = -\frac{\log(1-k_2) + k_2 + \frac{1}{2}k_2^2}{\sqrt{-2\left(\log(1-k_1) + k_1 + \frac{1}{2}k_1^2\right)}} \asymp \kappa^{3/2}. \tag{6}$$

(6) holds uniformly for ε satisfying the conditions of Theorem 2.4 and $\kappa^{3/2}$ on the right side is free of ε. If $\kappa \to \infty$, then $\frac{\sigma_2^2}{\sigma_1} \to \infty$, and so $\Phi\left(\frac{\sigma_2^2}{\sigma_1} - z_{1-\alpha}\right)$ approaches one. The power behavior merely relies on κ while being free of ε. Our result is closely relating to Banerjee and Ma (2017) who investigate the asymptotic power of the classic LR test which nonetheless requires $0 < \kappa < 1$. Montanari and Sen (2016) proposed an efficient method based on semidefinite program but their size and power are not explicitly quantifiable like ours.

2.3. Monte-Carlo approximation

Despite its theoretically nice properties, the test statistic Y_n^ε might be computationally infeasible. This can be easily seen from (5), i.e., Y_n^ε can be viewed as the average of the quantity $g_n^\varepsilon(\sigma)$ over the entire space of configurations $\{\pm\}^n$, where

$$g_n^\varepsilon(\sigma) = \prod_{u<v}\left(\frac{p_{uv}^\varepsilon(\sigma)}{p_0}\right)^{A_{uv}}\left(\frac{q_{uv}^\varepsilon(\sigma)}{q_0}\right)^{1-A_{uv}}.$$

The computational effort for the direct averages scales exponentially with n. So an accurate and computationally efficient approximation of Y_n^ε would be needed for practical use.

In this section, we consider the classical Monte Carlo (MC) method which randomly chooses a set of M configurations $\sigma[1], \ldots, \sigma[M]$ from $\{\pm\}^n$. The MC method works directly under the bounded-degree regime. The average Y_n^ε is naturally approximated by the sample mean of $g_n^\varepsilon(\sigma[l])$'s, which may substantially reduce computational cost if $M \ll 2^n$. However, a small choice of M may result in inaccurate approximation. An interesting question is how small M can be to ensure valid approximation. More explicitly, we aim to find an order of M such that the following approximation becomes valid:

$$\widehat{Y}_n^\varepsilon \equiv \frac{1}{M}\sum_{l=1}^{M} g_n^\varepsilon(\sigma[l]) = Y_n^\varepsilon + o_P(1). \tag{7}$$

The following theorem shows that the validity of (7) is possible.

Theorem 2.5. *Suppose $\kappa \geq 1$, ε satisfies (4), and $M \gg \exp\left(\frac{n\kappa_\varepsilon}{2}\right)$. Then (7) holds under H_0. Moreover, if $(a_\varepsilon - b_\varepsilon)(a-b) < a+b$, then (7) holds under H_1.*

According to Theorem 2.5, (7) becomes valid when only $M \gg \exp\left(\frac{n\kappa_\varepsilon}{2}\right)$ configurations are used, e.g., $M \asymp (\log n)^c \exp\left(\frac{n\kappa_\varepsilon}{2}\right)$ for a constant $c > 0$. When κ_ε is close to zero, this may substantially reduce the computational cost from $O(2^n)$ to nearly $O([\exp(\kappa_\varepsilon/2)]^n)$. However, MC method still requires heavy computation. A more efficient method would be highly useful.

3. Numerical studies

In this section, we examine the performance of the proposed testing procedure through simulation studies in Section 3.1, and through real-world data sets in Section 3.2.

3.1. Simulation

The empirical performance of our test statistic $\widehat{Y}_n^\varepsilon$ is demonstrated through simulation studies. We also compared our method with the spectral method proposed by Bickel and Sarkar (Bickel and Sarkar, 2016), the subgraph count method proposed by Gao and Lafferty (Gao and Lafferty, 2017a), the long-cycle counting method proposed by Mossel, Neeman and Sly in Mossel et al. (2015) and the Bethe Hessian estimation method proposed by Le and Levina in Le and Levina (2015). Note that the Bethe Hessian method in Le and Levina (2015) is to estimate the number of communities. To make the comparison, we calculate the correct estimation rates under H_0 and H_1 respectively. Firstly, we assume that both a and b are known and compare the four methods. Then we use the estimated a and b proposed in Mossel et al. (2015) to evaluate our test procedure. We evaluated the size and power of various methods at significance level 0.05. For size, data were generated from $\mathcal{G}\left(n, \frac{a+b}{2n}\right)$. For power, data were generated from $\mathcal{G}\left(n, \frac{a}{n}, \frac{b}{n}\right)$. Both size and power were calculated as proportions of rejections based on 500 independent experiments.

We examined various choices of a, b, n, ε for $\widehat{Y}_n^\varepsilon$. For convenience, denote $a = 2.5 + c$, $b = 2.5 - c$. We chose $n = 20, 30, 40, 45$ and $(c, \varepsilon) = (2.10, 1.10), (2.15, 1.15), (2.25, 1.25), (2.35, 1.35)$. The ε in each case was chosen to be approximately $c/2$. The corresponding values of SNR $\kappa = \frac{(a-b)^2}{2(a+b)}$ are $1.76, 1.85, 2.03, 2.21$. We chose $100n^3 e^{\frac{n\kappa_\varepsilon}{2}}$ samples for MC approximations according to Theorem 2.5 for calculating $\widehat{Y}_n^\varepsilon$. Table 1 summarizes the size and power of our test with the true a and b. For all cases, the sizes of the $\widehat{Y}_n^\varepsilon$ are close to the 0.05 nominal level indicating the validity of the test. For each choice of (c, ε), the power increases along with n. For any fixed n, the power increases as κ increases, consistent with Theorem 2.4 which states that the power should increase with κ.

Table 1
Power (Size) of ε-LR test based on various choices of c, ε, n with true a, b.

(c, ϵ)	κ	$n = 20$	$n = 30$	$n = 40$	$n = 45$
(2.10, 1.10)	1.76	0.572 (0.058)	0.654 (0.042)	0.730 (0.050)	0.812 (0.048)
(2.15, 1.15)	1.85	0.598 (0.054)	0.684 (0.052)	0.764 (0.050)	0.852 (0.040)
(2.25, 1.25)	2.03	0.626 (0.056)	0.704 (0.044)	0.802 (0.042)	0.910 (0.044)
(2.35, 1.35)	2.21	0.648 (0.044)	0.736 (0.042)	0.888 (0.058)	1.000 (0.042)

Table 2
Power (Size) of BS's spectral test based on various choices of c, n.

c	κ	$n = 20$	$n = 30$	$n = 40$	$n = 45$
2.10	1.76	0.308 (0.070)	0.260 (0.048)	0.266 (0.058)	0.300 (0.054)
2.15	1.85	0.330 (0.070)	0.280 (0.048)	0.288 (0.058)	0.314 (0.054)
2.25	2.03	0.364 (0.070)	0.336 (0.048)	0.348 (0.058)	0.362 (0.054)
2.35	2.21	0.400 (0.070)	0.376 (0.048)	0.402 (0.058)	0.402 (0.054)

Table 3
Power (Size) of GL's subgraph count test based on various choices of c, n.

c	κ	$n = 20$	$n = 30$	$n = 40$	$n = 45$
2.10	1.76	0.156 (0.05)	0.216 (0.05)	0.196 (0.05)	0.168 (0.05)
2.15	1.85	0.216 (0.05)	0.224 (0.05)	0.204 (0.05)	0.206 (0.05)
2.25	2.03	0.296 (0.05)	0.234 (0.05)	0.246 (0.05)	0.300 (0.05)
2.35	2.21	0.306 (0.05)	0.350 (0.05)	0.328 (0.05)	0.336 (0.05)

Table 4
Power(Size) by counting the number of $k_n = 3$ cycle by MNS with true a, b.

c	κ	$n = 20$	$n = 30$	$n = 40$	$n = 45$
2.10	1.76	0.172(0.058)	0.188(0.058)	0.218(0.056)	0.236(0.056)
2.15	1.85	0.198 (0.050)	0.206(0.054)	0.256(0.060)	0.256(0.054)
2.25	2.03	0.234 (0.046)	0.252(0.054)	0.272(0.052)	0.278(0.052)
2.35	2.21	0.250(0.052)	0.278(0.056)	0.294(0.058)	0.310(0.054)

Table 5
Proportion of correct estimation of k under $H_1(H_0)$ of LL's methods based on various choices of c, n.

c	κ	$n = 20$	$n = 30$	$n = 40$	$n = 45$
2.10	1.76	0.224 (0.948)	0.270 (0.976)	0.362 (0.980)	0.418 (0.978)
2.15	1.85	0.272 (0.970)	0.384 (0.970)	0.428 (0.974)	0.452 (0.974)
2.25	2.03	0.322 (0.966)	0.414 (0.960)	0.522 (0.966)	0.536 (0.978)
2.35	2.21	0.414 (0.952)	0.490 (0.982)	0.608 (0.966)	0.610 (0.964)

Table 6
Power (Size) of ε-LR test based on various choices of c, ε, n with estimated a, b.

(c, ϵ)	κ	$n = 20$	$n = 30$	$n = 40$	$n = 45$
(2.10, 1.10)	1.76	0.454 (0.062)	0.516 (0.060)	0.534 (0.058)	0.538 (0.056)
(2.15, 1.15)	1.85	0.510 (0.060)	0.540 (0.056)	0.542 (0.052)	0.546 (0.050)
(2.25, 1.25)	2.03	0.512 (0.060)	0.554 (0.054)	0.588 (0.046)	0.610 (0.050)
(2.35, 1.35)	2.21	0.550 (0.058)	0.610 (0.056)	0.680 (0.054)	0.726 (0.052)

Tables 2–4 summarize the size and power of BS's spectral method, GL's subgraph count method and MNS's long-cycle counting method with the true a and b. It is worth mentioning that the sizes of the three methods are free of c since the null models under various values of c are equivalent and the sizes of three methods are uniquely determined by the common null model. Due to the high sparsity of the simulated networks, subgraph counts are generally small, and we obtained the critical values for GL's method based on resampling instead of using asymptotic distribution. It is observed that the three methods achieve much smaller power (less than 0.402) than ε-LR while maintaining the correct size.

Tables 5 presents the correct estimation rates of LL with the true a and b (we take $r = \tilde{d}^{1/2}$ and $t = 5$ as proposed in Le and Levina (2015)). Under H_0, the LL's method has high estimation rates(above 0.950), while the rates are less than 0.610 under H_1, much smaller than the power of our proposed method.

Table 6 shows the size and power of the ε-LR method with estimated a and b. All the sizes are still close to the nominal 0.05 and the power has a similar pattern as that in Table 1. Even though there is a decrease in the power with estimated a and b, the powers are still significantly larger than that in Tables 2–5.

Table 7
Power (Size) of ε-LR test with estimated a, b, $n = 30$ and various r_+.

(c, ϵ)	κ	$r_+ = 0.5$	$r_+ = 0.55$	$r_+ = 0.6$	$r_+ = 0.75$
(2.10, 1.10)	1.76	0.516 (0.060)	0.500 (0.046)	0.474 (0.038)	0.282 (0.042)
(2.15, 1.15)	1.85	0.540 (0.056)	0.530 (0.052)	0.502 (0.044)	0.316 (0.056)
(2.25, 1.25)	2.03	0.554 (0.054)	0.546 (0.058)	0.512 (0.056)	0.302 (0.058)
(2.35, 1.35)	2.21	0.610 (0.056)	0.592 (0.056)	0.534 (0.054)	0.200 (0.040)

Table 8
P-values of three methods over communities C_I, C_{II}, C_{III} for Political Book Data.

Method	P-value		
	C_I	C_{II}	C_{III}
ε-LR	0.000	1.000	1.000
BS	0.000	0.000	0.000
GL	0.000	0.465	0.039

Table 9
Rejection proportions by three methods in different combination regimes for Political Book Data.

Method	Rejection proportion		
	C_I&C_{II}	C_I&C_{III}	C_{II}&C_{III}
ε-LR	100%	92%	83%
BS	100%	100%	100%
GL	100%	98%	100%

Next, we study the effect of unbalanced group sizes. Let $P(\sigma_i = +) = r_+$ and $P(\sigma_i = -) = r_-$ with $r_+ + r_- = 1$ and $r_+, r_- \in (0, 1)$, $i = 1, 2, \ldots, n$. On average, there are nr_+ and nr_- nodes in the '+' and '−' group respectively. In this simulation, we take $n = 30$, $r_+ = 0.5, 0.55, 0.60, 0.75$ and other parameters the same as in the previous simulation. The result is presented in Table 7. When r_+ increases, there is a decreasing trend in power. For r_+ close to 0.5 (balanced group size), the test has similar performance as $r_+ = 0.5$ (balanced case).

3.2. Real data analysis

In this section, we applied our procedure to analyze the political book data (Newman, 2006) which has 105 political books (nodes). Two books are connected if they were frequently co-purchased on Amazon. This data was analyzed by Zhao et al. (2011) who detected three communities. We used R package *igraph* based on a spin-glass model and simulated annealing to recover their findings, and denote the three communities by C_I, C_{II}, C_{III} which contain 20, 44, 41 nodes, respectively. Books within the same community are expected to demonstrate similar political tendencies. The aim of this study is to examine whether our method can detect the existence of the communities. Our ε-LR method was based on $M = 10^7 e^{n\delta}$ MC samples with $\delta \approx 0.01$ and $\varepsilon \approx \frac{\hat{a}+\hat{b}}{2}$, where $\widehat{a} = 21.633$, $\widehat{b} = 1.139$ are MLEs of a, b under H_1. We first examined whether H_0 is rejected (at 0.05) over each community. Table 8 summarizes the results. We find that all three methods rejected H_0 over C_I. This incorrect decision might be due to the small size of the first community. Moreover, ε-LR failed to reject H_0 over communities C_{II}, C_{III}; BS rejected H_0 over C_{II}, C_{III}; GL rejected H_0 over C_{III} while failed to reject H_0 over C_{II}.

We then examined whether H_0 is rejected for a subnetwork with nodes from two different communities. In particular, we uniformly sampled 20 nodes out of C_{II} without replacement, and combined it with C_I. Therefore, the combined network has two communities of 20 nodes each. We also examined the combination regimes C_I&C_{II} and C_{II}&C_{III} with 20 nodes uniformly sampled from C_{II} in the former and from both C_{II}, C_{III} in the latter (so each combination has 40 nodes in total). We repeated each combination regime 100 times and calculated the proportions that H_0 was rejected. Results are summarized in Table 9. It is observed that all three methods rejected H_0 100 times over C_I&C_{II} hence the success rates are 100%. The rejection rates of ε-LR over C_I&C_{III} and over C_{II}&C_{III} are 92% and 83% respectively. The success rates of BS 100%, and the success rates of GL are 98% and 100%, for both combination regimes.

