JUSTC

Article

http://justc.ustc.edu.cn

Received: February 10, 2025; Accepted: April 08, 2025

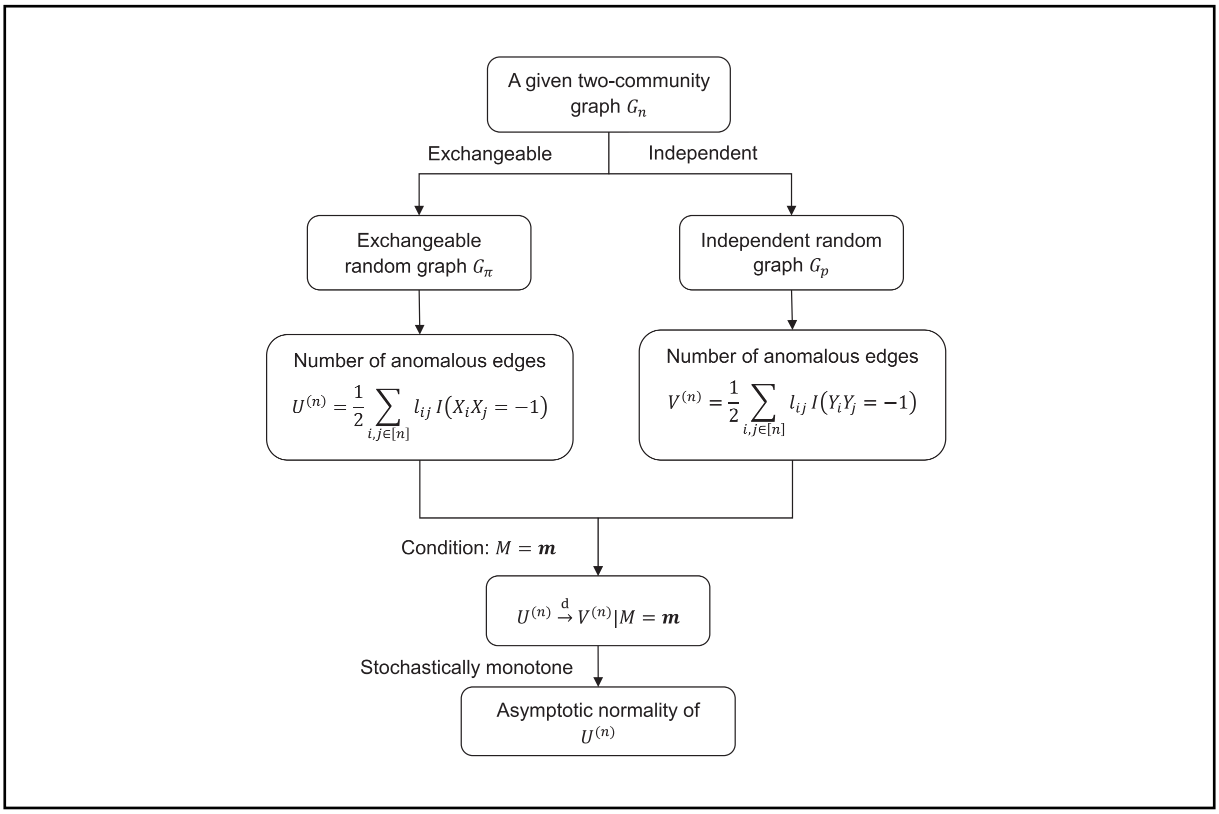
Limiting behavior of anomalous edges in exchangeable random graphs

Yaru Tian, and Qunqiang Feng ✉

Department of Statistics and Finance, School of Management, University of Science and Technology of China, Hefei 230026, China

✉Correspondence: Qunqiang Feng, E-mail: fengqq@ustc.edu.cn

Graphical abstract



The analysis of the asymptotic normality of the anomalous edges number.

Public summary

- Random network models with fixed topologies outperform those with free topologies in modeling real-world networks. In this paper, given a fixed graph, we propose a random graph framework in which vertices and edges are deterministic, but community labels are treated as random variables.
- As the graph size tends to infinity, we establish the asymptotic normality of the number of anomalous edges in our random graphs.
- This paper provides a rigorous criterion to assess whether detected communities are statistically distinguishable from random fluctuations. If the observed number of anomalous edges in a network significantly deviates from the asymptotic distribution predicted under exchangeability, it suggests the presence of non-random community structure.

Citation: Tian Y R, Feng Q Q. Limiting behavior of anomalous edges in exchangeable random graphs. *JUSTC*, **2025**, 55(0): . CSTR: 32290.14.JUSTC-2025-0013. DOI: 10.52396/JUSTC-2025-0013

Article
http://justc.ustc.edu.cn
Received: February 10, 2025; Accepted: April 08, 2025

Limiting behavior of anomalous edges in exchangeable random graphs

Yaru Tian, and Qunqiang Feng ✉

Department of Statistics and Finance, School of Management, University of Science and Technology of China, Hefei 230026, China

✉Correspondence: Qunqiang Feng, E-mail: fengqq@ustc.edu.cn

Cite This: *JUSTC*, **2025**, 55(X): (7pp)   Read Online

Abstract: Community structures are a common feature of many social networks. Anomalous edges are those edges that fall between distinct communities. According to a given network with two communities, random graphs are generated by assuming that the communities of vertices are exchangeable. In this paper, under some restriction conditions, the asymptotic normality of the number of anomalous edges in such random graphs is established as the graph size tends to infinity, where the method of transforming weak convergence into conditional weak convergence on the basis of stochastic monotonicity is utilized.

Keywords: network analysis; community structure; combinatorial central limit theorem; exchangeability; stochastically monotone

CLC number: O211.2   **Document code:** A

2020 Mathematics Subject Classification: 05C80; 60C05; 60F05

1 Introduction

The issue of community structure in social networks has been the focus of considerable attention for over twenty years[1–3]. There is no agreed-upon definition of a community in the network; the most widely accepted informal definition says that the vertices in networks are often found to cluster into tightly knit groups with a high density of within-group edges and a lower density of between-group edges[4]. A large number of real-world networks can be observed, where the individuals and the connections between them are abstracted by vertices and edges in the theoretical analysis, respectively. However, information on community structure is always absent. To address such problems, numerous techniques have been developed in the existing literature (see, for example, Refs. [5, 6] and references therein).

