CRAMÉR'S MODERATE DEVIATIONS OF MODULARITY IN NETWORK

YU MIAO AND QING YIN

ABSTRACT. Complex networks play a crucial role in understanding physical, biological, social and technological systems. One of the most relevant features of graphs representing real systems is community structure. In this paper, for a specific partition of a given network, we prove the Cramér's moderate deviations of modularity for the partition when the size of the network gets large.

1. Introduction

The modern science of networks has brought significant advances to our understanding of complex systems. Existing networks often display a high level of local inhomogeneity, with high edge density within certain groups of nodes and low edge density between these groups. A relevant feature of networks is community structure. Detecting communities is of great importance in understanding, analyzing, and organizing networks, as well as in making informed decisions (see [1, 11, 16, 18]).

Many approaches have been proposed for detecting community structure in networks. For example, Fortunato [8] provided some striking examples of real networks with community structure. In order to distinguish meaningful structural changes from random fluctuations, Rosvall and Bergstrom [21] presented a solution to this problem by using bootstrap resampling accompanied by significance clustering. Lancichinetti et al. [12] described a measure aimed at quantifying the statistical significance of single communities. Zhang and Chen [26] introduced a statistical framework for modularity based network community detection. Under this framework, a hypothesis testing procedure is developed to determine the significance of an identified community structure. Ma and Barnett [14] proved that the largest eigenvalue and modularity are asymptotically uncorrelated, which suggests the need for inference directly on modularity itself when the network is large.

Li and Qi [13] proposed a way of evaluating the significance of any given partition by considering whether this particular partition can arise simply from randomness under the assumption that there is no underlying community structure in the network. They further described the global null hypothesis called free labeling. Under this null hypothesis, they derived the asymptotic distribution of modularity. Moreover, they

2020 *Mathematics Subject Classification.* 05C82, 60F05.

Key words and phrases. Modularity; Cramér's moderate deviations; Network.

This work is supported by National Natural Science Foundation of China (NSFC-11971154).

performed a simulation study to validate the asymptotic behavior and further used some well-known real network data for illustration. The significance of the partition is defined based on this asymptotic distribution, which can help assess its goodness. Two different partitions can also be compared statistically. Simulation studies and real data analyses are performed for illustration. The model for a specific partition of a given network is as follows.

We consider an undirected graph G consisting of n vertices $\{v_1, v_2, \cdots, v_n\}$ and m edges $\{e_1, e_2, \cdots, e_m\}$. Let A_{ij} be the number of edges between vertex v_i and vertex v_j, for $1 \le i, j \le n$, and $k_i(n)$ denote the degree of vertex v_i, which is the number of edges connected to vertex v_i (for convenience, we write k_i instead of $k_i(n)$). In this paper, we discuss a simple graph, for which A_{ij} is 0 or 1 if $i \ne j$ and $A_{ii} = 0$. Then it is easy to see that

$$k_i = \sum_{j=1}^{n} A_{ij} = \sum_{j=1}^{n} A_{ji}, \quad \text{for} \quad 1 \le i \le n$$

and

$$\sum_{i=1}^{n} k_i = 2m, \qquad \text{for} \quad 1 \le i \le n.$$

Let C denote a partition of network G (using the existing community detection method, see Fortunato [8]), i.e., each vertex v_i $(1 \le i \le n)$ is associated with a group label or color $c_i \in \{1, 2, \cdots, K\}$, where K is the total number of communities by the partition, and we denote $C = (c_1, c_2, \cdots, c_n)$. Newman [17] introduced the following modularity of the partition C,

$$Q_n(C) = \frac{1}{2m} \sum_{i,j} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta_{c_i, c_j} = \frac{1}{2m} \sum_{i,j} B_{ij} \delta_{c_i, c_j}, \tag{1.1}$$

where

$$\delta_{c_i, c_j} = \begin{cases} 1 & \text{if } \ c_i = c_j \\ 0 & \text{otherwise} \end{cases}$$

and

$$B_{ij} = A_{ij} - \frac{k_i k_j}{2m}, \qquad 1 \le i, j \le n. \tag{1.2}$$

It is not difficult to check that $-1 < Q_n(C) < 1$, and $Q_n(C)$ is the weighted sum of B_{ij} over all pairs of vertices i, j that fall in the same groups. It measures the extent to which vertices of the same type are connected to each other in a network.

For a given partition C of the network, we are interested in whether this partition could be obtained by randomly assigning colors to the vertices. The global null hypothesis H_0 is that the colors are assigned to vertices randomly, regardless of the structure of the network. The probability that a given vertex is labeled as group 1 is $p_1 = |Col(1)|/n$, where $Col(1)$ is the cardinality of the set of vertices with color 1; the probability is $p_2 = |Col(2)|/n$ for group 2, and so on. For any $1 \le k \le K$, it is easy to check that

$$p_1 + p_2 + \cdots + p_K = 1, \qquad p_k \ge 0.$$

The labeling of different vertices is assumed to be independent so H_0 is also called *free labeling*.

Assume that the partition $C = (c_1, c_2, \ldots, c_n)$ is a random vector, where $c_1, c_2, \ldots, c_n$ are independent identically distributed random variables, and have the following distribution

$$\mathbb{P}(c_1 = j) = p_j, \quad 1 \le j \le K.$$

Denote

$$p_{(l)} = \sum_{k=1}^{K} p_k^l, \quad \text{for} \quad l = 1, 2, \cdots \tag{1.3}$$

and

$$\bar{h}(c_i, c_j) = \delta_{c_i, c_j} - p_{c_i} - p_{c_j} + p_{(2)}, \quad 1 \le i \ne j \le n. \tag{1.4}$$

In this case, we denote $Q_n(C)$ by Q_n to avoid confusion. Li and Qi [13] proved the following asymptotic normality of Q_n under some conditions:

$$\frac{Q_n - \mu_n}{\sigma_n} \xrightarrow{d} N(0, 1), \tag{1.5}$$

where μ_n and σ_n^2 are given by

$$\mu_n = \mathbb{E}[Q_n] = -\frac{1 - p_{(2)}}{4m^2} \sum_{i=1}^{n} k_i^2, \tag{1.6}$$

$$\sigma_n^2 = Var(Q_n) = \frac{p_{(2)} + p_{(2)}^2 - 2p_{(3)}}{2m^2} \sum_{1 \le i \ne j \le n} B_{ij}^2 + \frac{p_{(3)} - p_{(2)}^2}{m^2} \sum_{i=1}^{n} B_{ii}^2. \tag{1.7}$$

Yin et al. [25] proved the moderate deviation principle of the modularity estimator for the specific partition of a given network. Miao and Yin [15] studied the Berry-Esseen bound and strong law of large numbers of modularity in network when the size of the network gets large.

Let us recall the development of the Cramér moderate deviations as follows. Let $(\eta_i)_{i \ge 1}$ be a sequence of independent and identically distributed centered real random variables satisfying Cramér's condition:

$$\mathbb{E} \exp\{c_0 |\eta_1|\} < \infty, \tag{1.8}$$

where c_0 is a positive constant. Denote $\mathbb{E}\eta_1^2 = \sigma^2$ and $S_n = \sum_{i=1}^n \eta_i$. Cramér [2] established an asymptotic expansion of the probabilities of moderate deviations for the partial sums, that is, for all $0 < x = o(n^{1/2})$,

$$\left| \ln \frac{\mathbb{P}\left(S_n/(\sigma\sqrt{n}) > x\right)}{1 - \Phi(x)} \right| = O\left(\frac{1 + x^3}{\sqrt{n}}\right), \quad \text{as} \quad n \to \infty, \tag{1.9}$$

where $\Phi(x)$ is the standard normal distribution function. Cramér's moderate deviations for sums of independent random variables have been studied by many authors, (see, for instance, Feller [7], Petrov [19] and [20], Sakhanenko [22], Saulis and Statulevičius [23] and Statulevičius [24]). Grama and Haeusler [10] developed a new approach for proving large deviation results for martingales based on a change of probability measure. It extends to the case of martingales the conjugate distribution technique due to Cramér. Fan et al. [4] gave an expansion of large deviation probabilities for martingales, which extends the classical result due to Cramér to the case of martingale differences satisfying the conditional Bernstein condition. Fan et al. [5] proved a Cramér moderate deviation expansion for martingales with differences having finite conditional moments of order $2+\rho$, $\rho \in (0,1]$, and finite one-sided conditional exponential moments. Fan [3] derived Cramér-type moderate deviations for stationary sequences of bounded random variables. Fan and Shao [6] extended the classical Cramér result to the cases of normalized martingales and standardized martingales, with martingale differences satisfying the conditional Bernstein condition.

Based on the above discussions, the main purpose of this paper is to establish Cramér's moderate deviations of modularity in network by using the martingale approximation and Cramér's moderate deviations for martingales from Fan [3] and Fan and Shao [6]. The paper is organized as follows, our main results are stated and discussed in Section 2. In Section 3, the preliminary lemmas are stated. Proofs of main results are obtained in Section 4. Throughout the paper, the symbol M denotes a positive constant which is not necessarily the same one in each appearance.