4. Discussions

The work of Decelle et al. (2011) implies that extension of the current work to multi-community setting is highly important but nontrivial. As far as we know, only a few works address such settings but mostly in community detection. For instance, Neeman and Netrapalli (2014) provide a sufficient condition for impossible detection; Abbe and Sandon (2018) present an information-theoretic phase transition for the SNR to yield successful detection which strengthens the work of Neeman and Netrapalli (2014).

The test statistic $\widehat{Y}_n^{\varepsilon}$ can be viewed as a type of partition function over Gibbs field. Popular approximations of partition functions in statistical physics include MC approximations and mean-field approximations. This paper only considers the former while leaves the latter as a future topic. Mean-field approximation has proven to work well in dense magnetism such as Curie–Weiss model (see Weiss (1907)). Recently, validity of mean-field approximation was established by Basak and Mukherjee (2017) in the sparser settings which satisfy the so-called "mean-field assumption", i.e., the trace of the squared adjacency matrix is $o_P(n)$. This assumption fails in our setting in that the trace becomes $O_P(n)$. Additional theory is needed to extend the results of Basak and Mukherjee (2017).

Appendix. Proofs

In this section, we prove the main results of this paper. Our asymptotic results are derived based on the following Proposition A.1 which was proved by Janson in Janson (1995). For arbitrary non-negative integer x, let $[x]_j$ denote the descending factorial $x(x-1)\cdots(x-j+1)$.

Proposition A.1. *Let $\lambda_i > 0$, $i = 1, 2, \ldots$, be constants and suppose that for each n there are random variables X_{in}, $i = 1, 2, \ldots$, and Y_n (defined on the same probability space) such that X_{in} is non-negative integer valued and $E\{Y_n\} \neq 0$ (at least for large n), and furthermore the following conditions are satisfied:*

(A1) $X_{in} \xrightarrow{d} Z_i$ as $n \to \infty$, jointly for all i, where $Z_i \sim \text{Poisson}(\lambda_i)$ are independent Poisson random variables;
(A2) $\mathbb{E}\{Y_n[X_{1n}]_{j_1}\cdots[X_{kn}]_{j_k}\}/\mathbb{E}\{Y_n\} \to \prod_{i=1}^k \mu_i^{j_i}$, as $n \to \infty$, for some $\mu_i \geq 0$ and every finite sequence $j_1, \ldots, j_k$ of non-negative integers;
(A3) $\sum_{i=1}^\infty \lambda_i\delta_i^2 < \infty$, where $\delta_i = \mu_i/\lambda_i - 1$;
(A4) $\mathbb{E}\{Y_n^2\}/(\mathbb{E}\{Y_n\})^2 \to \exp\left(\sum_{i=1}^\infty \lambda_i\delta_i^2\right)$.

Then

$$\frac{Y_n}{\mathbb{E}\{Y_n\}} \xrightarrow{d} W \equiv \prod_{i=1}^{\infty}(1+\delta_i)^{Z_i}\exp(-\lambda_i\delta_i), \quad \text{as } n \to \infty.$$

Remark A.1. Janson (1995) (Janson, 1995) showed that the infinite product defining W in Proposition A.1 converges in L^2 a.s. with $\mathbb{E}W = 1$ and $\mathbb{E}W^2 = \exp\left(\sum_{i\geq 1}\lambda_i\delta_i^2\right)$.

Before proofs, we need the following lemma.

Lemma A.2. *For a random graph G with vertex $1, \ldots, n$, let X_{mn} be the number of m-cycles of G, for $m \geq 3$. Let $\lambda_m = \frac{1}{2m}\left(\frac{a+b}{2}\right)^m$ and $\delta_m = \left(\frac{a-b}{a+b}\right)^m$.*

1. *Under $G \sim \mathcal{G}(n, p_0)$, for any $k \geq 3$, $\{X_{mn}\}_{m=3}^k$ jointly converge to independent Poisson variables with mean λ_m.*
2. *Under $G \sim \mathcal{G}\left(n, \frac{a}{n}, \frac{b}{n}\right)$, for any $k \geq 3$, $\{X_{mn}\}_{m=3}^k$ jointly converge to independent Poisson variables with mean $\lambda_m(1+\delta_m)$.*

Proof of Lemma A.2. The first part was well known (see Mossel et al. (2015)). We only prove the second part.

Denote $\mathbb{E}_1$ the expectation based on hypothesis H_1. Let H be a graph on a subset of $[n]$ with vertex set $\mathcal{V}(H)$ and edge set $\mathcal{E}(H)$. Use 1_H to denote the 0–1 random variable that is 1 when $\mathcal{E}(H) \subseteq \mathcal{E}(G)$ and $P(H)$ for the probability that $1_H = 1$. For $3 \leq m \leq k$, let $H_{m1}, \ldots, H_{mj_m}$ be a j_m-tuple of distinct m-cycles. Then

$$\prod_{m=3}^{k}[X_{mn}]_{j_m} = \sum_{(H_{mi})}\prod_{m=3}^{k}\prod_{i=1}^{j_m}1_{H_{mi}},$$

where the sum ranges over all tuples of distinct cycles $\{H_{mi} : 3 \leq m \leq k, 1 \leq i \leq j_m\}$; each H_{mi} is an m-cycle and all cycles are distinct. Let A be the set of all such tuples of cycles for which the cycles are vertex-disjoint and let $\bar{A}$ be its complement, i.e., any tuple of $\bar{A}$ contains two cycles with at least one common vertex. Then

$$\begin{aligned}\mathbb{E}_1\prod_{m=3}^{k}[X_{mn}]_{j_m} &= \sum_{(H_{mi})}\mathbb{E}_1\prod_{m=3}^{k}\prod_{i=1}^{j_m}1_{H_{mi}} \\ &= \sum_{(H_{mi})\in A}\mathbb{E}_1\prod_{m=3}^{k}\prod_{i=1}^{j_m}1_{H_{mi}} + \sum_{(H_{mi})\in\bar{A}}\mathbb{E}_1\prod_{m=3}^{k}\prod_{i=1}^{j_m}1_{H_{mi}} \end{aligned} \tag{8}$$

Since the number of m-cycles on a graph of s vertexes is $\frac{s!}{(s-m)!2m}$ (two directions and m distinct starting vertexes give us $2m$ the same m-cycles), one gets that $|A| = \frac{n!}{(n-M)!}\prod_{m=3}^{k}\left(\frac{1}{2m}\right)^{j_m}$ with $M = \sum_{m=3}^{k} mj_m$ (see also (Bollobás, 2001, Chapter

4) for more complete derivation). Meanwhile, take τ uniformly from $\{\pm\}^n$ and define τ^{mi} be the restriction of τ on the vertexes of H_{mi}, and define $N_{mi} = \sum_{(u,v)\in\mathcal{E}(H_{mi})} 1(\tau_u^{mi} \neq \tau_v^{mi})$. The τ^{mi}'s are independent thanks to the vertex disjointness of H_{mi}'s. Following Mossel et al. (2015, Lemma 3.3) one can show that $P(N_{mi} = l) = 2^{-m+1}\binom{m}{l}$ for even $l \in [0, m]$ and zero for odd l. Then one has

$$\begin{aligned}
\mathbb{E}_1 \prod_{m=3}^{k}\prod_{i=1}^{j_m} 1_{H_{mi}} &= \mathbb{E}_\tau \mathbb{E}_1\{\prod_{m=3}^{k}\prod_{i=1}^{j_m} 1_{H_{mi}}|\tau\}\\
&= \mathbb{E}_\tau \prod_{m=3}^{k}\prod_{i=1}^{j_m}\prod_{(u,v)\in\mathcal{E}(H_{mi})}\left(\frac{a}{n}\right)^{1(\tau_u=\tau_v)}\left(\frac{b}{n}\right)^{1(\tau_u\neq\tau_v)}\\
&= \mathbb{E}_\tau \prod_{m=3}^{k}\prod_{i=1}^{j_m}\prod_{(u,v)\in\mathcal{E}(H_{mi})}\left(\frac{a}{n}\right)^{1(\tau_u^{mi}=\tau_v^{mi})}\left(\frac{b}{n}\right)^{1(\tau_u^{mi}\neq\tau_v^{mi})}\\
&= \mathbb{E}_\tau \prod_{m=3}^{k}\prod_{i=1}^{j_m}\left(\frac{a}{n}\right)^{m-N_{mi}}\left(\frac{b}{n}\right)^{N_{mi}}.
\end{aligned}$$

Since τ is broken into disjoint and independent $(\tau^{mi})_{3\le m\le k, 1\le i\le j_m}$, the above is equal to

$$\begin{aligned}
&\prod_{m=3}^{k}\prod_{i=1}^{j_m}\mathbb{E}_{\tau^{mi}}\left(\frac{a}{n}\right)^{m-N_{mi}}\left(\frac{b}{n}\right)^{N_{mi}}\\
&= \prod_{m=3}^{k}\prod_{i=1}^{j_m} 2^{-m}\left[\left(\frac{a+b}{n}\right)^m + \left(\frac{a-b}{n}\right)^m\right] = n^{-M}\prod_{m=3}^{k}\left[\left(\frac{a+b}{2}\right)^m(1+\delta_m)\right].
\end{aligned}$$

Then the first part of (8) becomes

$$\begin{aligned}
&|A| \times n^{-M}\prod_{m=3}^{k}\left[\left(\frac{a+b}{2}\right)^m(1+\delta_m)\right]\\
&= \frac{n!}{(n-M)!n^M}\prod_{m=3}^{k}(\lambda_m(1+\delta_m))^{j_m} \overset{n\to\infty}{\to} \prod_{m=3}^{k}(\lambda_m(1+\delta_m))^{j_m}.
\end{aligned}$$

On the other hand, for any $(H_{mi}) \in \bar{A}$, $H := \cup H_{mi}$ has at most $M-1$ vertexes and M edges, and $|\mathcal{E}(H)| > |\mathcal{V}(H)|$. Since

$$\mathbb{E}_1\{\prod_{m=3}^{k}\prod_{i=1}^{j_m} 1_{H_{mi}}|\tau\} = \prod_{(u,v)\in\mathcal{E}(H)}\left(\frac{a}{n}\right)^{1(\tau_u=\tau_v)}\left(\frac{b}{n}\right)^{1(\tau_u\neq\tau_v)} \le \left(\frac{\max\{a,b\}}{n}\right)^{|\mathcal{E}(H)|},$$

and there are $\binom{n}{|\mathcal{V}(H)|}|\mathcal{V}(H)|!$ graphs isomorphic to H, then

$$\sum_{H' \text{ is isomorphic to } H}\mathbb{E}_1\{1_{H'}|\tau\} \le \left(\frac{\max\{a,b\}}{n}\right)^{|\mathcal{E}(H)|}\binom{n}{|\mathcal{V}(H)|}|\mathcal{V}(H)|! \to 0.$$

Since there are a bounded number of isomorphism classes, the second part of (8) tends to zero as $n \to \infty$. Hence, $\mathbb{E}_1\prod_{m=3}^{k}[X_{mn}]_{j_m} \to \prod_{m=3}^{k}(\lambda_m(1+\delta_m))^{j_m}$ for any $k \ge 3$ and integers $j_3, \ldots, j_k$. It follows by Wormald (1999, Lemma 2.8) that the desirable result holds. □

A.1. Proofs in Section 2.1

Proof of Theorem 2.1. Let $\mathbb{E}_0$ denote the expectations under hypotheses H_0. We will use Proposition A.1 to prove the result, for which we will check the Conditions A1 to A4 therein. Some of the details are rooted in Mossel et al. (2015). To ease reading, we provide the detailed proofs. Obviously, $\mathbb{E}_0 Y_n^\varepsilon = 1$.

Let X_{mn} be the number of m-cycles of $G \sim \mathcal{G}(n, p_0)$, for $m \ge 3$. Following Lemma A.2 Part 1, for any $k \ge 3$, $\{X_{mn}\}_{m=3}^k$ jointly converge to independent Poisson variables with mean $\lambda_m = \frac{1}{2m}\left(\frac{a+b}{2}\right)^m$. This verifies Condition A1.

To check Condition A2, let $H = (H_{mi})_{3\le m\le k, 1\le i\le j_m}$ be a tuple of short cycles of disjoint vertexes; each H_{mi} is an m-cycle, $M = \sum_{m=3}^k m j_m$, and the vertexes of H_{mi}'s are disjoint. Let $\sigma^{1mi}, \sigma^{2mi}$ be the restrictions of σ over $\mathcal{V}(H_{mi})$ and $[n]\backslash\mathcal{V}(H_{mi})$, and σ^1, σ^2 be the restrictions of σ over $\mathcal{V}(H)$ and $[n]\backslash\mathcal{V}(H)$. By direct examinations we have

$$\mathbb{E}_0 Y_n^\varepsilon 1_H = 2^{-n}\sum_{\sigma\in\{\pm\}^n}\mathbb{E}_0 1_H \prod_{u<v}\left(\frac{p_{uv}^\varepsilon(\sigma)}{p_0}\right)^{A_{uv}}\left(\frac{q_{uv}^\varepsilon(\sigma)}{q_0}\right)^{1-A_{uv}}$$

$$
\begin{aligned}
&= 2^{-n} \sum_{\sigma\in\{\pm\}^n} \mathbb{E}_0 1_H \prod_{(u,v)\in\mathcal{E}(H)} \left(\frac{p^{\varepsilon}_{uv}(\sigma)}{p_0}\right)^{A_{uv}} \left(\frac{q^{\varepsilon}_{uv}(\sigma)}{q_0}\right)^{1-A_{uv}} \\
&\quad \times \prod_{(u,v)\in\overline{\mathcal{E}(H)}} \left(\frac{p^{\varepsilon}_{uv}(\sigma)}{p_0}\right)^{A_{uv}} \left(\frac{q^{\varepsilon}_{uv}(\sigma)}{q_0}\right)^{1-A_{uv}} \\
&= 2^{-n} \sum_{\sigma\in\{\pm\}^n} \mathbb{E}_0 1_H \prod_{(u,v)\in\mathcal{E}(H)} \left(\frac{p^{\varepsilon}_{uv}(\sigma)}{p_0}\right)^{A_{uv}} \left(\frac{q^{\varepsilon}_{uv}(\sigma)}{q_0}\right)^{1-A_{uv}}. \qquad (9)
\end{aligned}
$$