The stochastic block model (SBM)[6–10] is arguably the simplest statistical network model with communities, which has attracted much attention from network analysis and data science. Many statistical network analyses using SBM and its various variants assume that the size of each community is of the same order, which is beneficial for theoretical studies (see, e.g., Refs. [4, 11]). However, many researches have observed that the distributions of community sizes follow power laws in the large networks[4, 12, 13]. This means that community sizes can vary different, and most of them are very small[14, 15]. Moreover, SBM and its variants still ignore many other features of the real-world networks (see, e.g., Ref. [16]). For instance, many other network models have a common feature–exchangeability, which is enforced for models on graphs of a given size[17]. Using exchangeability, sparse graphs are studied in Ref. [18], which inspire a novel statistical model[19] to find the connections and reveal the structure in sparse graphs. In addition, Ref. [20] proposed an exchangeability-aware sum-product network model, which is based on explicit modeling of distributions over partially exchangeable.

Recent empirical studies[21, 22] demonstrate that random network models with fixed topologies outperform those with free topologies in modeling real-world networks. Here, fixed topology refers to models that preserve specific network properties, such as edge information, while allowing other attributes, such as vertex communities and edge signs, to be randomized. Building on this approach, we propose a random graph framework in which vertices and edges are deterministic, but community labels are treated as random variables. Specifically, our model generates new graphs by shuffling community assignments of vertices within an observed graph, while retaining its structural backbone. Assuming full exchangeability of all vertex community labels is often unrealistic. As highlighted in Ref. [23], social networks exhibit structural regularities—including degree distributions, degree assortativity, and homophily —that correlate with community organization. Thus, we introduce a more nuanced assumption: community labels of vertices with identical degrees are statistically exchangeable. This constraint aligns with observed patterns in network structure while maintaining computational and theoretical tractability.

The anomalous edges are edges that connect between distinct communities in a graph. When the vertices in a real-world network are clustered into different communities, the number of anomalous edges should be "small". To measure what is the appropriate level of smallness, our main concern here is to study the limiting behavior of such edges as the graph size tends to infinity.

The rest of the paper is organized as follows. In Section 2, we describe formally the generation of an exchangeable random graph on the basis of a given graph with communities, and give a definition for the anomalous edges in it. In Section 3, we show the asymptomatic normality of the number of such edges under some mild conditions.

2 The number of anomalous edges

We first introduce some necessary notations as follows. For any two vectors $\boldsymbol{a}=(a_1,a_2,\cdots,a_n)$ and $\boldsymbol{b}=(b_1,b_2,\cdots,b_n)$, let us denote $\boldsymbol{a}\le\boldsymbol{b}$ if $b_i-a_i\ge 0$ for all $1\le i\le n$. For any real x, let $\mathrm{sgn}(x)$ be the sign function. For a matrix A, denote by $A_{i,j}$ the (i,j)-entry of A. For any event A, let $I(A)$ be the indicator of A. We also write $\overset{\mathrm{d}}{=}$ for the distributional equality.

For simplicity, in this paper we consider a fixed simple, undirected graph $G_n=G(n,E,W)$ with only two communities, where $[n]:=\{1,2,\cdots,n\}$ is the vertex set, $E\subset[n]^2$ is the edge set, and $W\in\{-1,1\}^n$ is the vector of vertex labels representing the communities. More precisely, $W_i=1$ if vertex i is in the first community, and $W_i=-1$ otherwise. By convention, we say that vertex i is negative if $W_i=-1$. For any two distinct vertices $i,j\in[n]$, define $l_{ij}=1$ if there is an edge between i and j, and 0, otherwise. Then the degree of vertex $i\in[n]$ can be defined as

$$d_i=\sum_{j\ne i} l_{ij},$$

where we always assume that $l_{ii}=0$, i.e., self-loops are not allowed in G_n. We let $d_{\max}=\max_{i\in[n]} d_i$ denote the maximum degree of vertices, and $d_-=\max_{i\in[n]} d_i I(W_i=-1)$ the maximum degree of negative vertices. To avoid meaningless situations, we always assume that each vertex in G_n has degree greater than 1. We say that the edge between i and j is anomalous if it exists and $W_i\ne W_j$.

Recall that a sequence of random variables is said to be exchangeable, if their joint distribution is the same for any permutation of these random variables[24]. To associate a random graph with the fixed graph G_n, we now assume that the vertex labels are exchangeable random variables for the same degrees. More precisely, this random graph has the distribution $G_\pi=G_\pi(n,E,X)$, where the label vector X is uniformly distributed over a (disjoint) composition of uniform permutations, given by

$$X([n]):=(X(1)\circ\cdots\circ X(d_{\max}))([n]). \tag{1}$$

That is, $X(d)$ is a random vector uniformly distributed over permutations of the set of vertices with degree d, for any $d=1,2,\cdots,d_{\max}$.

Let $X_i\in\{-1,1\}$ be the vertex label of vertex i in a random graph with distribution G_π (for simplicity, we abuse the notation G_π for a random graph from now on). Then the number of anomalous edges in the random graph G_π can be expressed as

$$U^{(n)}=\frac{1}{2}\sum_{i,j\in[n]} l_{ij} I(X_iX_j=-1). \tag{2}$$

In the next section, we shall consider the asymptotic normality of $U^{(n)}$ under some mild conditions, as $n\to\infty$.

3 Asymptotic normality

In this section, we show the asymptotic normality for the number of anomalous edges in G_π under several regularity conditions.