2. Main results

In order to obtain the main results, we need to introduce the following symbols:

$$\delta_n = \left(\frac{p_{(2)} + p_{(2)}^2 - 2p_{(3)}}{m}\right)^{1/2}, \tag{2.1}$$

$$\varepsilon_n = \frac{2\max_{1\le i\le n} k_i}{\sqrt{m(p_{(2)} + p_{(2)}^2 - 2p_{(3)})}}, \tag{2.2}$$

$$\eta_n = \frac{\sqrt{64e\max_{1\le i\le n} k_i}}{\sqrt{m^{1/2}(p_{(2)} + p_{(2)}^2 - 2p_{(3)})}} = \frac{\sqrt{32e}}{\sqrt{p_{(2)} + p_{(2)}^2 - 2p_{(3)}}}\sqrt{\varepsilon_n} \tag{2.3}$$

and

$$\gamma_n = \frac{4e\max_{1\le i\le n} k_i}{\sqrt{m(p_{(2)} + p_{(2)}^2 - 2p_{(3)})}} = 4e\varepsilon_n. \tag{2.4}$$

Theorem 2.1. *Assume that the degree sequence $\{k_i, 1 \le i \le n\}$ satisfy the following conditions: $\varepsilon_n = o(1)$ as $n \to \infty$,*

$$\eta_n \le \frac{1}{2} \quad \text{and} \quad \varepsilon_n \le \frac{1}{8e} \quad \text{for all } n. \tag{2.5}$$

Then it holds that for all $0 \le x = o(\eta_n^{-1})$,

$$\left|\ln\frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n} > x\right)}{1-\Phi(x)}\right|$$

$$\leq M\Big(x^3(\varepsilon_n+\eta_n)+x^2\gamma_n|\ln\gamma_n|+(1+x)(\varepsilon_n|\ln\varepsilon_n|+\eta_n|\ln\eta_n|+\gamma_n|\ln\gamma_n|)\Big)$$
$$\leq M\Big(x^3\eta_n+x^2\eta_n^2|\ln\eta_n|+(1+x)\eta_n|\ln\eta_n|\Big).$$

Moreover, the same result holds true when replacing $Q_n-\mu_n$ is replaced by μ_n-Q_n.

Remark 2.1. *For the degree sequence $\{k_i, 1\leq i\leq n\}$, we assume that there exists a sequence $\{l_n, n\geq 1\}$ such that*

$$d_1 l_n \leq k_i \leq d_2 l_n \quad \textit{for} \quad 1\leq i\leq n,$$

where

$$d_1=\frac{(p_{(2)}+p_{(2)}^2-2p_{(3)})^2}{65^4e^2} \quad \textit{and} \quad d_2=\frac{(p_{(2)}+p_{(2)}^2-2p_{(3)})^2}{64^4e^2}.$$

If $l_n\leq(\log n)^5$, then for any $n\geq 1$,

$$\eta_n=\frac{\sqrt{64e\max_{1\leq i\leq n}k_i}}{\sqrt{m^{1/2}(p_{(2)}+p_{(2)}^2-2p_{(3)})}}\leq\frac{1}{2}n^{-1/4}(\log n)^{5/4}\leq\frac{1}{2}.$$

Similarly, we have $\varepsilon_n\leq 1/(8e)$. Hence, (2.5) holds.

Remark 2.2. *Another example is that for every $1\leq i\leq n$, the degree k_i of vertex v_i can be controlled uniformly by some positive constant, i.e.,*

$$\max_{1\leq j\leq n}k_j\leq\frac{(p_{(2)}+p_{(2)}^2-2p_{(3)})^{1/2}}{16e},$$

then $\varepsilon_n<1/(8e)$. Moreover, if

$$\max_{1\leq j\leq n}k_j\leq\frac{p_{(2)}+p_{(2)}^2-2p_{(3)}}{256e},$$

then $\eta_n\leq 1/2$.

Remark 2.3. *For the network G considered in the present paper, since $A_{ii}=0$ and $A_{ij}\in\{0,1\}$ for $i\neq j$, then we have*

$$0\leq m\leq n(n-1)/2 \quad \textit{and} \quad 0\leq\max_{1\leq i\leq n}k_i\leq n-1.$$

Notice that, if $\max_{1\leq i\leq n}k_i=n\sqrt{(p_{(2)}+p_{(2)}^2-2p_{(3)})}/(16e)$, then $\varepsilon_n\geq 1/(8e)$. If $\max_{1\leq i\leq n}k_i=n(p_{(2)}+p_{(2)}^2-2p_{(3)})/(64e)$, then $\eta_n\geq\frac{1}{2}$. Hence, Theorem 2.1 does not hold.

From Theorem 2.1, we have the following result about the equivalence to the normal tail.

Corollary 2.1. *Under the conditions in Theorem 2.1, it holds that for all $0\leq x=o(\eta_n^{-1})$,*

$$\frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}>x\right)}{1-\Phi(x)}=1+o(1). \tag{2.6}$$

Moreover, the same result holds true when replacing $Q_n-\mu_n$ is replaced by μ_n-Q_n.

Corollary 2.2. *Under the conditions in Theorem 2.1, we have*

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}\le x\right)-\Phi(x)\right|\le M\eta_n|\ln\eta_n|.$$

Remark 2.4. *Assume that the degree sequence $\{k_i, 1\le i\le n\}$ satisfies the following condition:*

$$\frac{\max_{1\le i\le n}k_i}{\sqrt{m}}\le Mn^{-1/2},$$

then, we derive that

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}\le x\right)-\Phi(x)\right|\le Mn^{-1/4}\log n.$$

Next we give the Cramér moderate deviation for the normalized sums $\dfrac{Q_n-\mu_n}{\sigma_n}$.

Theorem 2.2. *Under the conditions in Theorem 2.1, it holds that for all $0\le x=o(\eta_n^{-1})$,*

$$\left|\ln\frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\sigma_n}>x\right)}{1-\Phi(x)}\right|\le M\Big(x^3\eta_n+x^2\eta_n^2|\ln\eta_n|+(1+x)\eta_n|\ln\eta_n|\Big) \tag{2.7}$$

and

$$\frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\sigma_n}>x\right)}{1-\Phi(x)}=1+o(1). \tag{2.8}$$

Moreover, the same results hold true when replacing $Q_n-\mu_n$ is replaced by μ_n-Q_n.

Corollary 2.3. *Under the conditions in Theorem 2.2, we have*

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left(\frac{Q_n-\mu_n}{\sigma_n}\le x\right)-\Phi(x)\right|\le M\eta_n|\ln\eta_n|.$$

3. Some important lemmas

Let $(\xi_i,\mathcal{F}_i)_{i=0,\dots,n}$ be a finite sequence of martingale differences, defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$, where $\xi_0=0$, $\{\emptyset,\Omega\}=\mathcal{F}_0\subseteq\ldots\subseteq\mathcal{F}_n\subseteq\mathcal{F}$ are increasing σ-fields and $(\xi_i)_{i=1,\dots,n}$ are allowed to depend on n. Set

$$X_0=0,\qquad X_k=\sum_{i=1}^k\xi_i,\quad k=1,\dots,n.$$

Then $(X_i,\mathcal{F}_i)_{i=0,\dots,n}$ is a martingale. Denote by $\langle X\rangle$ the quadratic characteristic of the martingale $X=(X_k,\mathcal{F}_k)_{k=0,\dots,n}$, that is

$$\langle X\rangle_0=0,\qquad \langle X\rangle_k=\sum_{i=1}^k\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1}),\quad k=1,\dots,n.$$

In the sequel we shall use the following conditions:

(i) There exists a number $\varepsilon_n \in (0, \frac{1}{2}]$ such that

$$|\mathbb{E}\left(\xi_i^k|\mathcal{F}_{i-1}\right)| \leq \frac{1}{2}k!\varepsilon_n^{k-2}\mathbb{E}\left(\xi_i^2|\mathcal{F}_{i-1}\right), \quad \text{for all } k \geq 2 \text{ and } 1 \leq i \leq n;$$

(ii) There exist a number $\eta_n \in (0, \frac{1}{2}]$ and a positive constant M such that for all $x > 0$,

$$\mathbb{P}\left(|\langle X\rangle_n - 1| \geq x\right) \leq M \exp\left\{-x\eta_n^{-2}\right\}.$$

In the proof of Theorem 2.1, we need the following Cramér's moderate deviations for martingales, which is a simple consequence of Theorem 2.2 of Fan and Shao [6].