Since σ is broken into σ^1 and σ^2 which are supported on $\mathcal{V}$ and its complement respectively, and $p^{\varepsilon}_{uv}(\sigma)$, $q^{\varepsilon}_{uv}(\sigma)$ only depend on σ^1 when $(u, v) \in \mathcal{E}(H)$, (9) is equal to the following

$$
\begin{aligned}
&2^{-n} \sum_{\sigma^1\in\{\pm\}^{\mathcal{V}(H)}} \sum_{\sigma^2\in\{\pm\}^{[n]\setminus\mathcal{V}(H)}} \mathbb{E}_0 1_H \prod_{(u,v)\in\mathcal{E}(H)} \left(\frac{p^{\varepsilon}_{uv}(\sigma^1)}{p_0}\right)^{A_{uv}} \left(\frac{q^{\varepsilon}_{uv}(\sigma^1)}{q_0}\right)^{1-A_{uv}} \\
&= 2^{-M} \sum_{\sigma^1\in\{\pm\}^{\mathcal{V}(H)}} \mathbb{E}_0 1_H \prod_{(u,v)\in\mathcal{E}(H)} \left(\frac{p^{\varepsilon}_{uv}(\sigma^1)}{p_0}\right)^{A_{uv}} \left(\frac{q^{\varepsilon}_{uv}(\sigma^1)}{q_0}\right)^{1-A_{uv}}. \qquad (10)
\end{aligned}
$$

Since $1_H = 1$ implies $\mathcal{E}(H) \subset \mathcal{E}(G)$, any $(u, v) \in \mathcal{V}(H)$ leads to $A_{uv} = 1$. Meanwhile, $\mathbb{E}_0 1_H = p_0^M$, hence (10) equals

$$
\begin{aligned}
&2^{-M} p_0^M \sum_{\sigma^1\in\{\pm\}^{\mathcal{V}(H)}} \prod_{(u,v)\in\mathcal{E}(H)} \left(\frac{p^{\varepsilon}_{uv}(\sigma^1)}{p_0}\right) \\
&= \mathbb{E}_{\sigma^1} \prod_{(u,v)\in\mathcal{E}(H)} p^{\varepsilon}_{uv}(\sigma^1) = \prod_{m=3}^{k}\prod_{i=1}^{j_m} \mathbb{E}_{\sigma^{1mi}} \prod_{(u,v)\in\mathcal{E}(H_{mi})} p^{\varepsilon}_{uv}(\sigma^{1mi}) = n^{-M} \prod_{m=3}^{k}\prod_{i=1}^{j_m} \mathbb{E}_{\sigma^{1mi}} a_{\varepsilon}^{m-N_{mi}} b_{\varepsilon}^{N_{mi}},
\end{aligned}
$$

where $N_{mi} = \sum_{(u,v)\in\mathcal{E}(H_{mi})} 1(\sigma^{1mi}_u \neq \sigma^{1mi}_v)$, the number of edges over H_{mi} with distinct end points. Following the proof of Mossel et al. (2015, Lemma 3.3),

$$
\mathbb{E}_{\sigma^{1mi}} a_{\varepsilon}^{m-N_{mi}} b_{\varepsilon}^{N_{mi}} = 2^{-m}\left[(a_\varepsilon+b_\varepsilon)^m + (a_\varepsilon-b_\varepsilon)^m\right] = \left(\frac{a+b}{2}\right)^m \left(1+\left(\frac{a_\varepsilon-b_\varepsilon}{a+b}\right)^m\right).
$$

Hence,

$$
\mathbb{E}_0 Y_n^{\varepsilon} 1_H = n^{-M} \prod_{m=3}^{k} \left(\left(\frac{a+b}{2}\right)^m \left(1+\left(\frac{a_\varepsilon-b_\varepsilon}{a+b}\right)^m\right)\right)^{j_m}.
$$

Let A be the set of tuples $(H_{mi})_{3\le m\le k, 1\le i\le j_m}$ for which the cycles are vertex-disjoint and let $\bar{A}$ be its complement. Using $|A| = \frac{n!}{(n-M)!}\prod_{m=3}^{k}\left(\frac{1}{2m}\right)^{j_m}$ (see proof of Lemma A.2) we get that

$$
\begin{aligned}
\sum_{H\in A} \mathbb{E}_0 Y_n^{\varepsilon} 1_H &= |A| n^{-M} \prod_{m=3}^{k} \left(\left(\frac{a+b}{2}\right)^m \left(1+\left(\frac{a_\varepsilon-b_\varepsilon}{a+b}\right)^m\right)\right)^{j_m} \\
&\overset{n\to\infty}{\to} \prod_{m=3}^{k} \left(\frac{1}{2m}\left(\frac{a+b}{2}\right)^m \left(1+\left(\frac{a_\varepsilon-b_\varepsilon}{a+b}\right)^m\right)\right)^{j_m} \\
&= \prod_{m=3}^{k} \left(\lambda_m(1+\delta_m^{\varepsilon})\right)^{j_m},
\end{aligned}
$$

where $\delta_m^{\varepsilon} = \left(\frac{a_\varepsilon-b_\varepsilon}{a+b}\right)^m$. Similar to (9) one gets that, for $H \in \bar{A}$,

$$
\begin{aligned}
\mathbb{E}_0 Y_n^{\varepsilon} 1_H &= 2^{-n} \sum_{\sigma\in\{\pm\}^n} \mathbb{E}_0 1_H \prod_{(u,v)\in\mathcal{E}(H)} \left(\frac{p^{\varepsilon}_{uv}(\sigma)}{p_0}\right)^{A_{uv}} \left(\frac{q^{\varepsilon}_{uv}(\sigma)}{q_0}\right)^{1-A_{uv}} \\
&= 2^{-n} \sum_{\sigma\in\{\pm\}^n} \prod_{(u,v)\in\mathcal{E}(H)} \frac{p^{\varepsilon}_{uv}(\sigma)}{p_0} \times P_0(H) \\
&\le a_{\varepsilon}^{|\mathcal{E}(H)|} n^{-|\mathcal{E}(H)|},
\end{aligned}
$$

where the last inequality follows from $p^{\varepsilon}_{uv}(\sigma) \le a_\varepsilon/n$ and $P_0(H) = p_0^{|\mathcal{E}(H)|}$. So

$$\sum_{H' \text{ is isomorphic to } H} \mathbb{E}_0 Y_n^{\varepsilon} 1_H \le a_\varepsilon^{|\mathcal{E}(H)|} n^{-|\mathcal{E}(H)|} \binom{n}{|\mathcal{V}(H)|} |\mathcal{V}(H)|! \to 0,$$

which leads to $\sum_{H \in \bar{A}} \mathbb{E}_0 Y_n^{\varepsilon} 1_H \to 0$ using a similar argument as the proof of Lemma A.2 Part 2. So as $n \to \infty$,

$$\mathbb{E}_0 Y_n^{\varepsilon} [X_{3n}]_{j_3} \cdots [X_{kn}]_{j_k} = \sum_{H \in A} \mathbb{E}_0 Y_n^{\varepsilon} 1_H + \sum_{H \in \bar{A}} \mathbb{E}_0 Y_n^{\varepsilon} 1_H \to \prod_{m=3}^{k} \left(\lambda_m (1 + \delta_m^{\varepsilon}) \right)^{j_m},$$

which verifies Condition A2.

Condition A3 holds due to the following trivial fact:

$$\sum_{m \ge 3} \lambda_m (\delta_m^{\varepsilon})^2 = \sum_{m \ge 3} \frac{1}{2m} \left(\frac{(a_\varepsilon - b_\varepsilon)^2}{2(a+b)} \right)^m = \sum_{m \ge 3} \frac{1}{2m} \left(\frac{(a-b-2\varepsilon)^2}{2(a+b)} \right)^m < \infty.$$

In the end let us check Conditions A4. Let $N_{uv}^{\sigma\tau} = 1(\sigma_u = \sigma_v) + 1(\tau_u = \tau_v)$. Note that

$$\begin{aligned}
\mathbb{E}_0 (Y_n^{\varepsilon})^2 &= 4^{-n} \sum_{\sigma,\tau \in \{\pm\}^n} \prod_{u<v} \mathbb{E}_0 \left(\frac{p^{\varepsilon}_{uv}(\sigma) p^{\varepsilon}_{uv}(\tau)}{p_0^2} \right)^{A_{uv}} \left(\frac{q^{\varepsilon}_{uv}(\sigma) q^{\varepsilon}_{uv}(\tau)}{q_0^2} \right)^{1-A_{uv}} \\
&= 4^{-n} \sum_{\sigma,\tau \in \{\pm\}^n} \prod_{u<v} \left(\frac{p^{\varepsilon}_{uv}(\sigma) p^{\varepsilon}_{uv}(\tau)}{p_0} + \frac{q^{\varepsilon}_{uv}(\sigma) q^{\varepsilon}_{uv}(\tau)}{q_0} \right) \\
&= 4^{-n} \sum_{\sigma,\tau \in \{\pm\}^n} \prod_{u<v} \left(\frac{1}{p_0} \left(\frac{a_\varepsilon}{n}\right)^{N_{uv}^{\sigma\tau}} \left(\frac{b_\varepsilon}{n}\right)^{2-N_{uv}^{\sigma\tau}} + \frac{1}{q_0} \left(1 - \frac{a_\varepsilon}{n}\right)^{N_{uv}^{\sigma\tau}} \left(1 - \frac{b_\varepsilon}{n}\right)^{2-N_{uv}^{\sigma\tau}} \right) \\
&= 4^{-n} \sum_{\sigma,\tau \in \{\pm\}^n} \prod_{N_{uv}^{\sigma\tau}=0} \left(\frac{1}{p_0} \left(\frac{b_\varepsilon}{n}\right)^2 + \frac{1}{q_0} \left(1 - \frac{b_\varepsilon}{n}\right)^2 \right) \\
&\quad \times \prod_{N_{uv}^{\sigma\tau}=2} \left(\frac{1}{p_0} \left(\frac{a_\varepsilon}{n}\right)^2 + \frac{1}{q_0} \left(1 - \frac{a_\varepsilon}{n}\right)^2 \right) \\
&\quad \times \prod_{N_{uv}^{\sigma\tau}=1} \left(\frac{1}{p_0} \left(\frac{a_\varepsilon}{n}\right) \left(\frac{b_\varepsilon}{n}\right) + \frac{1}{q_0} \left(1 - \frac{a_\varepsilon}{n}\right) \left(1 - \frac{b_\varepsilon}{n}\right) \right).
\end{aligned}$$

It is easy to check that

$$\begin{aligned}
\frac{1}{p_0} \left(\frac{b_\varepsilon}{n}\right)^2 + \frac{1}{q_0} \left(1 - \frac{b_\varepsilon}{n}\right)^2 &= 1 + \gamma_n^{\varepsilon} + O(n^{-3}) \\
\frac{1}{p_0} \left(\frac{a_\varepsilon}{n}\right)^2 + \frac{1}{q_0} \left(1 - \frac{a_\varepsilon}{n}\right)^2 &= 1 + \gamma_n^{\varepsilon} + O(n^{-3}) \\
\frac{1}{p_0} \left(\frac{a_\varepsilon}{n}\right) \left(\frac{b_\varepsilon}{n}\right) + \frac{1}{q_0} \left(1 - \frac{a_\varepsilon}{n}\right) \left(1 - \frac{b_\varepsilon}{n}\right) &= 1 - \gamma_n^{\varepsilon} + O(n^{-3}),
\end{aligned} \tag{11}$$

where $\gamma_n^{\varepsilon} = \frac{\kappa_\varepsilon}{n} + \frac{(a_\varepsilon - b_\varepsilon)^2}{4n^2}$, $\kappa_\varepsilon = \frac{(a_\varepsilon - b_\varepsilon)^2}{2(a+b)}$. Let

$$s_+ = \#\{(u,v) : u < v, \sigma_u \sigma_v \tau_u \tau_v = +\}, s_- = \#\{(u,v) : u < v, \sigma_u \sigma_v \tau_u \tau_v = -\}.$$

Let $\rho = \frac{1}{n} \sum_{u=1}^{n} \sigma_u \tau_u$. Following Mossel et al. (2015), we have $s_+ = \frac{n^2}{4}(1+\rho^2) - \frac{n}{2}$ and $s_- = \frac{n^2}{4}(1-\rho^2)$. Then using the approximation technique in Mossel et al. (2015), i.e., Lemmas 5.3, 5.4, 5.5 therein, it holds that

$$\begin{aligned}
\mathbb{E}_0 (Y_n^{\varepsilon})^2 &= 4^{-n} \sum_{\sigma,\tau} (1 + \gamma_n^{\varepsilon} + O(n^{-3}))^{s_+} (1 - \gamma_n^{\varepsilon} + O(n^{-3}))^{s_-} \\
&= (1 + o(1)) 4^{-n} \sum_{\sigma,\tau} (1 + \gamma_n^{\varepsilon})^{\frac{n^2}{4}(1+\rho^2) - \frac{n}{2}} (1 - \gamma_n^{\varepsilon})^{\frac{n^2}{4}(1-\gamma^2)} \\
&= (1 + o(1)) \exp\left(-\kappa_\varepsilon^2/4 - \kappa_\varepsilon/2\right) 4^{-n} \sum_{\sigma,\tau} \exp\left(\frac{\rho^2}{2} \left(n\kappa_\varepsilon + \frac{(a_\varepsilon - b_\varepsilon)^2}{4} \right) \right) \\
&= (1 + o(1)) \exp\left(-\kappa_\varepsilon^2/4 - \kappa_\varepsilon/2\right) \mathbb{E}_{\sigma\tau} \exp\left(\frac{\rho^2}{2} \left(n\kappa_\varepsilon + \frac{(a_\varepsilon - b_\varepsilon)^2}{4} \right) \right) \\
&\overset{n\to\infty}{\to} \exp\left(-\kappa_\varepsilon^2/4 - \kappa_\varepsilon/2\right) (1 - \kappa_\varepsilon)^{-1/2} = \exp\left(\sum_{m=3}^{\infty} \lambda_m (\delta_m^{\varepsilon})^2 \right).
\end{aligned}$$

This verifies Condition A4. The result of Theorem 2.1 follows from Proposition A.1. □

Proof of Theorem 2.2. Let X_{mn} be the number of m-cycles of G, for $m \geq 3$. Let $\lambda_m = \frac{1}{2m}\left(\frac{a+b}{2}\right)^m$ and $\delta_m = \left(\frac{a-b}{a+b}\right)^m$. It follows by Lemma A.2 Part 2 that, under H_1, $\{X_{mn}\}_{m=3}^k$ jointly converge to independent Poisson variables with mean $\lambda_m(1+\delta_m)$, verifying Condition A1 of Proposition A.1. This leaves us to check Conditions A2 to A4. Let $M = \sum_{m=3}^k mj_m$ for integers $j_3, \ldots, j_k$ and $k \geq 3$.