Since $\{X_iX_j\}_{i\ne j}$ is a non-independent sequence, the asymptotic distribution of sums of permutation-based random variables is usually complicated to investigate. Fortunately, we show below that the vertex labels are distributed as independent Rademacher random variables when the size of each community is given. So the distribution of our graph is a conditional distribution of a simple graph with independence. To describe more precisely, we define a new random graph $G_p=G_p(n,E,Y)$ with the same support as G_n, where $Y=\{Y_1,\cdots,Y_n\}$ is an independent sequence with the probability mass function

$$\mathbb{P}(Y_i=-1)=1-\mathbb{P}(Y_i=1)=p_d,\quad i\in V_d,$$

where $V_d=\{i: i\in[n], d_i=d\}$ is a set of vertices with the same degree d, the success rate $p_d=\dfrac{m_d}{n_d}$ with n_d and m_d being the numbers of vertices with degree d and negative vertices with degree d, respectively. Similar to (2), the number of anomalous edges in G_p is

$$V^{(n)}=\frac{1}{2}\sum_{i,j\in[n]} l_{ij} I(Y_iY_j=-1).$$

The relation between random graphs G_π and G_p is given in the following.

Lemma 1. Let M_d be the number of negative vertices with degree d in G_π, vectors $M=\{M_d\}_{d=0}^{d_-}$ and $\mathbf{m}=\{m_d\}_{d=0}^{d_-}$. By conditioning on that $M=\mathbf{m}$, we have

$$U^{(n)}\overset{\mathrm{d}}{=}(V^{(n)}|M=\mathbf{m}).$$

Proof. For any degree d, let $\mathbf{y}=\{y_i\}_{i\in V_d}$ be a vector with entries $y_i\in\{-1,1\}$. Recalling that (1), $(X_i)_{i\in V_d}$ is a uniform random permutation on V_d, we have

$$\mathbb{P}((X_i)_{i\in V_d}=\mathbf{y})=\frac{1}{\binom{n_d}{m_d}}.$$

On the other hand, we have

$$\begin{aligned}\mathbb{P}((Y_i)_{i\in V_d}=\mathbf{y}|M_d=m_d)&=\frac{\mathbb{P}((Y_i)_{i\in V_d}=\mathbf{y},M_d=m_d)}{\mathbb{P}(M_d=m_d)}\\&=\frac{p_d^{m_d}(1-p_d)^{n_d-m_d}}{\binom{n_d}{m_d}p_d^{m_d}(1-p_d)^{n_d-m_d}}=\frac{1}{\binom{n_d}{m_d}}.\end{aligned}$$

Therefore, we have $\mathcal{L}(U_d^{(n)})=\mathcal{L}(V_d^{(n)}|M_d=m_d)$ for any degree d, where $U_d^{(n)}$ and $V_d^{(n)}$ are the number of anomalous edges incident to vertices with degree d in G_π and G_p, respectively. Since Y_i is independent for different degrees, we have $\mathcal{L}(U^{(n)})=\mathcal{L}(V^{(n)}|M=\mathbf{m})$, which completes the proof of

CSTR: 32290.14.JUSTC-2025-0013
DOI: 10.52396/JUSTC-2025-0013
JUSTC, **2025**, 55(X):

Lemma 1.

The above lemma provides a much easier way to analyze the distribution of the number of anomalous edges in G_π. However, in general, weak convergence does not lead to conditional weak convergence, so we need to introduce a restriction condition as follows.

Definition 1. Let V and M be random vectors in $\mathbb{R}^q$ and $\mathbb{R}^r$, respectively. If $\mathbb{P}(V \le v|M = \mathbf{m}_1) \ge \mathbb{P}(V \le v|M = \mathbf{m}_2)$ holds for any $v \in \mathbb{R}^q$ and $\mathbf{m}_1 \le \mathbf{m}_2$, we say that V is stochastically increasing with respect to M. In contrast, we say that V is stochastically decreasing with respect to M, if the inequality has a reverse direction, i.e.,

$$\mathbb{P}(V \le v|M = \mathbf{m}_1) \le \mathbb{P}(V \le v|M = \mathbf{m}_2).$$

Furthermore, we say that V is stochastically monotone if it is either stochastically increasing or decreasing.

Roughly speaking, the stochastic increasing property defined above captures the relationship wherein the elements of a random vector V exhibit a tendency to increase as the corresponding elements of another random vector V increase.

Let $T_\alpha^{(n)}$ be the number of edges with α negative endpoints in G_p for $\alpha = 0, 1, 2$. Then we have the following.

Lemma 2. The random vector $(T_2^{(n)}, T_1^{(n)} + T_2^{(n)})$ is stochastically increasing with respect to M, where M is the number of negative vertices in G_p.

Proof. Let $m = (m_0, \cdots, m_{d_-})$ and $\widetilde{m} = (\widetilde{m}_0, \cdots, \widetilde{m}_{d_-})$ be two vectors, where integers $0 \le m_d \le \widetilde{m}_d \le n-1$ for all $d = 0, \cdots, d_-$. Consider two graphs $G = (n, E, (X_i)_{i\in[n]})$ and $\widetilde{G} = (n, E, (\widetilde{X}_i)_{i\in[n]})$ with vertex labels X_i, $\widetilde{X}_i \in \{-1, 1\}$, such that $\sum_{i\in V_d} B_i = m_d$, $\sum_{i\in V_d} \widetilde{B}_i = \widetilde{m}_d$ for all $d = 0, \cdots, d_-$, where $B_i = \dfrac{1 - X_i}{2}$, $\widetilde{B}_i = \dfrac{1 - \widetilde{X}_i}{2}$. For the random permutation π and $r = 1, 2$, we further define

$$R_r = \sum_{i,j\in[n]} l_{ij} I\{B_{\pi(i)} + B_{\pi(j)} \ge r\}, \quad \tilde{R}_r = \sum_{i,j\in[n]} l_{ij} I\{\widetilde{B}_{\pi(i)} + \widetilde{B}_{\pi(j)} \ge r\}.$$