Lemma 3.1. ([6, Remark 2.1]) *Assume that the conditions (i) and (ii) are satisfied. Then for all* $0 \leq x = o(\min\{\varepsilon_n^{-1}, \eta_n^{-1}\})$, *we have*

$$\left|\ln \frac{\mathbb{P}(X_n > x)}{1 - \Phi(x)}\right| \leq M\Big(x^3(\varepsilon_n + \eta_n) + (1 + x)\left(\eta_n|\ln \eta_n| + \varepsilon_n|\ln \varepsilon_n|\right)\Big).$$

Lemma 3.2. ([9, (2.16)]) *If* $\xi_1, \xi_2, \ldots, \xi_n$ *are centered and independent random variables, then for any* $p \geq 2$,

$$\mathbb{E}\left|\sum_{i=1}^n \xi_i\right|^p \leq 2^p(p-1)^{p/2}\mathbb{E}\left(\sum_{i=1}^n \xi_i^2\right)^{p/2}.$$

Lemma 3.3. *Under the conditions in Lemma 3.2, if there exist positive constants* $\{b_i, 1 \leq i \leq n\}$, *such that* $|\xi_i| \leq b_i$ *for all* $1 \leq i \leq n$, *then for any* $x > 0$, *we have*

$$\mathbb{P}\left(\left|\sum_{i=1}^n \xi_i\right| \geq x\right) \leq \exp\left\{2 - \frac{x}{2eF_n}\right\}, \tag{3.1}$$

where

$$F_n = \left(\sum_{i=1}^n b_i^2\right)^{1/2}.$$

Proof. From Lemma 3.2, we have

$$\mathbb{E}\left|\sum_{i=1}^n \xi_i\right|^p \leq 2^p(p-1)^{p/2}\mathbb{E}\left(\sum_{i=1}^n \xi_i^2\right)^{p/2} \leq 2^p p^p \left(\sum_{i=1}^n b_i^2\right)^{p/2} =: (2F_n)^p p^p.$$

Let $p = x/2eF_n$ for any $x > 4eF_n$. By using the Markov's inequality, we have

$$\mathbb{P}\left(\left|\sum_{i=1}^n \xi_i\right| \geq x\right) \leq \exp\left\{-\frac{x}{2eF_n}\right\}.$$

Moreover, for any $0 < x < 4eF_n$, we have

$$\mathbb{P}\left(\left|\sum_{i=1}^n \xi_i\right| \geq x\right) \leq \exp\left\{2 - \frac{x}{2eF_n}\right\}.$$

Hence, for any $x > 0$, (3.1) holds. □

Lemma 3.4. ([9, (2.18)]) *Let $X_1, X_2, \cdots, X_n$ be a sequence of independent identically distributed random variables, and for any $1 \le i, j \le n$, $h_{i,j}(u, v) : \mathbb{R}^2 \to \mathbb{R}$ be measurable and symmetric with respect to its arguments. For any $1 \le i \ne j \le n$, assume that $\mathbb{E}(h_{i,j}(X_i, X_j)|X_j) = 0$, $\mathbb{E}(h_{i,j}(X_i, X_j)|X_i) = 0$ and $\mathbb{E}|h_{i,j}(X_1, X_2)|^p < \infty$ for some $p \ge 2$. Then we have*

$$\mathbb{E}\left|\sum_{1\le i<j\le n} h_{i,j}(X_i, X_j)\right|^p \le 4^p p^p \mathbb{E}\left(\sum_{1\le i<j\le n} h_{i,j}^2(X_i, X_j)\right)^{p/2}.$$

Lemma 3.5. *Under the conditions in Lemma 3.4, if there exist positive constants $\{a_{i,j}, 1 \le i, j \le n\}$, such that $|h_{i,j}(u, v)| \le a_{i,j}$ for all $1 \le i, j \le n$ and all $u, v \in \mathbb{R}$, then for any $x > 0$, we have*

$$\mathbb{P}\left(\left|\sum_{1\le i<j\le n} h_{i,j}(X_i, X_j)\right| > x\right) \le \exp\left\{2 - \frac{x}{4eD_n}\right\}, \tag{3.2}$$

where

$$D_n = \left(\sum_{1\le i<j\le n} a_{i,j}^2\right)^{1/2}.$$

Proof. From Lemma 3.4, we have

$$\mathbb{E}\left|\sum_{1\le i<j\le n} h_{i,j}(X_i, X_j)\right|^p \le 4^p p^p \left(\sum_{1\le i<j\le n} a_{i,j}^2\right)^{p/2} =: (4D_n)^p p^p.$$

Let $p = x/(4eD_n)$ for any $x > 8eD_n$. Then, by the Markov's inequality, we have

$$\mathbb{P}\left(\left|\sum_{1\le i<j\le n} h_{i,j}(X_i, X_j)\right| > x\right) \le \exp\left\{-\frac{x}{4eD_n}\right\}.$$

Moreover, for any $0 < x < 8eD_n$, we have

$$\mathbb{P}\left(\left|\sum_{1\le i<j\le n} h_{i,j}(X_i, X_j)\right| > x\right) \le \exp\left\{2 - \frac{x}{4eD_n}\right\}.$$

Hence, for any $x > 0$, (3.2) holds. □

Lemma 3.6. ([13, Lemma 4]) *Define $o_{ij} = \sum_{l=1}^n A_{il}A_{jl}$ for $1 \le i, j \le n$. Then*

$$\sum_{1\le i,j\le n} o_{ij}^2 \le \max_{1\le j\le n} k_j \sum_{i=1}^n k_i^2.$$

4. Proofs of main results

Proof of Theorem 2.1. Since

$$\sum_{i=1}^n B_{ij} = \sum_{j=1}^n B_{ij} = 0,$$

we can deduce that

$$\sum_{1\le i\neq j\le n} B_{ij} = \sum_{1\le i,j\le n} B_{ij} - \sum_{i=1}^n B_{ii} = -\sum_{i=1}^n B_{ii}$$

and

$$\begin{aligned} Q_n &= \frac{1}{2m}\sum_{i=1}^n B_{ii} + \frac{1}{2m}\sum_{1\le i\neq j\le n} B_{ij}\bar{h}(c_i,c_j) \\ &\quad + \frac{1}{2m}\sum_{1\le i\neq j\le n} B_{ij}(p_{c_i}+p_{c_j}) - \frac{p_{(2)}}{2m}\sum_{1\le i\neq j\le n} B_{ij} \\ &= \frac{1+p_{(2)}}{2m}\sum_{i=1}^n B_{ii} + \frac{1}{m}\sum_{1\le i<j\le n} B_{ij}\bar{h}(c_i,c_j) - \frac{1}{m}\sum_{i=1}^n B_{ii}p_{c_i} \\ &= \frac{1-p_{(2)}}{2m}\sum_{i=1}^n B_{ii} + \frac{1}{m}\sum_{1\le i<j\le n} B_{ij}\bar{h}(c_i,c_j) - \frac{1}{m}\sum_{i=1}^n B_{ii}(p_{c_i}-p_{(2)}). \qquad (4.1) \end{aligned}$$

Combining (1.6), (2.1) and the above equation, it holds that

$$\begin{aligned} \frac{Q_n-\mu_n}{\delta_n} &= \frac{1}{m\delta_n}\sum_{1\le i<j\le n} B_{ij}\bar{h}(c_i,c_j) - \frac{1}{m\delta_n}\sum_{i=1}^n B_{ii}(p_{c_i}-p_{(2)}) \\ &= \frac{1}{m\delta_n}\sum_{1\le i<j\le n} A_{ij}\overline{h}(c_i,c_j) - \frac{1}{2m\delta_n}\sum_{1\le i<j\le n} \frac{k_ik_j}{m}\overline{h}(c_i,c_j) \\ &\quad + \frac{1}{2m\delta_n}\sum_{i=1}^n \frac{k_i^2}{m}(p_{c_i}-p_{(2)}). \qquad (4.2) \end{aligned}$$

Denote

$$T_n = \frac{1}{m\delta_n}\sum_{j=1}^n\sum_{i=1}^{j-1} A_{ij}\overline{h}(c_i,c_j) =: \frac{1}{m\delta_n}\sum_{j=1}^n z_{nj}. \qquad (4.3)$$

Let $\mathcal{F}_j = \sigma(c_1,c_2,\ldots,c_j)$ denote the σ-algebra generated by $\{c_1,c_2,\ldots,c_j\}$ for $1\le j\le n$. Since $\mathbb{E}(\bar{h}(c_i,c_j)|\mathcal{F}_{j-1}) = \mathbb{E}(\bar{h}(c_i,c_j)|c_i) = 0$ for any $1\le i<j\le n$, it is easy to see that

$$\mathbb{E}\left(z_{nj}|\mathcal{F}_{j-1}\right) = 0 \quad \text{for} \quad 2\le j\le n.$$

Therefore, for each $n\ge 2$, $\{z_{nj},\ j=2,\ldots,n\}$ forms a martingale difference with respect to $\{\mathcal{F}_j\}$. Then $(T_i,\mathcal{F}_i)_{i=0,\ldots,n}$ is a martingale and

$$\langle T\rangle_n = \frac{1}{m^2\delta_n^2}\sum_{j=1}^n \mathbb{E}\left(z_{nj}^2|\mathcal{F}_{j-1}\right).$$

Next, we will verify that T_n satisfies the conditions (i) and (ii) in Lemma 3.1. By using (2.1) and the fact $|\overline{h}(c_i,c_j)|\le 2$, then for $1\le j\le n$, we can derive

$$\frac{1}{m\delta_n}|z_{nj}| = \frac{1}{m\delta_n}\left|\sum_{i=1}^{j-1} A_{ij}\overline{h}(c_i,c_j)\right| \le \frac{2k_j}{m\delta_n} \le \frac{2\max_{1\le j\le n} k_j}{\sqrt{m(p_{(2)}+p_{(2)}^2-2p_{(3)})}} = \varepsilon_n.$$

Hence, from the condition (2.5), the condition (i) holds. Since $A_{ij} \in \{0,1\}$, we have $A_{ij}^2 = A_{ij}$. It is enough to get that

$$\begin{aligned}\frac{1}{m^2\delta_n^2}\sum_{j=1}^{n}\mathbb{E}\left(z_{nj}^2|\mathcal{F}_{j-1}\right) =&\frac{1}{m^2\delta_n^2}\sum_{j=1}^{n}\mathbb{E}\left(\left(\sum_{i=1}^{j-1}A_{ij}\overline{h}(c_i,c_j)\right)^2\Big|\mathcal{F}_{j-1}\right)\\ =&\frac{1}{m^2\delta_n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}A_{ij}\mathbb{E}\left(\left(\overline{h}(c_i,c_j)\right)^2|\mathcal{F}_{j-1}\right)\\ &+\frac{2}{m^2\delta_n^2}\sum_{1\le i<l\le n-1}\sum_{j=l+1}^{n}A_{ij}A_{lj}\mathbb{E}\left(\overline{h}(c_i,c_j)\overline{h}(c_l,c_j)|\mathcal{F}_{j-1}\right).\end{aligned}$$