Check Condition A2. Denote $\mathbb{E}_1$ the expectation based on hypothesis H_1. Let X_{mn} be the number of m-cycles of G, for $m \geq 3$ and $[x]_j$ be the descending factorial. Define $M = \sum_{m=3}^k mj_m$ for $k \geq 3$ and integers $j_3, \ldots, j_k$. To check A2, notice that

$$\mathbb{E}_1 Y_n^{\varepsilon}[X_{3n}]_{j_3}\cdots[X_{kn}]_{j_k} = \sum_{(H_{mi})_{3\leq m\leq k, 1\leq i\leq j_m}} \mathbb{E}_1 Y_n^{\varepsilon} 1_{\cup H_{mi}} \tag{12}$$

$$= \sum_{(H_{mi})\in A} \mathbb{E}_1 Y_n^{\varepsilon} 1_{\cup H_{mi}} + \sum_{(H_{mi})\in \bar{A}} \mathbb{E}_1 Y_n^{\varepsilon} 1_{\cup H_{mi}}, \tag{13}$$

where the sum in (12) ranges over $\mathcal{H}$, the collection of all M-tuples of cycles $(H_{mi})_{3\leq m\leq k, 1\leq i\leq j_m}$ with each H_{mi} an m-cycle, and A in the sum of (13) is the set of such tuples for which the cycles are vertex-disjoint and let $\bar{A} = \mathcal{H}\backslash A$, i.e., $\bar{A}$ contains M-tuples of cycles $(H_{mi})_{3\leq m\leq k, 1\leq i\leq j_m}$ with at least one common vertex among those cycles. Let us look at the first part of (13). Take τ uniformly distributed from $\{\pm\}^n$. For any $H = (H_{mi}) \in \mathcal{H}$, define τ^1, τ^2 to be the restrictions of τ over $\mathcal{V}(H)$ and $[n]\backslash\mathcal{V}(H)$ respectively.

One can check that, for any $H \in \mathcal{H}$,

$$\begin{aligned}
&\mathbb{E}_1\{Y_n^{\varepsilon} 1_H|\tau\}\\
&= \mathbb{E}_1\left\{1_H 2^{-n}\sum_{\sigma}\prod_{u<v}\left(\frac{p_{uv}^{\varepsilon}(\sigma)}{p_0}\right)^{A_{uv}}\left(\frac{q_{uv}^{\varepsilon}(\sigma)}{q_0}\right)^{1-A_{uv}}\Bigg|\tau\right\}\\
&= 2^{-n}\sum_{\sigma}\mathbb{E}_1\left\{1_H\prod_{u<v}\left(\frac{p_{uv}^{\varepsilon}(\sigma)}{p_0}\right)^{A_{uv}}\left(\frac{q_{uv}^{\varepsilon}(\sigma)}{q_0}\right)^{1-A_{uv}}\Bigg|\tau\right\}\\
&= 2^{-n}\sum_{\sigma}\mathbb{E}_1\left\{1_H\prod_{(u,v)\in\mathcal{E}(H)}\left(\frac{p_{uv}^{\varepsilon}(\sigma)}{p_0}\right)^{A_{uv}}\left(\frac{q_{uv}^{\varepsilon}(\sigma)}{q_0}\right)^{1-A_{uv}}\Bigg|\tau\right\}\\
&\quad\times\prod_{(u,v)\in\overline{\mathcal{E}(H)}}\mathbb{E}_1\left\{\left(\frac{p_{uv}^{\varepsilon}(\sigma)}{p_0}\right)^{A_{uv}}\left(\frac{q_{uv}^{\varepsilon}(\sigma)}{q_0}\right)^{1-A_{uv}}\Bigg|\tau\right\}\\
&= 2^{-n}\sum_{\sigma}\mathbb{E}_1\{1_H|\tau\}\prod_{(u,v)\in\mathcal{E}(H)}\left(\frac{p_{uv}^{\varepsilon}(\sigma)}{p_0}\right)\prod_{(u,v)\in\overline{\mathcal{E}(H)}}\left(\frac{p_{uv}^{\varepsilon}(\sigma)p_{uv}(\tau)}{p_0}+\frac{q_{uv}^{\varepsilon}(\sigma)q_{uv}(\tau)}{q_0}\right)\\
&= 2^{-n}\sum_{\sigma}\prod_{(u,v)\in\mathcal{E}(H)}\left(\frac{p_{uv}^{\varepsilon}(\sigma^1)p_{uv}(\tau^1)}{p_0}\right)\prod_{(u,v)\in\overline{\mathcal{E}(H)}}\left(\frac{p_{uv}^{\varepsilon}(\sigma)p_{uv}(\tau)}{p_0}+\frac{q_{uv}^{\varepsilon}(\sigma)q_{uv}(\tau)}{q_0}\right),
\end{aligned}\tag{14}$$

which leads to that

$$\begin{aligned}
&\mathbb{E}_1 Y_n^{\varepsilon} 1_H\\
&= 2^{-2n}\sum_{\tau}\sum_{\sigma}\prod_{(u,v)\in\mathcal{E}(H)}\left(\frac{p_{uv}^{\varepsilon}(\sigma^1)p_{uv}(\tau^1)}{p_0}\right)\prod_{(u,v)\in\overline{\mathcal{E}(H)}}\left(\frac{p_{uv}^{\varepsilon}(\sigma)p_{uv}(\tau)}{p_0}+\frac{q_{uv}^{\varepsilon}(\sigma)q_{uv}(\tau)}{q_0}\right)\\
&= \mathbb{E}_{\sigma\tau}\prod_{(u,v)\in\mathcal{E}(H)}\left(\frac{p_{uv}^{\varepsilon}(\sigma^1)p_{uv}(\tau^1)}{p_0}\right)\prod_{(u,v)\in\overline{\mathcal{E}(H)}}\left(\frac{p_{uv}^{\varepsilon}(\sigma)p_{uv}(\tau)}{p_0}+\frac{q_{uv}^{\varepsilon}(\sigma)q_{uv}(\tau)}{q_0}\right)\\
&\equiv \mathbb{E}_{\sigma\tau}X_H^{\varepsilon}(\sigma^1,\tau^1)W_H^{\varepsilon}(\sigma,\tau)Z_H^{\varepsilon}(\sigma^2,\tau^2),
\end{aligned}\tag{15}$$

where

$$X_H^{\varepsilon}(\sigma^1,\tau^1) = \prod_{(u,v)\in\mathcal{E}(H)}\left(\frac{1}{p_0}p_{uv}^{\varepsilon}(\sigma^1)p_{uv}(\tau^1)\right),$$

$$W_H^{\varepsilon}(\sigma,\tau) = \prod_{(u,v)\in S_1(H)}\left(\frac{1}{p_0}p_{uv}^{\varepsilon}(\sigma)p_{uv}(\tau)+\frac{1}{q_0}q_{uv}^{\varepsilon}(\sigma)q_{uv}(\tau)\right)$$

and

$$Z_H^{\varepsilon}(\sigma^2,\tau^2) = \prod_{(u,v)\in S_2(H)}\left(\frac{1}{p_0}p_{uv}^{\varepsilon}(\sigma^2)p_{uv}(\tau^2)+\frac{1}{q_0}q_{uv}^{\varepsilon}(\sigma^2)q_{uv}(\tau^2)\right).$$

Here $S_1(H) = \{(u,v)\in\overline{\mathcal{E}(H)} : u\in\mathcal{V}(H) \text{ or } v\in\mathcal{V}(H)\}$ and $S_2(H) = \{(u,v)\in\overline{\mathcal{E}(H)} : u,v\notin\mathcal{V}(H)\}$.

We will show that $W_H^\varepsilon(\sigma,\tau)$ is uniformly bounded over σ,τ,H, and that

$$\sup_{H\in\mathcal{H}}|W_H^\varepsilon(\sigma,\tau)-1|\to 0,\ \ a.s. \tag{16}$$

To see this, observe that

$$\begin{aligned}W_H^\varepsilon(\sigma,\tau)=&\prod_{(u,v)\in\overline{\mathcal{E}(H)},u,v\in\mathcal{V}(H)}\left(\frac{1}{p_0}p_{uv}^\varepsilon(\sigma)p_{uv}(\tau)+\frac{1}{q_0}q_{uv}^\varepsilon(\sigma)q_{uv}(\tau)\right)\\&\times\prod_{v\in\mathcal{V}(H)}\prod_{u\notin\mathcal{V}(H)}\left(\frac{1}{p_0}p_{uv}^\varepsilon(\sigma)p_{uv}(\tau)+\frac{1}{q_0}q_{uv}^\varepsilon(\sigma)q_{uv}(\tau)\right).\end{aligned} \tag{17}$$

We note that

$$\frac{1}{p_0}p_{uv}^\varepsilon(\sigma)p_{uv}(\tau)+\frac{1}{q_0}q_{uv}^\varepsilon(\sigma)q_{uv}(\tau)=1+O(n^{-1}),$$

where the $O(n^{-1})$ term is uniform for u,v,σ,τ,H. The first product in (17) is therefore equal to $(1+O(n^{-1}))^{\binom{M}{2}-M}=1+o(1)$. We turn to the second product in (17). For any $v\in\mathcal{V}$, let

$$\begin{aligned}S_v^1&=\#\{u\notin\mathcal{V}(H):\sigma_u^2=\sigma_v^1,\tau_u^2=\tau_v^1\}\\S_v^2&=\#\{u\notin\mathcal{V}(H):\sigma_u^2=\sigma_v^1,\tau_u^2\neq\tau_v^1\}\\S_v^3&=\#\{u\notin\mathcal{V}(H):\sigma_u^2\neq\sigma_v^1,\tau_u^2=\tau_v^1\}\\S_v^4&=\#\{u\notin\mathcal{V}(H):\sigma_u^2\neq\sigma_v^1,\tau_u^2\neq\tau_v^1\}.\end{aligned}$$

Also let $S_{ll'}=\#\{u\notin\mathcal{V}(H):\sigma_u^2=l,\tau_u^2=l'\}$ and $N_{ll'}=\{v\in\mathcal{V}(H):\sigma_v^1=l,\tau_v^1=l'\}$ for $l,l'=\pm$. Then the second product in (17) equals

$$\begin{aligned}&\prod_{v\in\mathcal{V}(H)}\prod_{u\notin\mathcal{V}(H)}\left(\frac{1}{p_0}\left(\frac{a_\varepsilon}{n}\right)^{1(\sigma_u^2=\sigma_v^1)}\left(\frac{b_\varepsilon}{n}\right)^{1(\sigma_u^2\neq\sigma_v^1)}\left(\frac{a}{n}\right)^{1(\tau_u^2=\tau_v^1)}\left(\frac{b}{n}\right)^{1(\tau_u^2\neq\tau_v^1)}\right.\\&\left.+\frac{1}{q_0}\left(1-\frac{a_\varepsilon}{n}\right)^{1(\sigma_u^2=\sigma_v^1)}\left(1-\frac{b_\varepsilon}{n}\right)^{1(\sigma_u^2\neq\sigma_v^1)}\left(1-\frac{a}{n}\right)^{1(\tau_u^2=\tau_v^1)}\left(1-\frac{b}{n}\right)^{1(\tau_u^2\neq\tau_v^1)}\right)\\=&\prod_{v\in\mathcal{V}(H)}\left(\frac{1}{p_0}\left(\frac{a_\varepsilon}{n}\right)\left(\frac{a}{n}\right)+\frac{1}{q_0}\left(1-\frac{a_\varepsilon}{n}\right)\left(1-\frac{a}{n}\right)\right)^{S_v^1}\\&\times\left(\frac{1}{p_0}\left(\frac{a_\varepsilon}{n}\right)\left(\frac{b}{n}\right)+\frac{1}{q_0}\left(1-\frac{a_\varepsilon}{n}\right)\left(1-\frac{b}{n}\right)\right)^{S_v^2}\\&\times\left(\frac{1}{p_0}\left(\frac{b_\varepsilon}{n}\right)\left(\frac{a}{n}\right)+\frac{1}{q_0}\left(1-\frac{b_\varepsilon}{n}\right)\left(1-\frac{a}{n}\right)\right)^{S_v^3}\\&\times\left(\frac{1}{p_0}\left(\frac{b_\varepsilon}{n}\right)\left(\frac{b}{n}\right)+\frac{1}{q_0}\left(1-\frac{b_\varepsilon}{n}\right)\left(1-\frac{b}{n}\right)\right)^{S_v^4}\\=&(1+\widetilde{\gamma}_n^\varepsilon+O(n^{-3}))^{\sum_{v\in\mathcal{V}(H)}(S_v^1+S_v^4)}(1-\widetilde{\gamma}_n^\varepsilon+O(n^{-3}))^{\sum_{v\in\mathcal{V}(H)}(S_v^2+S_v^3)}.\end{aligned}$$

In the above we have used the following trivial facts:

$$\begin{aligned}\frac{1}{p_0}\left(\frac{a_\varepsilon}{n}\right)\left(\frac{a}{n}\right)+\frac{1}{q_0}\left(1-\frac{a_\varepsilon}{n}\right)\left(1-\frac{a}{n}\right)&=1+\widetilde{\gamma}_n^\varepsilon+O(n^{-3})\\\frac{1}{p_0}\left(\frac{a_\varepsilon}{n}\right)\left(\frac{b}{n}\right)+\frac{1}{q_0}\left(1-\frac{a_\varepsilon}{n}\right)\left(1-\frac{b}{n}\right)&=1-\widetilde{\gamma}_n^\varepsilon+O(n^{-3})\\\frac{1}{p_0}\left(\frac{b_\varepsilon}{n}\right)\left(\frac{a}{n}\right)+\frac{1}{q_0}\left(1-\frac{b_\varepsilon}{n}\right)\left(1-\frac{a}{n}\right)&=1-\widetilde{\gamma}_n^\varepsilon+O(n^{-3})\\\frac{1}{p_0}\left(\frac{b_\varepsilon}{n}\right)\left(\frac{b}{n}\right)+\frac{1}{q_0}\left(1-\frac{b_\varepsilon}{n}\right)\left(1-\frac{b}{n}\right)&=1+\widetilde{\gamma}_n^\varepsilon+O(n^{-3}),\end{aligned}$$

where $\widetilde{\gamma}_n^{\varepsilon} = \frac{\widetilde{\kappa}_\varepsilon}{n} + \frac{(a-b)(a_\varepsilon - b_\varepsilon)}{4n^2}$ and $\widetilde{\kappa}_\varepsilon = \frac{(a-b)(a_\varepsilon-b_\varepsilon)}{2(a+b)}$. Note that

$$\begin{aligned}\sum_{v\in\mathcal{V}(H)}(S_v^1+S_v^4) &= \sum_{\sigma_v^1=+,\tau_v^1=+}(S_{++}+S_{--}) + \sum_{\sigma_v^1=+,\tau_v^1=-}(S_{+-}+S_{-+})\\ &\quad + \sum_{\sigma_v^1=-,\tau_v^1=+}(S_{-+}+S_{+-}) + \sum_{\sigma_v^1=-,\tau_v^1=-}(S_{--}+S_{++})\\ &= (S_{++}+S_{--})(N_{++}+N_{--}) + (S_{+-}+S_{-+})(N_{+-}+N_{-+}) \equiv N_1,\end{aligned}$$

similarly,

$$\sum_{v\in\mathcal{V}(H)}(S_v^2+S_v^3) = (S_{+-}+S_{-+})(N_{++}+N_{--}) + (S_{++}+S_{--})(N_{+-}+N_{-+}) \equiv N_2.$$

So the second product in (17) equals

$$(1+o(1))(1+\widetilde{\gamma}_n^{\varepsilon})^{N_1}(1-\widetilde{\gamma}_n^{\varepsilon})^{N_2} = (1+o(1))\exp\left(\frac{N_1-N_2}{n}\widetilde{\kappa}_\varepsilon\right),$$

where the $o(1)$ term is uniform for u, v, σ, τ, H, thanks to $N_1, N_2 \le Mn$. By law of large number, $(N_1 - N_2)/n \to 0$, a.s., uniformly for $H \in \mathcal{H}$. Therefore (16) holds. The above analysis also shows that $W_H^{\varepsilon}(\sigma, \tau)$ is uniformly bounded over σ, τ, H.