Then R_r and $\tilde{R}_r$ stand for the numbers of edges with at least r negative vertices in G and $\widetilde{G}$, respectively. It thus follows by Lemma 1 that

$$(R_1, R_2) \overset{d}{=} (T_2^{(n)}, T_1^{(n)} + T_2^{(n)}|M = m),$$

and

$$(\widetilde{R}_1, \widetilde{R}_2) = (T_2^{(n)}, T_1^{(n)} + T_2^{(n)}|\widetilde{M} = \widetilde{m}).$$

Note that R_r and $\tilde{R}_r$ are random variables defined on the same probability space. To prove the desired stochastic monotonicity, it is now sufficient to show $R_r \le \tilde{R}_r$. Moreover, in fact, we only need to consider the special case that $\tilde{m}_1 = m_1 + 1$ and $m_d = \tilde{m}_d$ for $d = 2, \cdots, d_-$. Since π restricted to V_d is a uniform random permutation, without loss of generality we can further assume that $X_{i_0} = -1$, $\widetilde{X}_{i_0} = 1$ for some $i_0 \in V_1$ and $X_i = \widetilde{X}_i$ otherwise. Then we now have that

$$l_{ij} I\{B_{\pi(i)} + B_{\pi(j)} \ge r\} \le l_{ij} I\{\widetilde{B}_{\pi(i)} + \widetilde{B}_{\pi(j)} \ge r\}. \tag{3}$$

In fact, if $\pi(i) \ne i_0$ and $\pi(j) \ne i_0$, it is clear that (3) holds. Otherwise, if either i or j is the same as i_0 (say $i = i_0$ and $j \ne i_0$), and the above inequality reduces to

$$l_{ij} I\{B_j \ge r\} \le l_{ij} I\{\widetilde{B}_j \ge r\},$$

which is also clearly true. Therefore, we have $R_r \le \widetilde{R}_r$ for $r = 1, 2$.

Although the number of anomalous edges is not stochastically monotone with respect to the number of negative vertices, by Lemma 2 we know that $(T_2^{(n)}, T_1^{(n)} + T_2^{(n)})$ satisfies the property of stochastic monotonicity with respect to M. Based on this property, we can now transform the weak convergence into the conditional weak convergence, see e.g., Refs. [25, 26].

To show the asymptotic behavior of the anomalous edges, we require some mild assumptions as follows.

Assumption 1. The degree of negative vertices is not larger than some given integer $d_- < \infty$.

Assumption 2. It holds that $\sup_{i\in[n]} d_i^2 = o(n)$.

Assumption 3. For $d_1, d_2, d_3 = 0, \cdots, d_-$, define $V_{d_1,d_2} = V_{d_1} \times V_{d_2}$. We assume that both partial sums

$$\mathcal{U}_{d_1,d_2} := \frac{1}{n} \sum_{(i,j)\in V_{d_1,d_2}} l_{ij}, \quad \mathcal{V}_{d_1,d_2,d_3} := \frac{1}{n} \sum_{i\in V_{d_1}} \Big(\sum_{j\in V_{d_2}} l_{ij}\Big)\Big(\sum_{j'\in V_{d_3}} l_{ij'}\Big)$$

converge to a limit as $n \to \infty$. Separately, we also require that n_d/n and $p_d = m_d/n_d$ converge as $n \to \infty$.

Assumption 1 implies that the number of negative vertices in G_n is bounded, or the connection to other vertices is very sparse. This constraint is critical for Theorem 2.2 of Ref. [26], which fails to hold if the dimension d_- of $M^{(n)}$ is unbounded. Assumption 2 further restricts all vertex degrees to a bounded range, ensuring that the graph exhibits sparse density. This assumption is commonly valid in the scale-free networks (see, e.g., Ref. [27, Chapter 8]). Assumption 3 is a technical condition. This condition guarantees the existence of limiting covariance terms of the statistic $T_\alpha^{(n)}$ for $\alpha = 0, 1, 2$, which depends on fixed structural components of the graph. For example, $\mathcal{U}_{d_1,d_2}$, denoting the number of edges between vertices with degrees d_1 and d_2, requires this assumption to ensure asymptotic convergence. Collectively, these assumptions balance generality with tractability, enabling rigorous analysis of anomalous edge distributions under sparsity constraints.

To state our first result, we need more notation as follows. For any $i \in [n]$, let

$$s_i = 4p_{d_i}(1 - p_{d_i}),$$

$$t_{1,1}(i) = s_i \sum_{j\in[n]} \frac{l_{ij}}{2}(2p_{d_j} - 1),$$

$$t_{2,1}(i) = -s_i \sum_{j\in[n]} \frac{l_{ij}}{2} p_{d_j}.$$

Let $h_d(i) = -\dfrac{s_i}{2} I\{i \in V_d\}$ for a fixed degree d. Define a $(d_{\max} + 3) \times (d_{\max} + 3)$ matrix $\Sigma^{(n)}$, where

CSTR: 32290.14.JUSTC-2025-0013
DOI: 10.52396/JUSTC-2025-0013
JUSTC, **2025**, 55(X):

$$\Sigma_{1,1}^{(n)} = \frac{1}{n}\Big(\sum_{i\in[n]} t_{1,1}^2(i) + \sum_{i,j\in[n]} \frac{l_{ij}}{16} s_i^2 s_j^2\Big),$$
$$\Sigma_{1,2}^{(n)} = \Sigma_{2,1}^{(n)} = \frac{1}{n}\Big(\sum_{i\in[n]} t_{1,1}(i)t_{2,1}(i) - \sum_{i,j\in[n]} \frac{l_{ij}}{32} s_i^2 s_j^2\Big),$$
$$\Sigma_{2,2}^{(n)} = \frac{1}{n}\Big(\sum_{i\in[n]} t_{2,1}^2(i) + \sum_{i,j\in[n]} \frac{l_{ij}}{64} s_i^2 s_j^2\Big),$$
$$\Sigma_{1,3+d}^{(n)} = \Sigma_{3+d,1}^{(n)} = \frac{1}{n}\sum_{i\in[n]} t_{1,1}(i)h_d(i), \tag{4}$$
$$\Sigma_{2,3+d}^{(n)} = \Sigma_{3+d,2}^{(n)} = \frac{1}{n}\sum_{i\in[n]} t_{2,1}(i)h_d(i),$$
$$\Sigma_{3+d,3+d}^{(n)} = \frac{1}{n}\sum_{i\in[n]} h_d^2(i),$$
$$\Sigma_{3+d,3+j}^{(n)} = 0,$$
$$d, j = 0,\cdots, d^-.$$