For every $1 \le i < j \le n$, define

$$U_{ni} = \sum_{j=i+1}^{n}A_{ij}\mathbb{E}\left[\left(\overline{h}(c_i,c_j)\right)^2|\mathcal{F}_{j-1}\right], \tag{4.4}$$

then $\{U_{ni}, 1 \le i \le n-1, n \ge 2\}$ is an array of independent random variables with

$$|U_{ni}| \le 4\sum_{j=i+1}^{n}A_{ij}, \quad 1 \le i \le n-1$$

and

$$\begin{aligned}\frac{1}{m^2\delta_n^2}\sum_{i=1}^{n-1}\mathbb{E}U_{ni} =&\mathbb{E}\left[\frac{1}{m^2\delta_n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}A_{ij}\mathbb{E}\left(\left(\overline{h}(c_i,c_j)\right)^2|\mathcal{F}_{j-1}\right)\right]\\ =&\frac{1}{m^2\delta_n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}A_{ij}\mathbb{E}\left[\mathbb{E}\left(\left(\overline{h}(c_i,c_j)\right)^2|\mathcal{F}_{j-1}\right)\right]\\ =&\frac{1}{m^2\delta_n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}A_{ij}\mathbb{E}\left[\mathbb{E}\left(\left(\delta_{c_i,c_j}-p_{c_i}-p_{c_j}+p_{(2)}\right)^2|\mathcal{F}_{j-1}\right)\right]\\ =&\frac{1}{m^2\delta_n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}A_{ij}\mathbb{E}\Big[\mathbb{E}\Big(\left(\delta_{c_i,c_j}-p_{c_i}\right)^2+\left(p_{(2)}-p_{c_j}\right)^2\\ &+2\left(\delta_{c_i,c_j}-p_{c_i}\right)\left(p_{(2)}-p_{c_j}\right)|\mathcal{F}_{j-1}\Big)\Big]\\ =&\frac{1}{m^2\delta_n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}A_{ij}\mathbb{E}\Big[\mathbb{E}\Big(\delta_{c_i,c_j}^2+p_{c_i}^2-2\delta_{c_i,c_j}p_{c_i}+p_{(2)}^2+p_{c_j}^2-2p_{(2)}p_{c_j}\\ &+2\delta_{c_i,c_j}p_{(2)}-2\delta_{c_i,c_j}p_{c_j}-2p_{(2)}p_{c_i}+2p_{c_i}p_{c_j}|\mathcal{F}_{j-1}\Big)\Big]\\ =&\frac{1}{m^2\delta_n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}A_{ij}\mathbb{E}\left[p_{c_i}-3p_{c_i}^2+2p_{(2)}p_{c_i}-p_{(2)}^2+p_{(3)}\right]\\ =&\frac{1}{2m^2\delta_n^2}\left(p_{(2)}+p_{(2)}^2-2p_{(3)}\right)\sum_{1\le i,j\le n}A_{ij}\end{aligned}$$

$$=\frac{2m}{2m^2\delta_n^2}\left(p_{(2)}+p_{(2)}^2-2p_{(3)}\right)=1. \tag{4.5}$$

For $1\le i<l\le n$, denote

$$\overline{H}_{i,l}(c_i,c_l):=\sum_{j=l+1}^{n}A_{ij}A_{lj}\mathbb{E}\left(\overline{h}(c_i,c_j)\overline{h}(c_l,c_j)|\mathcal{F}_{j-1}\right), \tag{4.6}$$

then, by using the fact $|\overline{h}(c_i,c_j)|\le 2$, we can derive

$$\left|\overline{H}_{i,l}(c_i,c_l)\right|\le 4\sum_{j=l+1}^{n}A_{ij}A_{lj}.$$

Furthermore, for $1\le i<l<j\le n$, since

$$\begin{aligned}
&\mathbb{E}\left[\overline{h}(c_i,c_j)\overline{h}(c_l,c_j)\right]\\
=&\mathbb{E}\left[\mathbb{E}\left[\left(\delta_{c_i,c_j}-p_{c_i}-p_{c_j}+p_{(2)}\right)\left(\delta_{c_l,c_j}-p_{c_l}-p_{c_j}+p_{(2)}\right)|\mathcal{F}_{j-1}\right]\right]\\
=&\mathbb{E}\Big[\mathbb{E}\Big[\delta_{c_i,c_j}\delta_{c_l,c_j}-\delta_{c_i,c_j}p_{c_l}-\delta_{c_i,c_j}p_{c_j}+\delta_{c_i,c_j}p_{(2)}-\delta_{c_l,c_j}p_{c_i}+p_{c_i}p_{c_l}+p_{c_i}p_{c_j}-p_{c_i}p_{(2)}\\
&\quad-\delta_{c_l,c_j}p_{c_j}+p_{c_j}p_{c_l}+p_{c_j}^2-p_{c_j}p_{(2)}+\delta_{c_l,c_j}p_{(2)}-p_{(2)}p_{c_l}-p_{(2)}p_{c_j}+p_{(2)}^2|\mathcal{F}_{j-1}\Big]\Big]\\
=&\mathbb{E}\Big[\mathbb{E}\Big(\delta_{c_i,c_l}\frac{p_{c_i}+p_{c_l}}{2}-p_{c_i}p_{c_l}-p_{c_i}^2+p_{c_i}p_{(2)}-p_{c_i}p_{c_l}+p_{c_i}p_{c_l}+p_{c_i}p_{(2)}-p_{c_i}p_{(2)}\\
&\quad-p_{c_l}^2+p_{c_l}p_{(2)}+p_{(3)}-p_{(2)}^2+p_{c_l}p_{(2)}-p_{c_l}p_{(2)}-p_{(2)}^2+p_{(2)}^2|c_i\Big)\Big]\\
=&\mathbb{E}\left[\mathbb{E}\left(\delta_{c_i,c_l}\frac{p_{c_i}+p_{c_l}}{2}-p_{c_i}p_{c_l}-p_{c_i}^2+p_{c_i}p_{(2)}-p_{c_l}^2+p_{c_l}p_{(2)}+p_{(3)}-p_{(2)}^2|c_i\right)\right]=0,
\end{aligned} \tag{4.7}$$

then we have

$$\begin{aligned}
&\frac{2}{m^2\delta_n^2}\sum_{1\le i<l\le n-1}\mathbb{E}\overline{H}_{i,l}(c_i,c_l)\\
=&\frac{2}{m^2\delta_n^2}\sum_{1\le i<l\le n-1}\mathbb{E}\left[\sum_{j=l+1}^{n}A_{ij}A_{lj}\mathbb{E}\left(\overline{h}(c_i,c_j)\overline{h}(c_l,c_j)|\mathcal{F}_{j-1}\right)\right]\\
=&\frac{2}{m^2\delta_n^2}\sum_{1\le i<l\le n-1}\sum_{j=l+1}^{n}A_{ij}A_{lj}\mathbb{E}\left[\overline{h}(c_i,c_j)\overline{h}(c_l,c_j)\right]=0.
\end{aligned}$$

Therefore, it follows that

$$\begin{aligned}
&\mathbb{P}\left(|\langle T\rangle_n-1|\ge x\right)\\
=&\mathbb{P}\left(\left|\frac{1}{m^2\delta_n^2}\sum_{j=1}^{n}\mathbb{E}(z_{nj}^2|\mathcal{F}_{j-1})-1\right|\ge x\right)\\
\le&\mathbb{P}\left(\left|\frac{1}{m^2\delta_n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}A_{ij}\mathbb{E}\left(\left(\overline{h}(c_i,c_j)\right)^2|\mathcal{F}_{j-1}\right)-1\right|\ge\frac{x}{2}\right)
\end{aligned}$$

$$
\begin{aligned}
&+\mathbb{P}\left(\left|\frac{2}{m^2\delta_n^2}\sum_{1\le i<l\le n-1}\sum_{j=l+1}^{n}A_{ij}A_{lj}\mathbb{E}\left(\overline{h}(c_i,c_j)\overline{h}(c_l,c_j)|\mathcal{F}_{j-1}\right)\right|\ge\frac{x}{2}\right)\\
=&\,\mathbb{P}\left(\left|\frac{1}{m^2\delta_n^2}\sum_{i=1}^{n-1}(U_{ni}-\mathbb{E}U_{ni})\right|\ge\frac{x}{2}\right)+\mathbb{P}\left(\left|\frac{2}{m^2\delta_n^2}\sum_{1\le i<l\le n-1}\overline{H}_{i,l}(c_i,c_l)\right|\ge\frac{x}{2}\right).
\end{aligned}\tag{4.8}
$$

Firstly, consider the first part. By using the Lemma 3.3, for any $x>0$, we deduce that

$$
\begin{aligned}
&\mathbb{P}\left(\left|\frac{1}{m^2\delta_n^2}\sum_{i=1}^{n-1}(U_{ni}-\mathbb{E}U_{ni})\right|\ge\frac{x}{2}\right)\\
\le&\exp\left(2-\frac{xm^2\delta_n^2}{32e\left(\sum_{i=1}^{n-1}\left(\sum_{j=i+1}^{n}A_{ij}\right)^2\right)^{1/2}}\right)\\
\le&\exp\left(2-\frac{xm(p_{(2)}+p_{(2)}^2-2p_{(3)})}{32e\left(\sum_{i=1}^{n-1}\left(\sum_{j=i+1}^{n}A_{ij}\right)^2\right)^{1/2}}\right)\\
\le&\exp\left(2-\frac{x(p_{(2)}+p_{(2)}^2-2p_{(3)})\sqrt{m}}{32e\sqrt{\max_{1\le i\le n}k_i}}\right).
\end{aligned}\tag{4.9}
$$