Next let us analyze the term $Z_H^{\varepsilon}(\sigma^2, \tau^2)$. By Taylor expansions and direct examinations it can be checked that for $u, v \in [n]\backslash\mathcal{V}(H)$,

$$\frac{1}{p_0}p_{uv}^{\varepsilon}(\sigma^2)p_{uv}(\tau^2) + \frac{1}{q_0}q_{uv}^{\varepsilon}(\sigma^2)q_{uv}(\tau^2) = \begin{cases} 1+\widetilde{\gamma}_n^{\varepsilon} + O(n^{-3}), & \text{if } \sigma_u^2\sigma_v^2\tau_u^2\tau_v^2 = +\\ 1-\widetilde{\gamma}_n^{\varepsilon} + O(n^{-3}), & \text{if } \sigma_u^2\sigma_v^2\tau_u^2\tau_v^2 = -.\end{cases}$$

Let $s_+ = \#\{(u, v) : u, v \in [n]\backslash\mathcal{V}(H), u < v, \sigma_u^2\sigma_v^2\tau_u^2\tau_v^2 = +\}$ and $s_- = \#\{(u, v) : u, v \in [n]\backslash\mathcal{V}(H), u < v, \sigma_u^2\sigma_v^2\tau_u^2\tau_v^2 = -\}$. Let $\rho = \rho(\sigma^2, \tau^2) = \frac{1}{n-M}\sum_{u\in[n]\backslash\mathcal{V}(H)}\sigma_u^2\tau_u^2$. By direct examinations we have

$$\begin{aligned} Z_H^{\varepsilon}(\sigma^2,\tau^2) &= (1+\widetilde{\gamma}_n^{\varepsilon}+O(n^{-3}))^{s_+}(1-\widetilde{\gamma}_n^{\varepsilon}+O(n^{-3}))^{s_-}\\ &= (1+o(1))(1+\widetilde{\gamma}_n^{\varepsilon})^{s_+}(1-\widetilde{\gamma}_n^{\varepsilon})^{s_-}\\ &= (1+o(1))\exp\left(-\frac{\widetilde{\kappa}_\varepsilon^2(n-M)^2}{4n^2} - \frac{\widetilde{\kappa}_\varepsilon(n-M)}{2n}\right)\\ &\quad\times\exp\left(\frac{(\sqrt{n-M}\rho)^2}{2}\left(\frac{(a-b)(a_\varepsilon-b_\varepsilon)(n-M)}{4n^2} + \frac{\widetilde{\kappa}_\varepsilon(n-M)}{n}\right)\right). \end{aligned} \tag{18}$$

By the condition $(a-b)(a_\varepsilon - b_\varepsilon) < 2(a+b)/3$, $\widetilde{\kappa}_\varepsilon < 1$. Let $Z_n = \sqrt{n-M}\rho$. Let $\kappa_n = \frac{(a-b)(a_\varepsilon-b_\varepsilon)(n-M)}{4n^2} + \frac{\widetilde{\kappa}_\varepsilon(n-M)}{n}$ which is nonrandom tending to $\widetilde{\kappa}_\varepsilon$. By Hoeffding's inequality: for any $C > 0$,

$$P\left(\exp(\kappa_n Z_n^2/2) \ge C\right) \le 2C^{-1/\kappa_n}. \tag{19}$$

From (18) there exists a universal constant C_0 such that $Z_H^{\varepsilon}(\sigma^2, \tau^2) \le C_0\exp(\kappa_n Z_n^2/2)$, hence, it follows from (19) that for all $C > 0$,

$$P\left(Z_H^{\varepsilon}(\sigma^2,\tau^2) \ge C\right) \le 2(C/C_0)^{-1/\kappa_n}.$$

Therefore, by (19) we have that

$$\begin{aligned} &\mathbb{E}_{\sigma^2\tau^2}Z_H^{\varepsilon}(\sigma^2,\tau^2)1(Z_H^{\varepsilon}(\sigma^2,\tau^2)\ge C)\\ &= \int_0^\infty P\left(Z_H^{\varepsilon}(\sigma^2,\tau^2)1(Z_H^{\varepsilon}(\sigma^2,\tau^2)\ge C) > t\right)dt\\ &= CP\left(Z_H^{\varepsilon}(\sigma^2,\tau^2)\ge C\right) + \int_C^\infty P\left(Z_H^{\varepsilon}(\sigma^2,\tau^2) > t\right)dt\\ &\le 2C_0^{1/\kappa_n}C^{1-1/\kappa_n}/(1-\kappa_n). \end{aligned} \tag{20}$$

We can also show that, as $n \to \infty$,

$$\sup_{H\in\mathcal{H}}|\mathbb{E}_{\sigma^2\tau^2}Z_H^{\varepsilon}(\sigma^2,\tau^2) - \exp\left(-\widetilde{\kappa}_\varepsilon^2/4 - \widetilde{\kappa}_\varepsilon/2\right)(1-\widetilde{\kappa}_\varepsilon)^{-1/2}| \to 0, n\to\infty. \tag{21}$$

To see this, let $\rho_0 = \frac{1}{n-M}\sum_{u\in[n]}\sigma_u\tau_u$ and $r_H = \frac{1}{n-M}\sum_{u\in\mathcal{V}(H)}\sigma_u\tau_u$, therefore, $\rho = \rho_0 - r_H$. Let $Z_{0n} = \sqrt{n-M}\rho_0$. Then for any $H \in \mathcal{H}$, $|r_H| \le M/(n-M)$ which leads to

$$Z_{0n}^2 - 2M|\rho_0| \le Z_n^2 \le Z_{0n}^2 - 2M|\rho_0| + M^2/(n-M).$$

Both left and right hand sides in the above are free of H and converge to χ_1^2 thanks to $\rho_0 \to 0$, a.s. So

$$\sup_{H\in\mathcal{H}} |\mathbb{E}\exp(\kappa_n Z_n^2/2) - (1-\widetilde{\kappa}_\varepsilon)^{-1/2}| \to 0, \ \ n\to\infty.$$

This, together with (18), proves (21).

Next let us analyze $X_H^\varepsilon(\sigma^1,\tau^1)$. Assume $H=(H_{mi})_{3\le m\le k, 1\le i\le j_m}\in A$. For $3\le m\le k$ and $1\le i\le j_m$, let τ^{1mi}, σ^{1mi} be the restrictions of τ^1, σ^1 over the vertexes of H_{mi}. Since H_{mi} are vertex-disjoint, τ^{1mi}'s, σ^{1mi}'s are all independent. Let $N_{mi}=\sum_{(u,v)\in\mathcal{E}(H_{mi})} 1(\sigma_u^{1mi}\neq\sigma_v^{1mi})$, the number of edges over H_{mi} with distinct end points. Following the proof of Mossel et al. (2015, Lemma 3.3), we get that

$$\begin{aligned}
&\mathbb{E}_{\sigma^1\tau^1} X_H^\varepsilon(\sigma^1,\tau^1)\\
&=\mathbb{E}_{\sigma^1\tau^1}\prod_{(u,v)\in\mathcal{E}(H)}\left(\frac{1}{p_0}p_{uv}^\varepsilon(\sigma^1)p_{uv}(\tau^1)\right)\\
&=p_0^{-M}\prod_{m=3}^{k}\prod_{i=1}^{j_m}\mathbb{E}_{\sigma^1}\prod_{(u,v)\in\mathcal{E}(H_{mi})}p_{uv}^\varepsilon(\sigma^{1mi})\times\mathbb{E}_{\tau^1}\prod_{(u,v)\in\mathcal{E}(H_{mi})}p_{uv}(\tau^{1mi})\\
&=p_0^{-M}n^{-2M}\prod_{m=3}^{k}\prod_{i=1}^{j_m}\mathbb{E}_{\sigma^{1mi}}a_\varepsilon^{m-N_{mi}}b_\varepsilon^{N_{mi}}\times\mathbb{E}_{\tau^{1mi}}a^{m-N_{mi}}b^{N_{mi}}\\
&=p_0^{-M}n^{-2M}\prod_{m=3}^{k}\prod_{i=1}^{j_m}2^{-m}\left[(a_\varepsilon+b_\varepsilon)^m+(a_\varepsilon-b_\varepsilon)^m\right]\times 2^{-m}\left[(a+b)^m+(a-b)^m\right]\\
&=n^{-M}\prod_{m=3}^{k}\left[\left(\frac{a+b}{2}\right)^m(1+\delta_m)(1+\delta_m^\varepsilon)\right]^{j_m},
\end{aligned}\tag{22}$$

recalling $\delta_m=\left(\frac{a-b}{a+b}\right)^m$ and $\delta_m^\varepsilon=\left(\frac{a_\varepsilon-b_\varepsilon}{a+b}\right)^m$. Meanwhile, it is easy to see that $n^M X_H^\varepsilon(\sigma^1,\tau^1)$ is almost surely bounded and the bound is unrelated to the vertexes of H, i.e.,

$$n^M X_H^\varepsilon(\sigma^1,\tau^1)\le\left(\frac{2a_\varepsilon a}{a+b}\right)^M, \quad \forall\sigma^1,\tau^1\in\{\pm\}^{\mathcal{V}(H)}.\tag{23}$$

By (16), (20), (23), and bounded convergence theorem, we can show that

$$\sum_{H\in A}\mathbb{E}_{\sigma\tau}X_H^\varepsilon(\sigma^1,\tau^1)|W_H^\varepsilon(\sigma,\tau)-1|Z_H^\varepsilon(\sigma^2,\tau^2)\to 0.\tag{24}$$

More precisely, using $|A|=\frac{n!}{(n-M)!}\prod_{m=3}^k\left(\frac{1}{2m}\right)^{j_m}$ (see proof of Lemma A.2), (24) follows from the following

$$\begin{aligned}
&\sum_{H\in A}\mathbb{E}_{\sigma\tau}X_H^\varepsilon(\sigma^1,\tau^1)|W_H^\varepsilon(\sigma,\tau)-1|Z_H^\varepsilon(\sigma^2,\tau^2)\\
=&\sum_{H\in A}\mathbb{E}_{\sigma\tau}X_H^\varepsilon(\sigma^1,\tau^1)|W_H^\varepsilon(\sigma,\tau)-1|Z_H^\varepsilon(\sigma^2,\tau^2)1(Z_H^\varepsilon(\sigma^2,\tau^2)\le C)\\
&+\sum_{H\in A}\mathbb{E}_{\sigma\tau}X_H^\varepsilon(\sigma^1,\tau^1)|W_H^\varepsilon(\sigma,\tau)-1|Z_H^\varepsilon(\sigma^2,\tau^2)1(Z_H(\sigma^2,\tau^2)>C)\\
\lesssim&\ Cn^{-M}|A|\mathbb{E}_{\sigma\tau}\sup_{H\in A}|W_H^\varepsilon(\sigma,\tau)-1|+n^{-M}|A|\sup_{H\in A}\mathbb{E}_{\sigma\tau}Z_H^\varepsilon(\sigma^2,\tau^2)1(Z_H^\varepsilon(\sigma^2,\tau^2)>C)\\
\to&\ 0,
\end{aligned}$$

where the last limit follows by first taking $C\to\infty$ and then $n\to\infty$. By (15), (21) and (22), we have that

$$\begin{aligned}
&\sum_{H\in A}\mathbb{E}_1 Y_n^\varepsilon 1_H\\
=&\ |A|n^{-M}\prod_{m=3}^{k}\left[\left(\frac{a+b}{2}\right)^m(1+\delta_m^\varepsilon)(1+\delta_m)\right]^{j_m}\exp(-\widetilde{\kappa}_\varepsilon^2/4-\widetilde{\kappa}_\varepsilon/2)/\sqrt{1-\widetilde{\kappa}_\varepsilon}+o(1)\\
\overset{n\to\infty}{\to}&\prod_{m=3}^{k}(\lambda_m(1+\delta_m^\varepsilon)(1+\delta_m))^{j_m}\exp(-\widetilde{\kappa}_\varepsilon^2/4-\widetilde{\kappa}_\varepsilon/2)/\sqrt{1-\widetilde{\kappa}_\varepsilon},
\end{aligned}$$

recalling $\lambda_m=\frac{1}{2m}\left(\frac{a+b}{2}\right)^m$.

From (15), the uniform boundedness of $\mathbb{E}_{\sigma^2\tau^2}Z_H^{\varepsilon}(\sigma^2,\tau^2)$ and the uniform boundedness of $W_H^{\varepsilon}(\sigma,\tau)$, and the independence of $\sigma^1,\tau^1,\sigma^2,\tau^2$ that, there exists a constant C_1 s.t. for any $H\in\bar{A}$,

$$\mathbb{E}_1 Y_n^{\varepsilon}1_H \le C_1\mathbb{E}_{\sigma^1\tau^1}X_H^{\varepsilon}(\sigma^1,\tau^1).$$

Also notice from the definition of X_H^{ε} that

$$\begin{aligned}X_H^{\varepsilon}(\sigma^1,\tau^1) &= p_0^{-|\mathcal{E}(H)|}\prod_{(u,v)\in\mathcal{E}(H)}\left(\frac{a_\varepsilon}{n}\right)^{1(\sigma_u^1=\sigma_v^1)+1(\tau_u^1=\tau_v^1)}\left(\frac{b_\varepsilon}{n}\right)^{1(\sigma_u^1\neq\sigma_v^1)+1(\tau_u^1\neq\tau_v^1)}\\ &\le n^{-2|\mathcal{E}(H)|}p_0^{-|\mathcal{E}(H)|}a_\varepsilon^{2|\mathcal{E}(H)|} = n^{-|\mathcal{E}(H)|}\left(\frac{2a_\varepsilon^2}{a+b}\right)^{|\mathcal{E}(H)|}.\end{aligned}$$

Since there are at most $\binom{n}{|\mathcal{V}(H)|}|\mathcal{V}(H)|!$ graphs isomorphic to H, and $|\mathcal{E}(H)|>|\mathcal{V}(H)|$ for $H\in\bar{A}$, we get that, as $n\to\infty$,

$$\sum_{H' \text{ is isomorphic to } H}\mathbb{E}_1 Y_n^{\varepsilon}1_H \le C_1\left(\frac{2a_\varepsilon^2}{a+b}\right)^{|\mathcal{E}(H)|}n^{-|\mathcal{E}(H)|}\binom{n}{|\mathcal{V}(H)|}|\mathcal{V}(H)|!\to 0.$$

Since there is a bounded number of isomorphism classes, we get that the second part of (13) tends to zero as $n\to\infty$.