Define a random vector $T^{(n)} = (T_1^{(n)}, T_2^{(n)})$, where $T_\alpha^{(n)}$ is the number of edges with α negative vertices, $\alpha = 1,2$. Note that $V^{(n)} = T_1^{(n)}$. Then, for normalized vectors

$$\widetilde{T}^{(n)} := \frac{1}{\sqrt{n}}(T^{(n)} - \mathbb{E}T^{(n)}), \quad \widetilde{M}^{(n)} := \frac{1}{\sqrt{n}}(M^{(n)} - \mathbb{E}M^{(n)}),$$

we have the following asymptotic joint normality.

Theorem 1. If Assumptions 1, 2 and 3 hold, we have that as $n\to\infty$,

$$(\widetilde{T}^{(n)}, \widetilde{M}^{(n)}) \xrightarrow{\text{d}} N(0,\Sigma),$$

where the limiting covariance matrix $\Sigma = \lim_{n\to\infty}\Sigma^{(n)}$.

Proof. Our proof is divided into the following three steps. For simplicity, we omit the index n most of the time.

Step 1. Similarly to the Hoeffding decomposition, we begin by decomposing the centered versions of T_α's. Note that Y_i are independent for $i = 1,\cdots,n$, and

$$\mathbb{E}(Y_i) = r_i = 1 - 2p_{d_i}, \quad \mathrm{Var}(Y_i) = s_i^2 = 4p_{d_i}(1-p_{d_i}).$$

Write $\widetilde{Y}_i := \frac{Y_i - r_i}{s_i}$ and $Z_i := \frac{1-Y_i}{2}$. Note that Z_i is a Bernoulli random variable with success rate p_{d_i} for any negative vertex i. Then

$$T_1 = \sum_{i,j\in[n]} l_{ij} Z_i(1-Z_j) \quad \text{and} \quad T_2 = \frac{1}{2}\sum_{i,j\in[n]} l_{ij} Z_i Z_j$$

denote the number of edges with 1 and 2 negative vertices, respectively. For any $i,j\in[n]$, let

$$t_{1,2}(i,j) = -\frac{l_{ij}}{4} s_i s_j, \quad t_{2,2}(i,j) = \frac{l_{ij}}{8} s_i s_j.$$

We can now decompose

$$T_\alpha - \mathbb{E}(T_\alpha) = T_{\alpha,1} + T_{\alpha,2}, \tag{5}$$

where

$$T_{\alpha,1} = \sum_{i\in[n]} t_{\alpha,1}(i)\widetilde{Y}_i, \quad T_{\alpha,2} = \sum_{i,j\in[n]} t_{\alpha,2}(i,j)\widetilde{Y}_i\widetilde{Y}_j.$$

The proof of Eq. (5) is as follows. We first consider T_1. Define $Z_i' = 1 - Z_i$ for any $i\in[n]$. We thus have $Z_i' \sim \mathrm{Ber}(p'_{d_i})$ with $p'_{d_i} = 1 - p_{d_i}$. Note that

$$Z_i - p_{d_i} = \frac{1-Y_i}{2} - p_{d_i} = -\frac{1}{2}(Y_i - r_i) = -\frac{s_i}{2}\cdot\frac{Y_i - r_i}{s_i} = -\frac{1}{2}s_i\widetilde{Y}_i,$$

and similarly,

$$Z_i' - p'_{d_i} = 1 - Z_i - (1 - p_{d_i}) = \frac{1}{2}(Y_i - r_i) = \frac{s_i}{2}\cdot\frac{Y_i - r_i}{s_i} = \frac{1}{2}s_i\widetilde{Y}_i.$$

Therefore,

$$\begin{aligned}
T_1 - \mathbb{E}T_1 &= \sum_{i,j\in[n]} l_{ij}Z_i(1-Z_j) - \sum_{i,j\in[n]} l_{ij}p_{d_i}(1-p_{d_j}) = \\
&\sum_{i,j\in[n]} l_{ij}[(Z_i - p_{d_i})(Z_j' - p'_{d_j}) - 2p_{d_i}p'_{d_j} + Z_i p'_{d_j} + Z_j' p_{d_i}] = \\
&\sum_{i,j\in[n]} l_{ij}\Big(-\frac{1}{4}s_i s_j\widetilde{Y}_i\widetilde{Y}_j - \frac{1}{2}s_i p'_{d_j}\widetilde{Y}_i + \frac{1}{2}s_j p_{d_i}\widetilde{Y}_j\Big) = \\
&\sum_{i,j\in[n]} l_{ij}\Big[-\frac{1}{4}s_i s_j\widetilde{Y}_i\widetilde{Y}_j + \frac{1}{2}s_i(2p_{d_j} - 1)\widetilde{Y}_i\Big] = \\
&T_{1,1} + T_{1,2}.
\end{aligned}$$

Similarly, for $\alpha = 2$ we have

$$\begin{aligned}
T_2 - \mathbb{E}T_2 &= \sum_{i,j\in[n]} \frac{1}{2}l_{ij}Z_iZ_j - \sum_{i,j\in[n]} \frac{1}{2}l_{ij}p_{d_i}p_{d_j} = \\
&\frac{1}{2}\sum_{i,j\in[n]} l_{ij}[(Z_i - p_{d_i})(Z_j - p_{d_j}) - 2p_{d_i}p_{d_j} + Z_i p_{d_j} + Z_j p_{d_i}] = \\
&\frac{1}{2}\sum_{i,j\in[n]} l_{ij}\Big(\frac{1}{4}s_i s_j\widetilde{Y}_i\widetilde{Y}_j - \frac{1}{2}s_i p_{d_j}\widetilde{Y}_i - \frac{1}{2}s_j p_{d_i}\widetilde{Y}_j\Big) = \\
&\frac{1}{2}\sum_{i,j\in[n]} l_{ij}\Big(\frac{1}{4}s_i s_j\widetilde{Y}_i\widetilde{Y}_j - s_i p_{d_j}\widetilde{Y}_i\Big) = T_{2,1} + T_{2,2}.
\end{aligned}$$

Therefore, it holds (5).