Next, combining Lemma 3.5 and 3.6, for any $x>0$, we have

$$
\begin{aligned}
&\mathbb{P}\left(\left|\frac{2}{m^2\delta_n^2}\sum_{1\le i<l\le n-1}\overline{H}_{i,l}(c_i,c_l)\right|\ge\frac{x}{2}\right)\\
\le&\exp\left(2-\frac{xm^2\delta_n^2}{64e\left(\sum_{1\le i<l\le n-1}\left(\sum_{j=l+1}^{n}A_{ij}A_{lj}\right)^2\right)^{1/2}}\right)\\
\le&\exp\left(2-\frac{xm(p_{(2)}+p_{(2)}^2-2p_{(3)})}{64e\left(\sum_{1\le i<l\le n-1}\left(\sum_{j=l+1}^{n}A_{ij}A_{lj}\right)^2\right)^{1/2}}\right)\\
\le&\exp\left(2-\frac{x(p_{(2)}+p_{(2)}^2-2p_{(3)})\sqrt{m}}{64e\max_{1\le i\le n}k_i}\right).
\end{aligned}\tag{4.10}
$$

From (4.8), (4.9) and (4.10), we have

$$
\mathbb{P}\left(|\langle T\rangle_n-1|\ge x\right)\le M\exp\left(-\frac{x(p_{(2)}+p_{(2)}^2-2p_{(3)})\sqrt{m}}{64e\max_{1\le i\le n}k_i}\right).
$$

Let

$$\eta_n^2 = 64e \max_{1\le i\le n} k_i/(\sqrt{m}(p_{(2)} + p_{(2)}^2 - 2p_{(3)})),$$

then, for all $x > 0$, we have

$$\mathbb{P}\left(|\langle T\rangle_n - 1| \ge x\right) \le M \exp\left\{-x\eta_n^{-2}\right\}.$$

Hence, from Lemma 3.1, for all $0 \le x = o\left(\min\{\varepsilon_n^{-1}, \eta_n^{-1}\}\right) = o(\eta_n^{-1})$, we have

$$\left|\ln \frac{\mathbb{P}(T_n > x)}{1 - \Phi(x)}\right| \le M\Big(x^3(\varepsilon_n + \eta_n) + (1+x)\left(\eta_n|\ln \eta_n| + \varepsilon_n|\ln \varepsilon_n|\right)\Big). \tag{4.11}$$

By using (4.2), it holds that

$$\begin{aligned}
&\mathbb{P}\left(\frac{Q_n - \mu_n}{\delta_n} > x\right)\\
&= \mathbb{P}\left(\frac{1}{m\delta_n}\left(\sum_{1\le i<j\le n} A_{ij}\bar{h}(c_i, c_j) - \sum_{1\le i<j\le n} \frac{k_i k_j}{2m}\bar{h}(c_i, c_j) + \sum_{i=1}^n \frac{k_i^2}{2m}(p_{c_i} - p_{(2)})\right) > x\right)\\
&\le \mathbb{P}\Big(T_n > x(1 - \gamma_n|\ln \gamma_n|)\Big) + \mathbb{P}\left(-\frac{1}{m\delta_n}\sum_{1\le i<j\le n} \frac{k_i k_j}{2m}\bar{h}(c_i, c_j) > \frac{x}{2}\gamma_n|\ln \gamma_n|\right)\\
&\quad + \mathbb{P}\left(\frac{1}{m\delta_n}\sum_{i=1}^n \frac{k_i^2}{2m}(p_{c_i} - p_{(2)}) > \frac{x}{2}\gamma_n|\ln \gamma_n|\right)
\end{aligned} \tag{4.12}$$

and

$$\begin{aligned}
&\mathbb{P}\left(\frac{Q_n - \mu_n}{\delta_n} > x\right)\\
&= \mathbb{P}\left(\frac{1}{m\delta_n}\left(\sum_{1\le i<j\le n} A_{ij}\bar{h}(c_i, c_j) - \sum_{1\le i<j\le n} \frac{k_i k_j}{2m}\bar{h}(c_i, c_j) + \sum_{i=1}^n \frac{k_i^2}{2m}(p_{c_i} - p_{(2)})\right) > x\right)\\
&\ge \mathbb{P}\Big(T_n > x(1 + \gamma_n|\ln \gamma_n|)\Big) - \mathbb{P}\left(\frac{1}{m\delta_n}\sum_{1\le i<j\le n} \frac{k_i k_j}{2m}\bar{h}(c_i, c_j) > \frac{x}{2}\gamma_n|\ln \gamma_n|\right)\\
&\quad - \mathbb{P}\left(-\frac{1}{m\delta_n}\sum_{i=1}^n \frac{k_i^2}{2m}(p_{c_i} - p_{(2)}) > \frac{x}{2}\gamma_n|\ln \gamma_n|\right).
\end{aligned}$$

From Lemma 3.5 and

$$\gamma_n = \frac{4e \max_{1\le i\le n} k_i}{\sqrt{m(p_{(2)} + p_{(2)}^2 - 2p_{(3)})}}, \tag{4.13}$$

we deduce that for any $x > 0$,

$$\mathbb{P}\left(-\frac{1}{m\delta_n}\sum_{1\le i<j\le n} \frac{k_i k_j}{2m}\bar{h}(c_i, c_j) > \frac{x}{2}\gamma_n|\ln \gamma_n|\right)$$

$$
\begin{aligned}
&\le \mathbb{P}\left(\left|\frac{1}{m\delta_n}\sum_{1\le i<j\le n}\frac{k_ik_j}{2m}\bar{h}(c_i,c_j)\right| > \frac{x}{2}\gamma_n|\ln\gamma_n|\right)\\
&\le \exp\left\{2-\frac{m\delta_n\gamma_n|\ln\gamma_n|x}{4e\left(\sum_{1\le i<j\le n}\frac{k_i^2k_j^2}{m^2}\right)^{1/2}}\right\}\\
&\le \exp\left\{2-\frac{m\delta_n\gamma_n|\ln\gamma_n|x}{4e\max_{1\le i\le n}k_i}\right\}\\
&= \exp\left\{2-|\ln\gamma_n|x\right\}. \qquad (4.14)
\end{aligned}
$$

Moreover, for any $x>0$, by using Lemma 3.3 and (4.13), we have

$$
\begin{aligned}
&\mathbb{P}\left(\frac{1}{m\delta_n}\sum_{i=1}^{n}\frac{k_i^2}{2m}(p_{c_i}-p_{(2)}) > \frac{x}{2}\gamma_n|\ln\gamma_n|\right)\\
&\le \mathbb{P}\left(\left|\frac{1}{m\delta_n}\sum_{i=1}^{n}\frac{k_i^2}{2m}(p_{c_i}-p_{(2)})\right| > \frac{x}{2}\gamma_n|\ln\gamma_n|\right)\\
&\le \exp\left\{2-\frac{m^2\delta_n\gamma_n|\ln\gamma_n|x}{2e\left(\sum_{i=1}^{n}k_i^4\right)^{1/2}}\right\}\\
&\le \exp\left\{2-\frac{m^{3/2}\delta_n\gamma_n|\ln\gamma_n|x}{2e\max_{1\le i\le n}k_i^{3/2}}\right\}\\
&\le \exp\left\{2-\frac{2m^{1/2}|\ln\gamma_n|x}{\max_{1\le i\le n}k_i^{1/2}}\right\}\\
&\le \exp\left\{2-|\ln\gamma_n|x\right\}. \qquad (4.15)
\end{aligned}
$$

Notice that for all $x\ge 0$ and $|\tau|\le\frac{1}{2}$,

$$
\frac{1-\Phi(x+\tau)}{1-\Phi(x)} = \exp\left\{\theta\sqrt{2\pi}(1+x)|\tau|\right\} \qquad (4.16)
$$

and

$$
\frac{1}{\sqrt{2\pi}(1+x)}e^{-x^2/2} \le 1-\Phi(x) \le \frac{1}{\sqrt{\pi}(1+x)}e^{-x^2/2}, \qquad (4.17)
$$

where $|\theta|\le 1$. From (4.11)-(4.17), we deduce that for all $1\le x=o\left(\eta_n^{-1}\right)$,

$$
\begin{aligned}
&\frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}>x\right)}{1-\Phi(x)}\\
&\le \frac{\mathbb{P}\Big(T_n > x(1-\gamma_n|\ln\gamma_n|)\Big)}{1-\Phi(x)} + \frac{\mathbb{P}\left(-\frac{1}{m\delta_n}\sum_{1\le i<j\le n}\frac{k_ik_j}{2m}\bar{h}(c_i,c_j) > \frac{x}{2}\gamma_n|\ln\gamma_n|\right)}{1-\Phi(x)}\\
&\quad + \frac{\mathbb{P}\left(\frac{1}{m\delta_n}\sum_{i=1}^{n}\frac{k_i^2}{2m}(p_{c_i}-p_{(2)}) > \frac{x}{2}\gamma_n|\ln\gamma_n|\right)}{1-\Phi(x)}
\end{aligned}
$$