Hence, as $n\to\infty$,

$$(12)\to\prod_{m=3}^{k}(\lambda_m(1+\delta_m^{\varepsilon})(1+\delta_m))^{j_m}\exp(-\widetilde{\kappa}_\varepsilon^2/4-\widetilde{\kappa}_\varepsilon/2)/\sqrt{1-\widetilde{\kappa}_\varepsilon}.$$

As for $\mathbb{E}_1 Y_n^{\varepsilon}$, note that it is equal to

$$\mathbb{E}_1 Y_n^{\varepsilon} = 4^{-n}\sum_{\sigma,\tau}\prod_{u<v}\left(\frac{1}{p_0}p_{uv}^{\varepsilon}(\sigma)p_{uv}(\tau)+\frac{1}{q_0}q_{uv}^{\varepsilon}(\sigma)q_{uv}(\tau)\right).$$

Similar to (21), i.e., taking H therein as empty graph, one gets that

$$\mathbb{E}_1 Y_n^{\varepsilon}\to\exp(-\widetilde{\kappa}_\varepsilon^2/4-\widetilde{\kappa}_\varepsilon/2)/\sqrt{1-\widetilde{\kappa}_\varepsilon}. \tag{25}$$

Hence,

$$\frac{\mathbb{E}_1 Y_n^{\varepsilon}[X_{3n}]_{j_3}\cdots[X_{kn}]_{j_k}}{\mathbb{E}_1 Y_n^{\varepsilon}}\overset{n\to\infty}{\to}\prod_{m=3}^{k}(\lambda_m(1+\delta_m^{\varepsilon})(1+\delta_m))^{j_m}.$$

This verifies Condition A2.

Check Condition A3. Since $\frac{\lambda_m(1+\delta_m^{\varepsilon})(1+\delta_m)}{\lambda_m(1+\delta_m)}-1=\delta_m^{\varepsilon}$, and by (4), we have

$$\sum_{m\ge 3}\lambda_m(1+\delta_m)(\delta_m^{\varepsilon})^2=\sum_{m=3}^{\infty}\frac{1}{2m}\left(\frac{(a_\varepsilon-b_\varepsilon)^2}{2(a+b)}\right)^m+\sum_{m=3}^{\infty}\frac{1}{2m}\left(\frac{(a_\varepsilon-b_\varepsilon)^2(a-b)}{2(a+b)^2}\right)^m<\infty.$$

Check Condition A4. By direct examinations it can be checked that

$$\begin{aligned}\mathbb{E}_1\{(Y_n^{\varepsilon})^2|\tau\} &= 4^{-n}\sum_{\sigma}\sum_{\eta}\mathbb{E}_1\left\{\prod_{u<v}\left(\frac{p_{uv}^{\varepsilon}(\sigma)p_{uv}^{\varepsilon}(\eta)}{p_0^2}\right)^{A_{uv}}\left(\frac{q_{uv}^{\varepsilon}(\sigma)q_{uv}^{\varepsilon}(\eta)}{q_0^2}\right)^{1-A_{uv}}\middle|\tau\right\}\\ &= 4^{-n}\sum_{\sigma}\sum_{\eta}\prod_{u<v}\left(\frac{1}{p_0^2}p_{uv}^{\varepsilon}(\sigma)p_{uv}^{\varepsilon}(\eta)p_{uv}(\tau)+\frac{1}{q_0^2}q_{uv}^{\varepsilon}(\sigma)q_{uv}^{\varepsilon}(\eta)q_{uv}(\tau)\right)\end{aligned}$$

So

$$\begin{aligned}\mathbb{E}_1(Y_n^{\varepsilon})^2 &= 8^{-n}\sum_{\sigma}\sum_{\eta}\sum_{\tau}\prod_{u<v}\left(\frac{1}{p_0^2}p_{uv}^{\varepsilon}(\sigma)p_{uv}^{\varepsilon}(\eta)p_{uv}(\tau)+\frac{1}{q_0^2}q_{uv}^{\varepsilon}(\sigma)q_{uv}^{\varepsilon}(\eta)q_{uv}(\tau)\right)\\ &= \mathbb{E}_{\sigma\eta\tau}\prod_{u<v}\left(\frac{1}{p_0^2}p_{uv}^{\varepsilon}(\sigma)p_{uv}^{\varepsilon}(\eta)p_{uv}(\tau)+\frac{1}{q_0^2}q_{uv}^{\varepsilon}(\sigma)q_{uv}^{\varepsilon}(\eta)q_{uv}(\tau)\right),\end{aligned} \tag{26}$$

where σ, η, τ in the above expectation are independent and uniformly distributed over $\{\pm\}^n$. By Taylor expansion and straightforward (but exhaustive) calculations, it can be shown that

$$\begin{aligned}
\frac{1}{p_0^2}\left(\frac{a_\varepsilon}{n}\right)^2\left(\frac{a}{n}\right)+\frac{1}{q_0^2}\left(1-\frac{a_\varepsilon}{n}\right)^2\left(1-\frac{a}{n}\right) &= 1+\gamma_{2+}+O(n^{-3})\\
\frac{1}{p_0^2}\left(\frac{a_\varepsilon}{n}\right)^2\left(\frac{b}{n}\right)+\frac{1}{q_0^2}\left(1-\frac{a_\varepsilon}{n}\right)^2\left(1-\frac{b}{n}\right) &= 1+\gamma_{2-}+O(n^{-3})\\
\frac{1}{p_0^2}\left(\frac{a_\varepsilon}{n}\right)\left(\frac{b_\varepsilon}{n}\right)\left(\frac{a}{n}\right)+\frac{1}{q_0^2}\left(1-\frac{a_\varepsilon}{n}\right)\left(1-\frac{b_\varepsilon}{n}\right)\left(1-\frac{a}{n}\right) &= 1+\gamma_{1+}+O(n^{-3})\\
\frac{1}{p_0^2}\left(\frac{a_\varepsilon}{n}\right)\left(\frac{b_\varepsilon}{n}\right)\left(\frac{b}{n}\right)+\frac{1}{q_0^2}\left(1-\frac{a_\varepsilon}{n}\right)\left(1-\frac{b_\varepsilon}{n}\right)\left(1-\frac{b}{n}\right) &= 1+\gamma_{1-}+O(n^{-3})\\
\frac{1}{p_0^2}\left(\frac{b_\varepsilon}{n}\right)^2\left(\frac{a}{n}\right)+\frac{1}{q_0^2}\left(1-\frac{b_\varepsilon}{n}\right)^2\left(1-\frac{a}{n}\right) &= 1+\gamma_{0+}+O(n^{-3})\\
\frac{1}{p_0^2}\left(\frac{b_\varepsilon}{n}\right)^2\left(\frac{b}{n}\right)+\frac{1}{q_0^2}\left(1-\frac{b_\varepsilon}{n}\right)^2\left(1-\frac{b}{n}\right) &= 1+\gamma_{0-}+O(n^{-3})
\end{aligned}\tag{27}$$

where

$$\begin{aligned}
\gamma_{2+} &= \frac{(a_\varepsilon-b_\varepsilon)^2(2a_\varepsilon+b_\varepsilon)+x_\varepsilon}{n(a+b)^2}+\frac{3(a_\varepsilon-b_\varepsilon)^2+y_\varepsilon}{4n^2}\\
\gamma_{2-} &= -\frac{a_\varepsilon(a_\varepsilon-b_\varepsilon)^2+x_\varepsilon}{n(a+b)^2}-\frac{(a_\varepsilon-b_\varepsilon)^2+y_\varepsilon}{4n^2}\\
\gamma_{1+} &= -\frac{a_\varepsilon(a_\varepsilon-b_\varepsilon)^2+z_\varepsilon}{n(a+b)^2}-\frac{(a_\varepsilon-b_\varepsilon)^2}{4n^2}\\
\gamma_{1-} &= -\frac{b_\varepsilon(a_\varepsilon-b_\varepsilon)^2-z_\varepsilon}{n(a+b)^2}-\frac{(a_\varepsilon-b_\varepsilon)^2}{4n^2}\\
\gamma_{0+} &= -\frac{b_\varepsilon(a_\varepsilon-b_\varepsilon)^2+w_\varepsilon}{n(a+b)^2}-\frac{(a_\varepsilon-b_\varepsilon)^2+y_\varepsilon}{4n^2}\\
\gamma_{0-} &= \frac{(a_\varepsilon-b_\varepsilon)^2(a_\varepsilon+2b_\varepsilon)+w_\varepsilon}{n(a+b)^2}+\frac{3(a_\varepsilon-b_\varepsilon)^2+y_\varepsilon}{4n^2}
\end{aligned}$$

with

$$x_\varepsilon=\varepsilon(a_\varepsilon-b_\varepsilon)(3a_\varepsilon+b_\varepsilon),\ y_\varepsilon=4\varepsilon(a_\varepsilon-b_\varepsilon),\ z_\varepsilon=\varepsilon(a_\varepsilon-b_\varepsilon)^2,\ w_\varepsilon=\varepsilon(a_\varepsilon-b_\varepsilon)(a_\varepsilon+3b_\varepsilon).$$

Define $s_{r+}=\#\{(u,v):u<v, N_{uv}^{\sigma\tau}=r, \tau_u\tau_v=+\}$ and $s_{r-}=\#\{(u,v):u<v, N_{uv}^{\sigma\tau}=r, \tau_u\tau_v=-\}$, for $r=0,1,2$. Then it holds that

$$\begin{aligned}
\mathbb{E}_1(Y_n^\varepsilon)^2 &= \mathbb{E}_{\sigma\eta\tau}\prod_{r=0,1,2}(1+\gamma_{r+}+O(n^{-3}))^{s_{r+}}\times\prod_{r=0,1,2}(1+\gamma_{r-}+O(n^{-3}))^{s_{r-}}\\
&= (1+o(1))\mathbb{E}_{\sigma\eta\tau}\prod_{r=0,1,2}(1+\gamma_{r+})^{s_{r+}}\times\prod_{r=0,1,2}(1+\gamma_{r-})^{s_{r-}}.
\end{aligned}\tag{28}$$

Define

$$\begin{aligned}
\rho_1 &= \frac{1}{\sqrt{n}}\sum_{u=1}^n\sigma_u,\ \rho_2=\frac{1}{\sqrt{n}}\sum_{u=1}^n\eta_u,\ \rho_3=\frac{1}{\sqrt{n}}\sum_{u=1}^n\tau_u,\\
\rho_4 &= \frac{1}{\sqrt{n}}\sum_{u=1}^n\sigma_u\eta_u,\ \rho_5=\frac{1}{\sqrt{n}}\sum_{u=1}^n\sigma_u\tau_u,\ \rho_6=\frac{1}{\sqrt{n}}\sum_{u=1}^n\eta_u\tau_u,\ \rho_7=\frac{1}{\sqrt{n}}\sum_{u=1}^n\sigma_u\eta_u\tau_u.
\end{aligned}\tag{29}$$

Observe that

$$\begin{aligned}
s_{2+} &= \sum_{u<v}1(\sigma_u\sigma_v=+)1(\eta_u\eta_v=+)1(\tau_u\tau_v=+)\\
&= \frac{n^2}{16}-\frac{n}{2}+\frac{n}{16}\left(\rho_1^2+\rho_2^2+\rho_3^2+\rho_4^2+\rho_5^2+\rho_6^2+\rho_7^2\right)\\
s_{2-} &= \sum_{u<v}1(\sigma_u\sigma_v=+)1(\eta_u\eta_v=+)1(\tau_u\tau_v=-)\\
&= \frac{n^2}{16}+\frac{n}{16}\left(\rho_1^2+\rho_2^2-\rho_3^2+\rho_4^2-\rho_5^2-\rho_6^2-\rho_7^2\right)
\end{aligned}$$

$$s_{1+} = \sum_{u<v} 1(\sigma_u\sigma_v\eta_u\eta_v = -)1(\tau_u\tau_v = +) = \frac{n^2}{8} + \frac{n}{8}\left(\rho_3^2 - \rho_4^2 - \rho_7^2\right)$$
$$s_{1-} = \sum_{u<v} 1(\sigma_u\sigma_v\eta_u\eta_v = -)1(\tau_u\tau_v = -) = \frac{n^2}{8} - \frac{n}{8}\left(\rho_3^2 + \rho_4^2 - \rho_7^2\right)$$

$$\begin{aligned} s_{0+} &= \sum_{u<v} 1(\sigma_u\sigma_v = -)1(\eta_u\eta_v = -)1(\tau_u\tau_v = +) \\ &= \frac{n^2}{16} + \frac{n}{16}\left(-\rho_1^2 - \rho_2^2 + \rho_3^2 + \rho_4^2 - \rho_5^2 - \rho_6^2 + \rho_7^2\right) \\ s_{0-} &= \sum_{u<v} 1(\sigma_u\sigma_v = -)1(\eta_u\eta_v = -)1(\tau_u\tau_v = -) \\ &= \frac{n^2}{16} + \frac{n}{16}\left(-\rho_1^2 - \rho_2^2 - \rho_3^2 + \rho_4^2 + \rho_5^2 + \rho_6^2 - \rho_7^2\right). \end{aligned}$$