Recall that M_d is the number of negative vertices with degree d, $d = 0,\cdots,d_-$. Define $\check{M}_d = M_d - \mathbb{E}M_d$, and

$$B^{(n)} := \frac{1}{\sqrt{n}}(T_{1,1}^{(n)}, T_{1,2}^{(n)}, T_{2,1}^{(n)}, T_{2,2}^{(n)}, \check{M}_0^{(n)}, \cdots, \check{M}_{d_-}^{(n)})^\top$$

with its covariance matrix $\Gamma^{(n)} \in \mathbb{R}^{(d_-+5)\times(d_-+5)}$.

Step 2. We show the asymptotic normality of $B^{(n)}$. Note that

$$\check{M}_d = M_d - \mathbb{E}M_d = \sum_{i\in[n]} I\{i\in V_d\}(Z_i - p_{d_i}) = \sum_{i\in[n]} h_d(i)\widetilde{Y}_i.$$

By independence of $\widetilde{Y}_i$, the variance of $\check{M}_d$ is

$$\mathrm{Var}(\check{M}_d) = \mathrm{Var}\Big(\sum_{i\in[n]} h_d(i)\widetilde{Y}_i\Big) = \frac{s_i^2}{4}\mathrm{Var}\Big(\sum_{i\in V_d}\widetilde{Y}_i\Big) = \sum_{i\in V_d}\frac{s_i^2}{4} = \sum_{i\in[n]} h_d^2(i). \tag{6}$$

Under Assumption 3, we have the three results as follows:
(i) there exists a matrix Γ such that

$$\lim_{n\to\infty}\Gamma^{(n)} = \Gamma. \tag{7}$$

(ii) for $\alpha,\beta = 1,2$ and $d = 0,\cdots,d_-$,

CSTR: 32290.14.JUSTC-2025-0013
DOI: 10.52396/JUSTC-2025-0013
JUSTC, **2025**, 55(X):

$$\sup_{i\in[n]}\frac{1}{n}h_d^2(i)\to 0,\quad \sup_{i\in[n]}\frac{1}{n}t_{\alpha,1}^2(i)\to 0,\quad \sup_{i\in[n]}\frac{1}{n}\sum_{j\in[n]}t_{\alpha,2}^2(i,j)\to 0. \tag{8}$$

(iii)

$$\mathbb{E}\Big(\frac{1}{\sqrt{n}}T_{\alpha,\beta}\Big)^4\to 3(\Gamma^2_{2(\alpha-1)+\beta,2(\alpha-1)+\beta},\quad \mathbb{E}\Big(\frac{1}{\sqrt{n}}\check{M}_d\Big)^4\to 3(\Gamma_{d+5,d+5})^2. \tag{9}$$

We now prove (7). For $\alpha,\beta=1,2$ and $d=0,\cdots,d_-$, by (5) and (6), we have the covariances

$$\frac{1}{n}\mathrm{Cov}(T_{\alpha,1},T_{\beta,1})=\frac{1}{n}\sum_{i\in[n]}t_{\alpha,1}(i)t_{\beta,1}(i), \tag{10}$$

$$\frac{1}{n}\mathrm{Cov}(T_{\alpha,2},T_{\beta,2})=\frac{1}{n}\sum_{i,j\in[n]}t_{\alpha,2}(i,j)t_{\beta,2}(i,j), \tag{11}$$

$$\frac{1}{n}\mathrm{Cov}(T_{\alpha,1},\check{M}_d)=\frac{1}{n}\sum_{i\in[n]}t_{\alpha,1}(i)h_d(i), \tag{12}$$

$$\frac{1}{n}\mathrm{Cov}(\check{M}_d,\check{M}_d)=\frac{1}{n}\sum_{i\in[n]}h_d^2(i), \tag{13}$$

$$\frac{1}{n}\mathrm{Cov}(T_{\alpha,1},T_{\beta,2})=\frac{1}{n}\mathrm{Cov}(T_{\alpha,2},\check{M}_d)=0, \tag{14}$$

while all other covariances are zero. By Assumption 3, we have that $\mathcal{U}_{d_1,d_2}$ and $\mathcal{V}_{d_1,d_2,d_3}$ are convergent for $d_1,d_2,d_3=0,\cdots,d_-$. We show that Eqs. (10)–(14) is convergent by Assumption 3 in the following. For Eq. (10), we have

$$t_{\alpha,1}(i)=s_i\sum_{j\in[n]}l_{ij}f_\alpha(p_{d_j})$$

for some functions f_α. Note that $s_i^2=0$ for $d_i=d_-+1,\cdots,d_{\max}$, the covariances satisfy

$$\frac{1}{n}\mathrm{Cov}(T_{\alpha,1},T_{\beta,1})=\frac{1}{n}\sum_{i\in[n]}s_i^2\sum_{j\in[n]}\frac{l_{ij}}{2}f_\alpha(p_{d_j})\sum_{j'\in[n]}\frac{l_{ij'}}{2}f_\beta(p_{d_{j'}})=$$
$$\sum_{d_1=0}^{d_-}p_{d_1}(1-p_{d_1})F(d_1),$$