$$
\begin{aligned}
&\le \frac{\mathbb{P}\Big(T_n > x(1-\gamma_n|\ln\gamma_n|)\Big)}{1-\Phi\Big(x(1-\gamma_n|\ln\gamma_n|)\Big)} \cdot \frac{1-\Phi\Big(x(1-\gamma_n|\ln\gamma_n|)\Big)}{1-\Phi(x)} \\
&\quad + \frac{\mathbb{P}\left(-\frac{1}{m\delta_n}\sum_{1\le i<j\le n}\frac{k_ik_j}{2m}\bar{h}(c_i,c_j) > \frac{x}{2}\gamma_n|\ln\gamma_n|\right)}{1-\Phi(x)} \\
&\quad + \frac{\mathbb{P}\left(\frac{1}{m\delta_n}\sum_{i=1}^n\frac{k_i^2}{2m}(p_{c_i}-p_{(2)}) > \frac{x}{2}\gamma_n|\ln\gamma_n|\right)}{1-\Phi(x)} \\
&\le \exp\Big\{M\Big(x^3(\varepsilon_n+\eta_n)+(1+x)(\varepsilon_n|\ln\varepsilon_n|+\eta_n|\ln\eta_n|)\Big)\Big\} \times \frac{1-\Phi\Big(x(1-\gamma_n|\ln\gamma_n|)\Big)}{1-\Phi(x)} \\
&\quad + \frac{2e^2}{1-\Phi(x)}\exp\{-|\ln\gamma_n|x\} \\
&\le \exp\Big\{M\Big(x^3(\varepsilon_n+\eta_n)+(1+x)(\varepsilon_n|\ln\varepsilon_n|+\eta_n|\ln\eta_n|)\Big)\Big\} \\
&\quad \times \exp\Big\{\theta\sqrt{2\pi}(1+x)x\gamma_n|\ln\gamma_n|\Big\} + M(1+x)\exp\{x^2/2\}\exp\{-|\ln\gamma_n|x\} \\
&\le \exp\Big\{M\Big(x^3(\varepsilon_n+\eta_n)+(1+x)(\varepsilon_n|\ln\varepsilon_n|+\eta_n|\ln\eta_n|)\Big)\Big\} \times \exp\big\{Mx^2\gamma_n|\ln\gamma_n|\big\} \\
&\quad + M(1+x)\exp\left\{\frac{x^2}{2}-|\ln\gamma_n|x\right\} \\
&\le \exp\Big\{M\Big(x^3(\varepsilon_n+\eta_n)+x^2\gamma_n|\ln\gamma_n|+(1+x)(\varepsilon_n|\ln\varepsilon_n|+\eta_n|\ln\eta_n|)\Big)\Big\}, \qquad (4.18)
\end{aligned}
$$

where the last line follows by

$$
(1+x)\exp\left\{\frac{x^2}{2}-|\ln\gamma_n|x\right\} \le M\exp\left\{x^3\varepsilon_n\right\}.
$$

By an argument similar to the proof of the above equation, we deduce that for all $1 \le x = o(\eta_n^{-1})$,

$$
\frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n} > x\right)}{1-\Phi(x)} \ge \exp\Big\{-M\Big(x^3(\varepsilon_n+\eta_n)+x^2\gamma_n|\ln\gamma_n|+(1+x)(\varepsilon_n|\ln\varepsilon_n|+\eta_n|\ln\eta_n|)\Big)\Big\}. \qquad (4.19)
$$

Hence, combining (4.18) with (4.19), we have, for all $1 \le x = o(\eta_n^{-1})$,

$$
\begin{aligned}
&\left|\ln\frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n} > x\right)}{1-\Phi(x)}\right| \\
&\le M\Big(x^3(\varepsilon_n+\eta_n)+x^2\gamma_n|\ln\gamma_n|+(1+x)(\varepsilon_n|\ln\varepsilon_n|+\eta_n|\ln\eta_n|)\Big) \\
&\le M\Big(x^3\eta_n+x^2\eta_n^2|\ln\eta_n|+(1+x)\eta_n|\ln\eta_n|\Big)
\end{aligned}
$$

Next, we consider the case where $x \in [0,1]$. It follows that (4.11) holds also for $-z_{nj}$. Thus, from (4.11), (4.17) and the inequality $|e^x - 1| \le |x|e^{|x|}$, we have

$$\begin{aligned}
&\sup_{|x|\le 2}\Big|\mathbb{P}\left(T_n > x\right) - \left(1-\Phi(x)\right)\Big| \\
\le& \sup_{|x|\le 2}\Big(1-\Phi(x)\Big)\Big|e^{M\left(x^3(\varepsilon_n+\eta_n)+(1+x)(\eta_n|\ln\eta_n|+\varepsilon_n|\ln\varepsilon_n|)\right)} - 1\Big| \\
\le& \sup_{|x|\le 2} M\Big(1-\Phi(x)\Big)\Big(x^3(\varepsilon_n+\eta_n) + (1+x)\left(\eta_n|\ln\eta_n| + \varepsilon_n|\ln\varepsilon_n|\right)\Big) \\
\le& M\Big(\varepsilon_n + \eta_n + \eta_n|\ln\eta_n| + \varepsilon_n|\ln\varepsilon_n|\Big) \\
\le& M\Big(\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n|\Big).
\end{aligned}$$

By using (4.16), we have

$$\begin{aligned}
&\Big|\Big(1-\Phi(x+\gamma_n|\ln\gamma_n|)\Big) - \Big(1-\Phi(x)\Big)\Big| \\
=& \Big(1-\Phi(x)\Big)\left|\frac{\Big(1-\Phi(x+\gamma_n|\ln\gamma_n|)\Big)}{\Big(1-\Phi(x)\Big)} - 1\right| \\
\le& \frac{M}{1+x}\exp\{-x^2/2\}\Big|\exp\{M(1+x)\gamma_n|\ln\gamma_n|\} - 1\Big| \\
\le& \frac{M}{1+x}\exp\{-x^2/2\}(1+x)\gamma_n|\ln\gamma_n|\exp\{(1+x)\gamma_n|\ln\gamma_n|\} \\
\le& M\gamma_n|\ln\gamma_n|.
\end{aligned}$$

Hence, from the inequality $|e^x - 1| \le |x|e^{|x|}$, (4.14), (4.15) and the above inequality, we deduce for all $x \in [0,1]$,

$$\begin{aligned}
&\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n} > x\right) - \Big(1-\Phi(x)\Big) \\
\ge& \mathbb{P}\Big(T_n > x + \gamma_n|\ln\gamma_n|\Big) - \Big(1-\Phi(x)\Big) - \mathbb{P}\left(\frac{1}{m\delta_n}\sum_{1\le i<j\le n}\frac{k_ik_j}{2m}\bar{h}(c_i,c_j) > \frac{1}{2}\gamma_n|\ln\gamma_n|\right) \\
&- \mathbb{P}\left(-\frac{1}{m\delta_n}\sum_{i=1}^{n}\frac{k_i^2}{2m}(p_{c_i}-p_{(2)}) > \frac{1}{2}\gamma_n|\ln\gamma_n|\right) \\
\ge& \mathbb{P}\Big(T_n > x + \gamma_n|\ln\gamma_n|\Big) - \Big(1-\Phi(x+\gamma_n|\ln\gamma_n|)\Big) \\
&- \Big|\Big(1-\Phi(x+\gamma_n|\ln\gamma_n|)\Big) - \Big(1-\Phi(x)\Big)\Big| \\
&- \mathbb{P}\left(\frac{1}{m\delta_n}\sum_{1\le i<j\le n}\frac{k_ik_j}{2m}\bar{h}(c_i,c_j) > \frac{1}{2}\gamma_n|\ln\gamma_n|\right) \\
&- \mathbb{P}\left(-\frac{1}{m\delta_n}\sum_{i=1}^{n}\frac{k_i^2}{2m}(p_{c_i}-p_{(2)}) > \frac{1}{2}\gamma_n|\ln\gamma_n|\right)
\end{aligned}$$

$$
\begin{aligned}
&\geq -M\Big(\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n|\Big) - \Big|\Big(1-\Phi(x+\gamma_n|\ln\gamma_n|)\Big) - \Big(1-\Phi(x)\Big)\Big| \\
&\quad - \mathbb{P}\left(\frac{1}{m\delta_n}\sum_{1\leq i<j\leq n}\frac{k_ik_j}{2m}\bar{h}(c_i,c_j) > \frac{1}{2}\gamma_n|\ln\gamma_n|\right) \\
&\quad - \mathbb{P}\left(-\frac{1}{m\delta_n}\sum_{i=1}^{n}\frac{k_i^2}{2m}(p_{c_i}-p_{(2)}) > \frac{1}{2}\gamma_n|\ln\gamma_n|\right) \\
&\geq -M\Big(\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n| + \gamma_n|\ln\gamma_n|\Big).
\end{aligned}
$$

Similarly, we derive for all $x \in [0,1]$,

$$
\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n} \geq x\right) - \Big(1-\Phi(x)\Big) \leq M\Big(\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n| + \gamma_n|\ln\gamma_n|\Big).
$$