Using the above notation $\gamma_{r\pm}$'s and $s_{r\pm}$'s we can write the right hand side of (28) as

$$\prod_{r=0,1,2} (1+\gamma_{r+})^{s_{r+}} \times \prod_{r=0,1,2} (1+\gamma_{r-})^{s_{r-}} \equiv T_1 \times T_2,$$

where

$$\begin{aligned} T_1 &= (1+\gamma_{2+})^{\frac{n^2}{16}-\frac{n}{2}}(1+\gamma_{2-})^{\frac{n^2}{16}}(1+\gamma_{1+})^{\frac{n^2}{8}}(1+\gamma_{1-})^{\frac{n^2}{8}}(1+\gamma_{0+})^{\frac{n^2}{16}}(1+\gamma_{0-})^{\frac{n^2}{16}} \\ &= (1+o(1))\exp\Bigg(-\frac{(a_\varepsilon-b_\varepsilon)^2}{4(a+b)^4}\left[(a_\varepsilon-b_\varepsilon)^2(a_\varepsilon^2+a_\varepsilon b_\varepsilon+b_\varepsilon^2)+\varepsilon(a_\varepsilon-b_\varepsilon)(3a_\varepsilon^2+2a_\varepsilon b_\varepsilon+3b_\varepsilon^2)\right. \\ &\quad \left.+\varepsilon^2(3a_\varepsilon^2+2a_\varepsilon b_\varepsilon+3b_\varepsilon^2)\right] - \frac{(a_\varepsilon-b_\varepsilon)^2(2a_\varepsilon+b_\varepsilon)+x_\varepsilon}{2(a+b)^2}\Bigg) \\ &= (1+o(1))\exp\Bigg(-\frac{(a_\varepsilon-b_\varepsilon)^4}{16(a+b)^2} - \frac{(a_\varepsilon-b_\varepsilon)^4(a-b)^2}{16(a+b)^4} - \frac{(a_\varepsilon-b_\varepsilon)^2(a-b)^2}{8(a+b)^2} \\ &\quad -\frac{(a_\varepsilon-b_\varepsilon)^2(a-b)}{4(a+b)^2} - \frac{(a_\varepsilon-b_\varepsilon)^2}{4(a+b)} - \frac{(a-b)(a_\varepsilon-b_\varepsilon)}{2(a+b)}\Bigg) \\ &= (1+o(1))\exp\left(-\kappa_\varepsilon^2/4-\kappa_\varepsilon/2\right)\exp\left(-\widetilde{\kappa}_\varepsilon^2/2-\widetilde{\kappa}_\varepsilon\right)\exp\left(-\frac{(a_\varepsilon-b_\varepsilon)^4(a-b)^2}{16(a+b)^4} - \frac{(a_\varepsilon-b_\varepsilon)^2(a-b)}{4(a+b)^2}\right), \end{aligned}$$

and

$$\begin{aligned} T_2 &= (1+o(1))(1+\gamma_{2+})^{\frac{n}{16}\left(\rho_1^2+\rho_2^2+\rho_3^2+\rho_4^2+\rho_5^2+\rho_6^2+\rho_7^2\right)}(1+\gamma_{2-})^{\frac{n}{16}\left(\rho_1^2+\rho_2^2-\rho_3^2+\rho_4^2-\rho_5^2-\rho_6^2-\rho_7^2\right)} \\ &\quad \times(1+\gamma_{1+})^{\frac{n}{8}\left(\rho_3^2-\rho_4^2-\rho_7^2\right)}(1+\gamma_{1-})^{-\frac{n}{8}\left(\rho_3^2+\rho_4^2-\rho_7^2\right)} \\ &\quad \times(1+\gamma_{0+})^{\frac{n}{16}\left(-\rho_1^2-\rho_2^2+\rho_3^2+\rho_4^2-\rho_5^2-\rho_6^2+\rho_7^2\right)}(1+\gamma_{0-})^{\frac{n}{16}\left(-\rho_1^2-\rho_2^2-\rho_3^2+\rho_4^2+\rho_5^2+\rho_6^2-\rho_7^2\right)} \\ &= (1+o(1))\exp\Bigg(\frac{(a_\varepsilon-b_\varepsilon)^2(2a_\varepsilon+b_\varepsilon)+x_\varepsilon}{16(a+b)^2}\left(\rho_1^2+\rho_2^2+\rho_3^2+\rho_4^2+\rho_5^2+\rho_6^2+\rho_7^2\right) \\ &\quad -\frac{a_\varepsilon(a_\varepsilon-b_\varepsilon)^2+x_\varepsilon}{16(a+b)^2}\left(\rho_1^2+\rho_2^2-\rho_3^2+\rho_4^2-\rho_5^2-\rho_6^2-\rho_7^2\right) \\ &\quad -\frac{a_\varepsilon(a_\varepsilon-b_\varepsilon)^2+z_\varepsilon}{8(a+b)^2}\left(\rho_3^2-\rho_4^2-\rho_7^2\right) \\ &\quad +\frac{b_\varepsilon(a_\varepsilon-b_\varepsilon)^2-z_\varepsilon}{8(a+b)^2}\left(\rho_3^2+\rho_4^2-\rho_7^2\right) \\ &\quad +\frac{b_\varepsilon(a_\varepsilon-b_\varepsilon)^2+w_\varepsilon}{16(a+b)^2}\left(\rho_1^2+\rho_2^2-\rho_3^2-\rho_4^2+\rho_5^2+\rho_6^2-\rho_7^2\right) \\ &\quad -\frac{(a_\varepsilon-b_\varepsilon)^2(a_\varepsilon+2b_\varepsilon)+w_\varepsilon}{16(a+b)^2}\left(\rho_1^2+\rho_2^2+\rho_3^2-\rho_4^2-\rho_5^2-\rho_6^2+\rho_7^2\right)\Bigg) \\ &= (1+o(1))\exp\left(\frac{\kappa_\varepsilon}{2}\rho_4^2 + \frac{\widetilde{\kappa}_\varepsilon}{2}\left(\rho_5^2+\rho_6^2\right) + \frac{(a_\varepsilon-b_\varepsilon)^2(a-b)}{4(a+b)^2}\rho_7^2\right). \end{aligned}$$

We note that ρ_7 is independent of (ρ_4, ρ_5, ρ_6), and the condition $\kappa_\varepsilon < \widetilde{\kappa}_\varepsilon \in (0, 1/3)$ leads to uniform integrability of $\exp\left(\frac{\kappa_\varepsilon}{2}\rho_4^2 + \frac{\widetilde{\kappa}_\varepsilon}{2}\left(\rho_5^2+\rho_6^2\right)\right)$, and $\rho_4, \rho_5, \rho_6, \rho_7$ jointly converge in distribution to independent standard normal variables. Therefore, we have that

$$\begin{aligned}&\mathbb{E}_1(Y_n^\varepsilon)^2\\ \overset{n\to\infty}{\to}\ & \exp\left(-\kappa_\varepsilon^2/4-\kappa_\varepsilon/2\right)\exp\left(-\widetilde{\kappa}_\varepsilon^2/2-\widetilde{\kappa}_\varepsilon\right)\exp\left(-\frac{(a_\varepsilon-b_\varepsilon)^4(a-b)^2}{16(a+b)^4}-\frac{(a_\varepsilon-b_\varepsilon)^2(a-b)}{4(a+b)^2}\right)\\ &\times(1-\kappa_\varepsilon)^{-1/2}(1-\widetilde{\kappa}_\varepsilon)^{-1}\left(1-\frac{(a_\varepsilon-b_\varepsilon)^2(a-b)}{2(a+b)}\right)^{-1/2}.\end{aligned}$$

By (25) we get that

$$\begin{aligned}&\frac{\mathbb{E}_1(Y_n^\varepsilon)^2}{(\mathbb{E}_1Y_n^\varepsilon)^2}\\ \overset{n\to\infty}{\to}\ & \exp\left(-\kappa_\varepsilon^2/4-\kappa_\varepsilon/2\right)(1-\kappa_\varepsilon)^{-1/2}\\ &\times\exp\left(-\frac{(a_\varepsilon-b_\varepsilon)^4(a-b)^2}{16(a+b)^4}-\frac{(a_\varepsilon-b_\varepsilon)^2(a-b)}{4(a+b)^2}\right)\left(1-\frac{(a_\varepsilon-b_\varepsilon)^2(a-b)}{2(a+b)}\right)^{-1/2}\\ =\ & \exp\left(\sum_{m=3}^\infty\frac{1}{2m}\left(\frac{(a_\varepsilon-b_\varepsilon)^2}{2(a+b)}\right)^m\right)\times\exp\left(\sum_{m=3}^\infty\frac{1}{2m}\left(\frac{(a_\varepsilon-b_\varepsilon)^2(a-b)}{2(a+b)^2}\right)^m\right) \qquad (30)\\ =\ & \exp\left(\sum_{m=3}^\infty\lambda_m(1+\delta_m)(\delta_m^\varepsilon)^2\right),\end{aligned}$$

where (30) follows from the below trivial facts:

$$\exp\left(\sum_{m=3}^\infty\frac{1}{2m}\left(\frac{(a_\varepsilon-b_\varepsilon)^2}{2(a+b)}\right)^m\right)=\exp\left(-\kappa_\varepsilon^2/4-\kappa_\varepsilon/2\right)(1-\kappa_\varepsilon)^{-1/2}$$

$$\begin{aligned}\exp\left(\sum_{m=3}^\infty\frac{1}{2m}\left(\frac{(a_\varepsilon-b_\varepsilon)^2(a-b)}{2(a+b)^2}\right)^m\right)&=\exp\left(-\frac{(a_\varepsilon-b_\varepsilon)^4(a-b)^2}{16(a+b)^4}-\frac{(a_\varepsilon-b_\varepsilon)^2(a-b)}{4(a+b)^2}\right)\\ &\quad\times\left(1-\frac{(a_\varepsilon-b_\varepsilon)^2(a-b)}{2(a+b)}\right)^{-1/2}.\end{aligned}$$

This verifies Condition A4.

In the end, notice that by (25),

$$\begin{aligned}\mathbb{E}_1Y_n^\varepsilon \overset{n\to\infty}{\to}\ & \exp\left(-\widetilde{\kappa}_\varepsilon^2/4-\widetilde{\kappa}_\varepsilon/2\right)(1-\widetilde{\kappa}_\varepsilon)^{-1/2}\\ =\ & \exp\left(\sum_{m=3}^\infty\frac{1}{2m}\widetilde{\kappa}_\varepsilon^m\right)\\ =\ & \exp\left(\sum_{m=3}^\infty\frac{1}{2m}\left(\frac{(a_\varepsilon-b_\varepsilon)(a-b)}{2(a+b)}\right)^m\right)=\exp\left(\sum_{m=3}^\infty\lambda_m\delta_m\delta_m^\varepsilon\right).\end{aligned}$$

The it follows by Proposition A.1 that

$$Y_n^\varepsilon\overset{n\to\infty}{\to}\exp\left(\sum_{m=3}^\infty\lambda_m\delta_m\delta_m^\varepsilon\right)\prod_{m=3}^\infty(1+\delta_m^\varepsilon)^{Z_m^1}\exp\left(-\lambda_m(1+\delta_m)\delta_m^\varepsilon\right)=W_1^\varepsilon,$$

where Z_m^1 are independent Poisson variable with mean $\lambda_m(1+\delta_m)$. □

A.2. Proofs in Section 2.2

Before proofs, we need the following technical lemma.

Lemma A.3. *Suppose that $\{c_{ml}\}_{m,l=1}^\infty$ is a real sequence satisfying (1) $\lim_{M\to\infty}\lim_{l\to\infty}\sum_{m=M}^\infty c_{ml}^2=0$, and (2) for any $m\geq 1$, $\lim_{l\to\infty}c_{ml}=c_m$. Furthermore, for any $l\geq 1$, $\{N_{ml}\}_{m=1}^\infty$ are independent random variables of zero mean and unit variance, and for any $m\geq 1$, $N_{ml}\overset{d}{\to}N(0,1)$ as $l\to\infty$. Then, as $l\to\infty$, $\sum_{m=1}^\infty c_{ml}N_{ml}\overset{d}{\to}N(0,\sum_{m=1}^\infty c_m^2)$.*

Proof of Lemma A.3. Notice that c_m is a square summable sequence. To see this, note that for any $M < N$,

$$\sum_{m=M}^{N} c_m^2 = \lim_{l\to\infty} \sum_{m=M}^{N} c_{ml}^2 \le \lim_{l\to\infty} \sum_{m=M}^{\infty} c_{ml}^2,$$

and hence, taking $N \to \infty$ on the left side we have,

$$\sum_{m=M}^{\infty} c_m^2 \le \lim_{l\to\infty} \sum_{m=M}^{\infty} c_{ml}^2,$$

leading to $\lim_{M\to\infty} \sum_{m=M}^{\infty} c_m^2 \le \lim_{M\to\infty} \lim_{l\to\infty} \sum_{m=M}^{\infty} c_{ml}^2 = 0$; see (1). Hence $\sum_{m=1}^{\infty} c_m^2 < \infty$.

For arbitrary M and $\delta > 0$, define an event $\mathcal{E}_{Ml} = \{|\sum_{m=M}^{\infty} c_{ml} N_{ml}| < \delta\}$. Since $E|\sum_{m=M}^{\infty} c_{ml} N_{ml}|^2 = \sum_{m=M}^{\infty} c_{ml}^2$, by condition (1) we can choose l and M large so that $E|\sum_{m=M}^{\infty} c_{ml} N_{ml}|^2 \le \delta^3$, and so $P(\mathcal{E}_{Ml}) \ge 1 - \delta$ by Chebyshev inequality. By independence and asymptotic normality of N_{ml} for $1 \le m \le M-1$, and condition (2), one has $\sum_{m=1}^{M-1} c_{ml} N_{ml} \xrightarrow{d} N(0, \sum_{m=1}^{M-1} c_m^2)$ as $l \to \infty$. Define $T_l = \sum_{m=1}^{\infty} c_{ml} N_{ml}$. Hence, for any $z \in \mathbb{R}$,

$$\begin{aligned} P\left(T_l \le z\right) &\le P(T_l \le z, \mathcal{E}_{Ml}) + \delta \\ &\le P\left(\sum_{m=1}^{M-1} c_{ml} N_{ml} \le z + \delta\right) + \delta \overset{l\to\infty}{\to} \Phi\left(\frac{z+\delta}{\sqrt{\sum_{m=1}^{M-1} c_m^2}}\right) + \delta. \end{aligned}$$

Taking $\delta \to 0$ and $M \to \infty$ in the above, we have $\limsup_{l\to\infty} P(T_l \le z) \le \Phi\left(\frac{z}{\sqrt{\sum_{m=1}^{\infty} c_m^2}}\right)$. Likewise one can show that $\liminf_{l\to\infty} P(T_l \le z) \ge \Phi\left(\frac{z}{\sqrt{\sum_{m=1}^{\infty} c_m^2}}\right)$. Then we have $\lim_{l\to\infty} P(T_l \le z) = \Phi\left(\frac{z}{\sqrt{\sum_{m=1}^{\infty} c_m^2}}\right)$. Proof completed. □

Proof of Theorem 2.4. The proof follows by Lemma A.3. We will analyze the distributions of W_0^ε and W_1^ε. Define $\Delta_\varepsilon = \sum_{m=3}^{\infty} \lambda_m \left(\log(1+\delta_m^\varepsilon) - \delta_m^\varepsilon\right)$. Since, as $a + b \to \infty$,

$$\sqrt{\lambda_m}\log(1+\delta_m^\varepsilon) \to \sqrt{\frac{1}{2m}k_1^m},\ \sqrt{\lambda_m(1+\delta_m)}\log(1+\delta_m^\varepsilon) \to \sqrt{\frac{1}{2m}k_1^m},\ \lambda_m\delta_m\log(1+\delta_m^\varepsilon) \to \frac{1}{2m}k_2^m,$$

and

$$\frac{Z_m^0 - \lambda_m}{\sqrt{\lambda_m}} \xrightarrow{d} N(0,1),\ \frac{Z_m^1 - \lambda_m(1+\delta_m)}{\sqrt{\lambda_m(1+\delta_m)}} \xrightarrow{d} N(0,1).$$