where

$$F(d_1)=\frac{1}{n}\sum_{i\in V_{d_1}}\Bigg(\sum_{j\in[n]}l_{ij}f_\alpha\big(p_{d_j}\big)\Bigg)\Bigg(\sum_{j'\in[n]}l_{ij'}f_\alpha\big(p_{d_{j'}}\big)\Bigg)=$$
$$\frac{1}{n}\sum_{i\in V_{d_1}}\Bigg(\sum_{d_2=0}^{d_-}f_\alpha(p_{d_2})\sum_{j\in V_{d_2}}l_{ij}\Bigg)\Bigg(\sum_{d_3=0}^{d_-}f_\beta(p_{d_3})\sum_{j'\in V_{d_3}}l_{ij'}\Bigg)=$$
$$\frac{1}{n}\sum_{i\in V_{d_1}}\sum_{d_2=0}^{d_-}\sum_{d_3=0}^{d_-}f_\alpha(p_{d_2})f_\beta(p_{d_3})\sum_{j\in V_{d_2}}l_{ij}\sum_{j'\in V_{d_3}}l_{ij'}=$$
$$\sum_{d_2=0}^{d_-}\sum_{d_3=0}^{d_-}f_\alpha(p_{d_2})f_\beta(p_{d_3})\mathcal{V}_{d,d_2,d_3}.$$

With the above results, by Assumption 3 we can claim that $\frac{1}{n}\mathrm{Cov}(T_{\alpha,1},T_{\beta,1})$ is convergent as $n\to\infty$. Similarly, for some constant $c_{\alpha,\beta}$,

$$\frac{1}{n}\mathrm{Cov}(T_{\alpha,2},T_{\beta,2})=\frac{c_{\alpha,\beta}}{n}\sum_{i,j\in[n]}s_i^2s_j^2l_{ij}=$$
$$\frac{c_{\alpha,\beta}}{n}\sum_{d_1,d_2=0}^{d_-}4p_{d_1}(1-p_{d_1})4p_{d_2}(1-p_{d_2})\sum_{(i,j)\in V_{d_1,d_2}}l_{ij}=$$
$$c_{\alpha,\beta}\sum_{d_1,d_2=0}^{d_-}16p_{d_1}(1-p_{d_1})p_{d_2}(1-p_{d_2})\mathcal{U}_{d_1,d_2},$$

$$\frac{1}{n}\mathrm{Cov}(T_{\alpha,1},\check{M}_d)=\frac{1}{n}\sum_{i\in[n]}-\frac{s_i^2}{2}I\{i\in V_d\}\sum_{j\in[n]}l_{ij}f_\alpha(p_{d_j})=$$
$$\frac{1}{n}\sum_{i\in V_d}-2p_d(1-p_d)\sum_{d_2=0}^{d_-}f_\alpha(p_{d_2})\sum_{j\in V_{d_2}}l_{ij}=$$
$$-\sum_{d_2=0}^{d_-}2p_d(1-p_d)f_\alpha(p_{d_2})\mathcal{U}_{d,d_2},$$

and

$$\frac{1}{n}\mathrm{Cov}(\check{M}_d,\check{M}_d)=\frac{1}{n}\sum_{i\in[n]}\frac{1}{4}s_i^2I\{i\in V_d\}=\frac{1}{n}n_dp_d(1-p_d).$$

Combining Eqs. (10)–(14) with Assumption 3, we have that Eq. (7) holds.

Next, we prove (8). By Assumption 2, as $n\to\infty$, we have

$$\sup_{i\in[n]}\frac{1}{n}h_d^2(i)=\frac{1}{4n}I\{i\in V_d\}s_i^2=\frac{1}{n}p_d(1-p_d)\to 0,$$
$$\sup_{i\in[n]}\frac{1}{n}t_{\alpha,1}^2(i)=\sup_{i\in[n]}\frac{1}{n}\Big(s_i\sum_{j\in[n]}\frac{1}{2}l_{ij}f(p_{d_{(j)}})\Big)^2\le$$
$$\sup_{i\in[n]}\frac{1}{n}\Big(\sum_{j\in[n]}l_{ij}\Big)^2=\sup_{i\in[n]}\frac{1}{n}d_{(i)}^2\to 0,$$
$$\sup_{i\in[n]}\frac{1}{n}\sum_{j\in[n]}t_{\alpha,2}^2(i,j)=\sup_{i\in[n]}\frac{1}{n}\sum_{j\in[n]}(c_\alpha s_is_jl_{ij})^2\le\sup_{i\in[n]}\frac{1}{n}\sum_{j\in[n]}l_{ij}^2\le$$
$$\sup_{i\in[n]}\frac{1}{n}\Big(\sum_{j\in[n]}l_{ij}\Big)^2=\sup_{i\in[n]}\frac{1}{n}d_{(i)}^2\to 0.$$

For (9), let us first consider $T_{1,1}$. It follows that

$$\begin{aligned}
&\mathbb{E}\Big(\frac{1}{\sqrt{n}}T_{1,1}\Big)^4=\\
&\frac{1}{n^2}\sum_{i,j,k,l\in[n]}t_{1,1}(i)t_{1,1}(j)t_{1,1}(k)t_{1,1}(l)\mathbb{E}(\widetilde{Y}_i\widetilde{Y}_j\widetilde{Y}_k\widetilde{Y}_l)=\\
&\frac{1}{n^2}\sum_{i\in[n]}t_{1,1}^4(i)\mathbb{E}(\widetilde{Y}_i^4)+\frac{1}{n^2}\binom{4}{2}\sum_{1\le i\ne j\le n}t_{1,1}^2(i)t_{1,1}^2(j)\mathbb{E}(\widetilde{Y}_i^2)\mathbb{E}(\widetilde{Y}_j^2)=\\
&\frac{1}{n^2}\sum_{i\in[n]}t_{1,1}^4(i)\mathbb{E}(\widetilde{Y}_i^4-3)+\\
&\frac{3}{n^2}\Big(\sum_{1\le i\ne j\le n}2t_{1,1}^2(i)t_{1,1}^2(j)\mathbb{E}(\widetilde{Y}_i^2)\mathbb{E}(\widetilde{Y}_j^2)+\sum_{i\in[n]}t_{1,1}^4(i)\mathbb{E}(\widetilde{Y}_i^2)\Big)=\\
&\frac{1}{n^2}\sum_{i\in[n]}t_{1,1}^4(i)\mathbb{E}(\widetilde{Y}_i^4-3)+3\Big(\frac{1}{n}\sum_{i\in[n]}t_{1,1}^2(i)\Big)^2=\\
&\frac{1}{n^2}\sum_{i\in[n]}t_{1,1}^4(i)\mathbb{E}(\widetilde{Y}_i^4-3)+3(\Gamma_{1,1}^{(n)})^2.
\end{aligned} \tag{15}$$