From (4.17), we have

$$
\begin{aligned}
\frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n} > x\right)}{1-\Phi(x)} - 1 &\leq \frac{M}{1-\Phi(x)}\Big(\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n| + \gamma_n|\ln\gamma_n|\Big) \\
&\leq M(1+x)\Big(\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n| + \gamma_n|\ln\gamma_n|\Big),
\end{aligned}
$$

which implies

$$
\begin{aligned}
\frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n} > x\right)}{1-\Phi(x)} &\leq 1 + M(1+x)\Big(\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n| + \gamma_n|\ln\gamma_n|\Big) \\
&\leq 1 + Mx^3(\varepsilon_n+\eta_n) + Mx^2\gamma_n|\ln\gamma_n| \\
&\qquad + M(1+x)\Big(\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n| + \gamma_n|\ln\gamma_n|\Big).
\end{aligned}
$$

Then for all $x \in [0,1]$, it holds that

$$
\begin{aligned}
&\ln\frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n} > x\right)}{1-\Phi(x)} \\
&\leq \ln\Big(1 + Mx^3(\varepsilon_n+\eta_n) + Mx^2\gamma_n|\ln\gamma_n| + M(1+x)(\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n| + \gamma_n|\ln\gamma_n|)\Big) \\
&\leq M\Big(x^3(\varepsilon_n+\eta_n) + x^2\gamma_n|\ln\gamma_n| + (1+x)(\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n| + \gamma_n|\ln\gamma_n|)\Big).
\end{aligned}
$$

Similarly, we derive for all $x \in [0,1]$,

$$
\begin{aligned}
&\ln\frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n} > x\right)}{1-\Phi(x)} \\
&\quad \geq -M\Big(x^3(\varepsilon_n+\eta_n) + x^2\gamma_n|\ln\gamma_n| + (1+x)(\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n| + \gamma_n|\ln\gamma_n|)\Big).
\end{aligned}
$$

Therefore, for all $x \in [0,1]$, we have

$$
\left|\ln\frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n} > x\right)}{1-\Phi(x)}\right|
$$

$$\leq M\Big(x^3(\varepsilon_n+\eta_n)+x^2\gamma_n|\ln\gamma_n|+(1+x)(\varepsilon_n|\ln\varepsilon_n|+\eta_n|\ln\eta_n|+\gamma_n|\ln\gamma_n|)\Big)$$
$$\leq M\Big(x^3\eta_n+x^2\eta_n^2|\ln\eta_n|+(1+x)\eta_n|\ln\eta_n|\Big),$$

which implies the desired result for all $0 \leq x = o(\eta_n^{-1})$. □

Proof of Corollary 2.1. By using Theorem 2.1, we have

$$\ln\frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}>x\right)}{1-\Phi(x)}\leq M\Big(x^3\eta_n+x^2\eta_n^2|\ln\eta_n|+(1+x)\eta_n|\ln\eta_n|\Big),$$

which implies that

$$\frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}>x\right)}{1-\Phi(x)}\leq \exp\Big\{M\Big(x^3\eta_n+x^2\eta_n^2|\ln\eta_n|+(1+x)\eta_n|\ln\eta_n|\Big)\Big\}.$$

Then (2.6) holds for all $0 \leq x = o(\eta_n^{-1})$. □

Proof of Corollary 2.2. Let $\kappa_n=\min\{\varepsilon_n^{-1/4},\eta_n^{-1/4},\gamma_n^{-1/4}\}=\eta_n^{-1/4}$, then we have

$$\begin{aligned}
&\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}\leq x\right)-\Phi(x)\right|\\
&\leq \sup_{|x|\leq\kappa_n}\left|\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}\leq x\right)-\Phi(x)\right|+\sup_{|x|>\kappa_n}\left|\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}\leq x\right)-\Phi(x)\right|\\
&\leq \sup_{|x|\leq\kappa_n}\left|\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}\leq x\right)-\Phi(x)\right|+\sup_{x<-\kappa_n}\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}\leq x\right)+\sup_{x<-\kappa_n}\Phi(x)\\
&\quad+\sup_{x>\kappa_n}\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}> x\right)+\sup_{x>\kappa_n}\Big(1-\Phi(x)\Big).
\end{aligned}\tag{4.20}$$

By Theorem 2.1, (4.17) and the inequality $|e^x-1|\leq|x|e^{|x|}$, we deduce that

$$\begin{aligned}
&\sup_{|x|\leq\kappa_n}\left|\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}\leq x\right)-\Phi(x)\right|\\
&\leq \sup_{0<x\leq\kappa_n}\left|\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}> x\right)-\Big(1-\Phi(x)\Big)\right|+\sup_{-\kappa_n\leq x<0}\left|\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}\leq x\right)-\Phi(x)\right|\\
&\leq \sup_{0<x\leq\kappa_n}M\Big(1-\Phi(x)\Big)\Big(x^3(\varepsilon_n+\eta_n)+x^2\gamma_n|\ln\gamma_n|+(1+x)(\varepsilon_n|\ln\varepsilon_n|+\eta_n|\ln\eta_n|\\
&\quad+\gamma_n|\ln\gamma_n|)\Big)+\sup_{-\kappa_n\leq x<0}M\Phi(x)\Big(|x|^3(\varepsilon_n+\eta_n)+x^2\gamma_n|\ln\gamma_n|\\
&\quad+(1+|x|)(\varepsilon_n|\ln\varepsilon_n|+\eta_n|\ln\eta_n|+\gamma_n|\ln\gamma_n|)\Big)\\
&\leq \sup_{0<x\leq\kappa_n}\frac{M}{1+x}e^{-x^2/2}\Big(x^3(\varepsilon_n+\eta_n)+x^2\gamma_n|\ln\gamma_n|+(1+x)(\varepsilon_n|\ln\varepsilon_n|+\eta_n|\ln\eta_n|
\end{aligned}$$

$$
\begin{aligned}
&+ \gamma_n|\ln\gamma_n|\Big)\Big) + \sup_{-\kappa_n\le x<0}\frac{M}{1-x}e^{-x^2/2}\Big(|x|^3(\varepsilon_n+\eta_n) + x^2\gamma_n|\ln\gamma_n| \\
&+ (1+|x|)(\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n| + \gamma_n|\ln\gamma_n|)\Big) \\
\le\ & \sup_{0<x\le\kappa_n}\left\{\frac{M}{x}\Big(x^3(\varepsilon_n+\eta_n) + x^2\gamma_n|\ln\gamma_n|\Big) + M(\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n| + \gamma_n|\ln\gamma_n|)\Big)\right\} \\
&+ \sup_{-\kappa_n\le x<0}\left\{\frac{M}{|x|}\Big(|x|^3(\varepsilon_n+\eta_n) + x^2\gamma_n|\ln\gamma_n|\Big) + M\Big(\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n| + \gamma_n|\ln\gamma_n|\Big)\right\} \\
\le\ & \sup_{0<x\le\kappa_n} M\Big(x^2(\varepsilon_n+\eta_n) + x\gamma_n|\ln\gamma_n| + (\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n| + \gamma_n|\ln\gamma_n|)\Big) \\
&+ \sup_{-\kappa_n\le x<0} M\Big(x^2(\varepsilon_n+\eta_n) + |x|\gamma_n|\ln\gamma_n| + (\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n| + \gamma_n|\ln\gamma_n|)\Big) \\
\le\ & M\Big(\varepsilon_n^{3/2} + \eta_n^{3/2} + \gamma_n^{3/4}|\ln\gamma_n| + \varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n| + \gamma_n|\ln\gamma_n|\Big) \\
\le\ & M\Big(\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n| + \gamma_n|\ln\gamma_n|\Big) \\
\le\ & M\eta_n|\ln\eta_n|. \qquad (4.21)
\end{aligned}
$$

In addition, from the inequality (4.17), it is easy to see that

$$
\begin{aligned}
\sup_{x>\kappa_n}\Big(1-\Phi(x)\Big) = \sup_{x<-\kappa_n}\Phi(x) =& \Phi(-\kappa_n) = 1-\Phi(\kappa_n) \\
&\le \frac{1}{\sqrt{\pi}(1+\kappa_n)}e^{-\kappa_n^2/2} \le M\eta_n|\ln\eta_n|.
\end{aligned} \qquad (4.22)
$$

Using (4.21) and (4.22), we have

$$
\begin{aligned}
\sup_{x<-\kappa_n}\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}\le x\right) &= \mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}\le -\kappa_n\right) \\
&\le M\eta_n|\ln\eta_n| + \Phi(-\kappa_n) \le M\eta_n|\ln\eta_n|. \qquad (4.23)
\end{aligned}
$$

Similarly, it holds that

$$
\sup_{x>\kappa_n}\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n}> x\right) \le M\eta_n|\ln\eta_n|. \qquad (4.24)
$$

Therefore, combining the inequalities (4.20)-(4.24) together, the desired result is obtained. □

Proof of Theorem 2.2. It follows from (1.2) that

$$
\begin{aligned}
\sum_{1\le i\ne j\le n} B_{ij}^2 &= \sum_{1\le i\ne j\le n}\left(A_{ij}^2 + \frac{k_i^2k_j^2}{4m^2} - 2A_{ij}\frac{k_ik_j}{2m}\right) \\
&= \sum_{1\le i\ne j\le n} A_{ij} + \sum_{1\le i\ne j\le n}\frac{k_i^2k_j^2}{4m^2} - \sum_{1\le i\ne j\le n} 2A_{ij}\frac{k_ik_j}{2m} \\
&= 2m + \sum_{1\le i\ne j\le n}\frac{k_i^2k_j^2}{4m^2} - \sum_{1\le i\ne j\le n} A_{ij}\frac{k_ik_j}{m}.
\end{aligned}
$$