Therefore, by Lemma A.3 we have, as $a + b \to \infty$,

$$\log W_0^\varepsilon - \Delta_\varepsilon = \sum_{m=3}^{\infty} \frac{Z_m^0 - \lambda_m}{\sqrt{\lambda_m}} \times \sqrt{\lambda_m}\log(1+\delta_m^\varepsilon) \xrightarrow{d} N(0, \sigma_1^2),$$

and

$$\begin{aligned} \log W_1^\varepsilon - \Delta_\varepsilon &= \sum_{m=3}^{\infty} \frac{Z_m^1 - \lambda_m(1+\delta_m)}{\sqrt{\lambda_m(1+\delta_m)}} \times \sqrt{\lambda_m(1+\delta_m)}\log(1+\delta_m^\varepsilon) + \sum_{m=3}^{\infty} \lambda_m\delta_m\log(1+\delta_m^\varepsilon) \\ &\xrightarrow{d} N(\sigma_2^2, \sigma_1^2). \end{aligned}$$

Therefore, as $a + b \to \infty$,

$$1 - \alpha = P(W_0^\varepsilon \le w_\alpha^\varepsilon) = P\left(\frac{\log W_0^\varepsilon - \Delta_\varepsilon}{\sigma_1} \le \frac{\log w_\alpha^\varepsilon - \Delta_\varepsilon}{\sigma_1}\right),$$

which implies $\frac{\log w_\alpha^\varepsilon - \Delta_\varepsilon}{\sigma_1} \to z_{1-\alpha}$, and hence,

$$P(a, b, \varepsilon) = P(W_1^\varepsilon \ge w_\alpha^\varepsilon) = P\left(\frac{\log W_1{}^\varepsilon - \Delta_\varepsilon}{\sigma_1} \ge \frac{\log w_\alpha^\varepsilon - \Delta_\varepsilon}{\sigma_1}\right) \to \Phi\left(\frac{\sigma_2^2}{\sigma_1} - z_{1-\alpha}\right).$$

Proof completed. □

A.3. Proofs in Section 2.3

Proof of Theorem 2.5. Observe that

$$Var\left(\frac{1}{M}\sum_{l=1}^{M} g_n^{\varepsilon}(\sigma[l])\bigg|A\right) = \frac{1}{M}\left[\mathbb{E}_\sigma\left\{g_n^{\varepsilon}(\sigma)^2\big|A\right\} - \mathbb{E}_\sigma\left\{g_n^{\varepsilon}(\sigma)\big|A\right\}^2\right] \le \frac{1}{M}\mathbb{E}_\sigma\left\{g_n^{\varepsilon}(\sigma)^2\big|A\right\},$$

where the variance is taken w.r.t. $\sigma[l]$'s conditional on A_{uv}'s. So it is sufficient to deal with $\mathbb{E}_{A,\sigma} g_n^{\varepsilon}(\sigma)^2$. First, assume H_0 holds. Then it holds that

$$\mathbb{E}_{A,\sigma} g_n^{\varepsilon}(\sigma)^2 = \mathbb{E}_\sigma \prod_{u<v}\left(\frac{p_{uv}^{\varepsilon}(\sigma)^2}{p_0} + \frac{q_{uv}^{\varepsilon}(\sigma)^2}{q_0}\right) = (1+o(1))(1+\gamma_n^{\varepsilon})^{\frac{n(n-1)}{2}},$$

where $\gamma^{\varepsilon} = \frac{\kappa_\varepsilon}{n} + \frac{(a_\varepsilon - b_\varepsilon)^2}{4n^2}$, $\kappa_\varepsilon = \frac{(a_\varepsilon - b_\varepsilon)^2}{2(a+b)}$, and the last equality holds due to the following trivial fact:

$$\frac{p_{uv}^{\varepsilon}(\sigma)^2}{p_0} + \frac{q_{uv}^{\varepsilon}(\sigma)^2}{q_0} = 1 + \gamma_n^{\varepsilon} + O(n^{-3}), \text{ uniformly for } \sigma \in \{\pm\}^n.$$

Obviously, $(1+\gamma_n^{\varepsilon})^{\frac{n(n-1)}{2}} = \exp\left(\frac{n\kappa_\varepsilon}{2} - \frac{\kappa_\varepsilon^2}{4} - \frac{\kappa_\varepsilon}{2} + \frac{(a_\varepsilon-b_\varepsilon)^2}{8}\right)$, hence, $\frac{1}{M}\sum_{l=1}^{M} g_n^{\varepsilon}(\sigma[l]) = Y_n^{\varepsilon} + o_P(1)$ if $M \gg \exp\left(\frac{n\kappa_\varepsilon}{2}\right)$.

Next assume H_1 holds. Let $N_{++} = \#\{(u,v): u<v, \sigma_u\sigma_v = +, \tau_u\tau_v = +\}$, $N_{+-} = \#\{(u,v): u<v, \sigma_u\sigma_v = +, \tau_u\tau_v = -\}$, $N_{++} = \#\{(u,v): u<v, \sigma_u\sigma_v = -, \tau_u\tau_v = +\}$, $N_{++} = \#\{(u,v): u<v, \sigma_u\sigma_v = -, \tau_u\tau_v = -\}$. Similar to the expressions of $s_{r\pm}$ for $r = 0, 1, 2$ in the proof of Theorem 2.2, one can derive that

$$N_{++} = \frac{n^2}{8} - \frac{n}{2} + \frac{n}{8}(\rho_1^2 + \rho_3^2 + \rho_5^2), N_{+-} = \frac{n^2}{8} + \frac{n}{8}(\rho_1^2 - \rho_3^2 - \rho_5^2),$$
$$N_{-+} = \frac{n^2}{8} - \frac{n}{8}(\rho_1^2 - \rho_3^2 + \rho_5^2), N_{--} = \frac{n^2}{8} - \frac{n}{8}(\rho_1^2 + \rho_3^2 - \rho_5^2).$$

Following (27), one can check that

$$\begin{aligned}
&\mathbb{E}_{A,\sigma} g_n^{\varepsilon}(\sigma)^2 \\
&= 4^{-n}\sum_{\sigma,\tau}\prod_{u<v}\left(\frac{1}{p_0^2}p_{uv}^{\varepsilon}(\sigma)^2 p_{uv}(\tau) + \frac{1}{q_0^2}q_{uv}^{\varepsilon}(\sigma)^2 q_{uv}(\tau)\right) \\
&= 4^{-n}\sum_{\sigma,\tau}\left(\frac{1}{p_0^2}\left(\frac{a_\varepsilon}{n}\right)^2\left(\frac{a}{n}\right) + \frac{1}{q_0^2}\left(1-\frac{a_\varepsilon}{n}\right)^2\left(1-\frac{a}{n}\right)\right)^{N_{++}} \\
&\quad\times\left(\frac{1}{p_0^2}\left(\frac{a_\varepsilon}{n}\right)^2\left(\frac{b}{n}\right) + \frac{1}{q_0^2}\left(1-\frac{a_\varepsilon}{n}\right)^2\left(1-\frac{b}{n}\right)\right)^{N_{+-}} \\
&\quad\times\left(\frac{1}{p_0^2}\left(\frac{b_\varepsilon}{n}\right)^2\left(\frac{a}{n}\right) + \frac{1}{q_0^2}\left(1-\frac{b_\varepsilon}{n}\right)^2\left(1-\frac{a}{n}\right)\right)^{N_{-+}} \\
&\quad\times\left(\frac{1}{p_0^2}\left(\frac{b_\varepsilon}{n}\right)^2\left(\frac{b}{n}\right) + \frac{1}{q_0^2}\left(1-\frac{b_\varepsilon}{n}\right)^2\left(1-\frac{b}{n}\right)\right)^{N_{--}} \\
&= (1+o(1))\mathbb{E}_{\sigma\tau}(1+\gamma_{2+})^{N_{++}}(1+\gamma_{2-})^{N_{+-}}(1+\gamma_{0+})^{N_{-+}}(1+\gamma_{0-})^{N_{--}}.
\end{aligned}$$

It follows from direct examinations that

$$(1+\gamma_{2+})^{\frac{n^2}{8}-\frac{n}{2}}(1+\gamma_{2-})^{\frac{n^2}{8}}(1+\gamma_{0+})^{\frac{n^2}{8}}(1+\gamma_{0-})^{\frac{n^2}{8}} \asymp \exp\left(\frac{n\kappa_\varepsilon}{2}\right),$$

and

$$\begin{aligned}
&\mathbb{E}_{\sigma\tau}(1+\gamma_{2+})^{\frac{n}{8}(\rho_1^2+\rho_3^2+\rho_5^2)}(1+\gamma_{2-})^{\frac{n}{8}(\rho_1^2-\rho_3^2-\rho_5^2)}(1+\gamma_{0+})^{-\frac{n}{8}(\rho_1^2-\rho_3^2+\rho_5^2)}(1+\gamma_{0-})^{-\frac{n}{8}(\rho_1^2+\rho_3^2-\rho_5^2)} \\
= \;&(1+o(1))\mathbb{E}_{\sigma\tau}\exp\left(\frac{(a_\varepsilon-b_\varepsilon)^2(a-b)}{4(a+b)^2}\rho_3^2 + \frac{(a_\varepsilon-b_\varepsilon)(a-b)}{2(a+b)}\rho_5^2\right) \\
\overset{n\to\infty}{\to}\;&\left(1-\frac{(a_\varepsilon-b_\varepsilon)^2(a-b)}{2(a+b)^2}\right)^{-1/2}\left(1-\frac{(a_\varepsilon-b_\varepsilon)(a-b)}{a+b}\right)^{-1/2}.
\end{aligned}$$

The last limit follows by condition $(a_\varepsilon - b_\varepsilon)(a-b) < a+b$ and asymptotic independent standard normality of ρ_3 and ρ_5. Hence, $\mathbb{E}_{A,\sigma} g_n^{\varepsilon}(\sigma)^2 \lesssim \exp\left(\frac{n\kappa_\varepsilon}{2}\right)$, leading to $\frac{1}{M}\sum_{l=1}^{M} g_n^{\varepsilon}(\sigma[l]) = Y_n^{\varepsilon} + o_P(1)$ if $M \gg \exp\left(\frac{n\kappa_\varepsilon}{2}\right)$. □

References

Abbe, E., 2018. Community detection and stochastic block models: recent developments. J. Mach. Learn. Res. 18, 1–86.
Abbe, E., Sandon, C., 2018. Proof of the achievability conjectures for the general stochastic block model. Comm. Pure Appl. Math. 71 (7), 1334–1406.
Amini, A., Chen, A., Bickel, P., 2013. Pseudo-likelihood methods for community detection in large sparse networks. Ann. Statist. 41 (4), 2097–2122.
Amini, A., Levina, E., 2018. On semidefinite relaxations for the block model. Ann. Statist. 46 (1), 149–179.
Banerjee, D., 2018. Contiguity and non-reconstruction results for planted partition models: the dense case. Electron. J. Probab. 23, 28.
Banerjee, D., Ma, Z., 2017. Optimal hypothesis testing for stochastic block models with growing degrees. arXiv:1705.05305.
Basak, A., Mukherjee, S., 2017. Universality of the mean-field for the Potts model. Probab. Theory Related Fields 168, 557–600.
Bickel, P.J., Chen, A., 2009. A nonparametric view of network models and Newman–Girvan and other modularities. Proc. Natl. Acad. Sci. USA 106, 21068–21073.
Bickel, P.J., Sarkar, P., 2016. Hypothesis testing for automated community detection in networks. J. R. Stat. Soc. Ser. B 78, 253–273.
Bollobás, B., 2001. Random Graphs, second ed. Cambridge University Press.
Decelle, A., Krzakala, F., Moore, C., Zdeborová, F., 2011. Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. Phys. Rev. E 84, 066–106.
Fosdick, B.K., Hoff, P.D., 2015. Testing and modeling dependencies between a network and nodal attributes. J. Amer. Statist. Assoc. 110, 1047–1056.
Gao, C., Lafferty, J., 2017a. Testing for global network structure using small subgraph statistics. https://arxiv.org/pdf/1710.00862.pdf.
Gao, C., Lafferty, J., 2017b. Testing network structure using relations between small subgraph probabilities. https://arxiv.org/pdf/1704.06742.pdf.
Janson, S., 1995. Random regular graphs: asymptotic distributions and contiguity. Combin. Probab. Comput. 4, 369–405.
Le, C., Levina, E., 2015. Estimating the number of communities by spectral methods. https://arxiv.org/pdf/1507.00827.pdf.
Lei, J., 2016. A goodness-of-fit test for stochastic block models. Ann. Statist. 44, 401–424.
Leskovec, J., Lang, K.L., Dasgupta, A., Mahoney, M.W., 2008. Statistical properties of community structure in large social and information networks. In: Proceeding of the 17th International Conference on World Wide Web. ACM, pp. 695–704.
Maugis, P-A.G., Priebe, C.E., Olhede, S.C., Wolfe, P.J., 2017. Statistical inference for network samples using subgraph counts. https://arxiv.org/pdf/1701.00505.pdf.
Montanari, A., Sen, S., 2016. Semidefinite programs on sparse random graphs and their application to community detection, In: STOC '16 Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing, pp. 814–827.
Mossel, E., Neeman, J., Sly, A., 2015. Reconstruction and estimation in the planted partition model. Probab. Theory Related Fields 162, 431–461.
Mossel, E., Neeman, J., Sly, A., 2017. A proof of the block model threshold conjecture. Combinatorica 1–44. http://dx.doi.org/10.1007/s00493-016-3238-8.
Neeman, J., Netrapalli, P., 2014. Non-reconstructability in the stochastic block model. https://arxiv.org/abs/1404.6304.
Newman, M.E.J., 2006. Finding community structure in networks using the eigenvectors of matrices. Phys. Rev. E 74 (3), 036104.
Sarkar, P., Bickel, P., 2015. Role of normalization in spectral clustering for stochastic blockmodels. Ann. Statist. 43 (3), 962–990.
Strogatz, S.H., 2001. Exploring complex networks. Nature 4106825, 268–276.
Weiss, P., 1907. Lhypothèse du champ moléculaire et la propriété ferromagnétique. J. Phys. Theor. Appl. 6 (1), 661–690.
Wormald, N.C., 1999. Models of random regular graphs. In: London Mathematical Society Lecture Note Series, Cambridge University Press, pp. 239–298.
Zhao, Y., Levina, E., Zhu, J., 2011. Community extraction for social networks. Proc. Natn. Acad. Sci. USA 108, 7321–7326.
Zhao, Y., Levina, E., Zhu, J., 2012. Consistency of community detection in networks under degree-corrected stochastic block models. Ann. Statist. 40, 2266–2292.