Since $Var(p_{c_1} - p_{(2)}) = p_{(3)} - p_{(2)}^2$, from (1.7), (2.1) and the above equality, we deduce

$$
\begin{aligned}
\sigma_n^2 &= \frac{p_{(2)} + p_{(2)}^2 - 2p_{(3)}}{2m^2} \sum_{1\le i\ne j\le n} B_{ij}^2 + \frac{p_{(3)} - p_{(2)}^2}{m^2} \sum_{i=1}^n B_{ii}^2 \\
&\ge \frac{p_{(2)} + p_{(2)}^2 - 2p_{(3)}}{2m^2} \sum_{1\le i\ne j\le n} B_{ij}^2 \\
&= \frac{p_{(2)} + p_{(2)}^2 - 2p_{(3)}}{2m^2} \left(2m + \sum_{1\le i\ne j\le n} \frac{k_i^2 k_j^2}{4m^2} - \sum_{1\le i\ne j\le n} A_{ij} \frac{k_i k_j}{m} \right) \\
&\ge \frac{p_{(2)} + p_{(2)}^2 - 2p_{(3)}}{m} \left(1 - \sum_{1\le i\ne j\le n} A_{ij} \frac{k_i k_j}{2m^2} \right) \\
&\ge \frac{p_{(2)} + p_{(2)}^2 - 2p_{(3)}}{m} \left(1 - \frac{1}{2m^2} \sqrt{\sum_{1\le i\ne j\le n} A_{ij}^2} \sqrt{\sum_{1\le i\ne j\le n} k_i^2 k_j^2} \right) \\
&\ge \frac{p_{(2)} + p_{(2)}^2 - 2p_{(3)}}{m} \left(1 - \sqrt{\frac{\left(\sum_{i=1}^n k_i^2\right)^2}{2m^3}} \right) \\
&\ge \frac{p_{(2)} + p_{(2)}^2 - 2p_{(3)}}{m} \left(1 - \frac{\sqrt{2}\max_{1\le i\le n} k_i}{\sqrt{m}} \right) \\
&\ge \delta_n^2 \left(1 - \frac{\sqrt{2}\max_{1\le i\le n} k_i}{\sqrt{m}} \right). \qquad (4.25)
\end{aligned}
$$

Note that

$$
\begin{aligned}
\left| \ln \frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\sigma_n} > x\right)}{1-\Phi(x)} \right| &= \left| \ln \frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n} > x\frac{\sigma_n}{\delta_n}\right)}{1-\Phi(x)} \right| \\
&= \left| \ln \left(\frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n} > x\frac{\sigma_n}{\delta_n}\right)}{1-\Phi\left(x\frac{\sigma_n}{\delta_n}\right)} \cdot \frac{1-\Phi\left(x\frac{\sigma_n}{\delta_n}\right)}{1-\Phi(x)} \right) \right| \\
&\le \left| \ln \frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n} > x\frac{\sigma_n}{\delta_n}\right)}{1-\Phi\left(x\frac{\sigma_n}{\delta_n}\right)} \right| + \left| \ln \frac{1-\Phi\left(x\frac{\sigma_n}{\delta_n}\right)}{1-\Phi(x)} \right|. \qquad (4.26)
\end{aligned}
$$

From Theorem 2.1, (4.25) and (4.16), we have

$$
\begin{aligned}
&\left| \ln \frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\delta_n} > x\frac{\sigma_n}{\delta_n}\right)}{1-\Phi\left(x\frac{\sigma_n}{\delta_n}\right)} \right| \\
&\le M\Big(x^3(\varepsilon_n+\eta_n) + x^2\gamma_n|\ln\gamma_n| + (1+x)(\varepsilon_n|\ln\varepsilon_n| + \eta_n|\ln\eta_n| + \gamma_n|\ln\gamma_n|)\Big) \qquad (4.27)
\end{aligned}
$$

and

$$
\begin{aligned}
\left| \ln \frac{1-\Phi\left(x\frac{\sigma_n}{\delta_n}\right)}{1-\Phi(x)} \right| &\le \left| \ln \frac{1-\Phi\left(x\left(1-\frac{\sqrt{2}\max_{1\le i\le n}k_i}{\sqrt{m}}\right)\right)}{1-\Phi(x)} \right| \\
&= \left| \ln \exp\left\{ \theta\sqrt{2\pi}(1+x)x\frac{\sqrt{2}\max_{1\le i\le n}k_i}{\sqrt{m}} \right\} \right| \\
&\le M(1+x)x\varepsilon_n \le M(1+x)x\eta_n^2. \qquad (4.28)
\end{aligned}
$$

Hence, from (4.26), (4.27) and (4.28), we deduce that

$$
\left| \ln \frac{\mathbb{P}\left(\frac{Q_n-\mu_n}{\sigma_n} > x\right)}{1-\Phi(x)} \right| \le M\Big(x^3\eta_n + x^2\eta_n^2|\ln\eta_n| + (1+x)\eta_n|\ln\eta_n|\Big),
$$

which implies that (2.8) holds. □

Proof of Corollary 2.3. The proof of Corollary 2.3 is similar to the proof of Corollary 2.2, so the proof can be omitted. □

References

[1] R. Alert and A. Barabási, Statistical mechanics of complex networks. Rev. Modern Phys. 74 (2002), no. 1, 47-97.

[2] H. Cramér, Sur un nouveau théorème-limite de la théorie des probabilités. Actualite's Sci. Indust. 763 (1938), 5-23.

[3] X. Q. Fan, Cramér-type moderate deviations for stationary sequences of bounded random variables. C. R. Math. Acad. Sci. Paris 357 (2019), no. 5, 463-477.

[4] X. Q. Fan, I. Grama and Q. S. Liu, Cramér large deviation expansions for martingales under Bernstein's condition. Stochastic Process. Appl. 123 (2013), no. 11, 3919-3942.

[5] X. Q. Fan, I. Grama and Q. S. Liu, Cramér moderate deviation expansion for martingales with one-sided Sakhanenko's condition and its applications. J. Theoret. Probab. 33 (2020), no. 2, 749-787.

[6] X. Q. Fan and Q. M. Shao, Cramér's moderate deviations for martingales with applications. Ann. Inst. Henri Poincaré Probab. Stat. (accepted) 2023. (arXiv preprint arXiv:2204.02562, 2022.)

[7] W. Feller, Generalization of a probability limit theorem of Cramér. Trans. Amer. Math. Soc. 54 (1943), 361-372.

[8] S. Fortunato, Community detection in graphs. Phys. Rep. 486 (2010), no. 3-5, 75-174.

[9] E. Giné, R. Latał and J. Zinn, Exponential and moment inequalities for U-statistics. High dimensional probability, II (Seattle, WA, 1999), 13-38, Birkhäuser Boston, Boston, MA, 2000.

[10] I. Grama and E. Haeusler, Large deviations for martingales via Cramér's method. Stochastic Process. Appl. 85 (2000), no. 2, 279-293.

[11] M. O. Jackson, Social and economic networks. Princeton University Press, Princeton, NJ, 2008.

[12] A. Lancichinetti, F. Radicchi and J. J. Ramasco, Statistical significance of communities in networks. Physical Review E, 2010.

[13] Y. Li and Y. C. Qi, Asymptotic distribution of modularity in networks. Metrika 83 (2020), no. 4, 467-484.

[14] R. Ma and L. Barnett, The asymptotic distribution of modularity in weighted signed networks. Biometrika 108 (2021), no. 1, 1-16.

[15] Y. Miao and Q. Yin, Berry-Esseen bound of modularity in network. Submitted.

[16] M. E. J. Newman, The structure and function of complex networks. SIAM Rev. 45 (2003), no. 2, 167-256.

[17] M. E. J. Newman, Modularity and community structure in networks. Proc. Natl. Acad. Sci. USA 103 (2006), no. 23, 8577-8582.

[18] M. E. J. Newman, Networks: an introduction. Oxford University Press, Oxford, 2010.

[19] V. V. Petrov, A generalization of Cramér's limit theorem. Uspekhi Math. Nauk, 9 (1954), 195-202.

[20] V. V. Petrov, Sums of independent random variables. Springer-Verlag, New York-Heidelberg, 1975.

[21] M. Rosvall and C. T. Bergstrom, Mapping change in large networks. PLoS ONE, 2010.

[22] A. I. Sakhanenko, Convergence rate in the invariance principle for non-identically distributed variables with exponential moments. Springer, New York, 1985.

[23] L. Saulis and V. A. Statulevičius, Limit theorems for large deviations. Kluwer Academic Publishers, 1978.

[24] V. A. Statulevičius, On large deviations. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 6 (1966), 133-144.

[25] Q. Yin, Y. Miao, Z. Wang and G. Y. Yang, Moderate deviation principle of modularity in network. Metrika, 87 (2024), no. 3, 281-298.

[26] J. F. Zhang and Y. J. Chen, A hypothesis testing framework for modularity based network community detection. Statist. Sinica 27 (2017), no. 1, 437-456.

(Y. Miao) College of Mathematics and Information Science, Henan Normal University, Henan Province, 453007, China; Henan Engineering Laboratory for Big Data Statistical Analysis and Optimal Control, Henan Normal University, Henan Province, 453007, China.

Email address: yumiao728@gmail.com; yumiao728@126.com

(Q. Yin) College of Mathematics and Information Science, Henan Normal University, Henan Province, 453007, China.

Email address: qingyin1282@163.com