CSTR: 32290.14.JUSTC-2025-0013
DOI: 10.52396/JUSTC-2025-0013
JUSTC, **2025**, 55(X):

Note that

$$\frac{1}{n^2}\sum_{i\in[n]} t_{1,1}^4(i) \le \frac{1}{n}\sup_{i\in[n]} t_{1,1}^2(i)\frac{1}{n}\sum_{i\in[n]} t_{1,1}^2(i) = \frac{1}{n}\sup_{i\in[n]} t_{1,1}^2(i)\Gamma_{1,1}^{(n)}.$$

Since it follows from the proof of (8) that $\frac{1}{n}\sup_{i\in[n]} t_{1,1}^2(i) \to 0$, the first term on the right-hand side of Eq. (15) converges to 0, which implies

$$\mathbb{E}\Big(\frac{1}{\sqrt{n}}T_{1,1}\Big)^4 \to 3(\Gamma_{1,1})^2.$$

Now we consider $T_{1,2}$. Similar to the discussion of $T_{1,1}$, we have that

$$\begin{aligned}\mathbb{E}\Big(\frac{1}{\sqrt{n}}T_{1,2}\Big)^4 = &\frac{1}{n^2}\sum_{i_k,j_k\in[n]} t_{1,2}(i_1,j_1)t_{1,2}(i_2,j_2)t_{1,2}(i_3,j_3)t_{1,4}(i_4,j_4)\times\\ &\mathbb{E}(\widetilde{Y}_{i_1}\widetilde{Y}_{j_1}\widetilde{Y}_{i_2}\widetilde{Y}_{j_2}\widetilde{Y}_{i_3}\widetilde{Y}_{j_3}\widetilde{Y}_{i_4}\widetilde{Y}_{j_4}) =\\ &\frac{1}{n^2}\sum_{i\in[n]} t_{1,2}^4(i,j)\mathbb{E}(\widetilde{Y}_i^4\widetilde{Y}_j^4)+\\ &\frac{\binom{4}{2}}{n^2}\sum_{1\le i_k\ne j_k\le n} t_{1,2}^2(i_1,j_1)t_{1,2}^2(i_2,j_2)\mathbb{E}(\widetilde{Y}_{i_1}^2\widetilde{Y}_{j_1}^2)\mathbb{E}(\widetilde{Y}_{i_2}^2\widetilde{Y}_{j_2}^2) =\\ &\frac{1}{n^2}\sum_{i\in[n]} t_{1,2}^4(i,j)\mathbb{E}(\widetilde{Y}_i^4\widetilde{Y}_j^4-3)+3(\Gamma_{2,2}^{(n)})^2 \to 3(\Gamma_{2,2})^2.\end{aligned}$$

Also in a similar way, we can further show

$$\mathbb{E}\Big(\frac{1}{\sqrt{n}}T_{2,1}\Big)^4 \to 3(\Gamma_{3,3})^2,\quad \mathbb{E}\Big(\frac{1}{\sqrt{n}}T_{2,2}\Big)^4 \to 3(\Gamma_{4,4})^2.$$

On the other hand, it follows that

$$\begin{aligned}\mathbb{E}\Big(\frac{1}{\sqrt{n}}\check{M}_d\Big)^4 = &\frac{1}{n^2}\mathbb{E}\Big(\sum_{i\in V_d} h(i)\widetilde{Y}_i\Big)^4 =\\ &\frac{1}{n^2}\sum_{i\in V_d} h^4(i)\mathbb{E}(\widetilde{Y}_i^4)+\\ &\frac{\binom{4}{2}}{n^2}\sum_{i,j\in V_d, i\ne j} h^2(i)h^2(j)\mathbb{E}(\widetilde{Y}_i^2)\mathbb{E}(\widetilde{Y}_j^2) =\\ &\frac{1}{n^2}\sum_{i\in V_d} h^4(i)\mathbb{E}(\widetilde{Y}_i^4-3)+3\Big(\frac{1}{n}\sum_{i\in V_d} h^2(i)\Big)^2 \to\\ &3(\Gamma_{d+5,d+5})^2,\end{aligned}$$

which completes the proof of (9). Combining (7)–(9) and Theorem 1.1 of Ref. [28], we thus have that as $n\to\infty$,

$$B^{(n)} \xrightarrow{\mathrm{d}} N(0,\Gamma).$$

Step 3. We show the joint asymptotic normality of $(\widetilde{T}^{(n)},\widetilde{M}^{(n)})$. Define a $(d_-+5)\times(d_-+5)$ matrix

$$Q = \begin{pmatrix} 1 & 1 & & & \\ & & 1 & 1 & \\ & & & & I_{d_-+1}\end{pmatrix},$$

where I_{d_-+1} is the identity matrix of size d_-+1. Note that $Q\Gamma Q^\top = \Sigma$. Then, we have

$$(\widetilde{T}^{(n)},\widetilde{M}^{(n)}) = QB^{(n)} \xrightarrow{\mathrm{d}} N(0,Q\Gamma Q^\top) = N(0,\Sigma),$$

as $n\to\infty$. The proof of Theorem 1 is complete.

Applying Theorem 1, we have the following result.

Theorem 2. If Assumptions 1, 2, and 3 hold, we have that

$$\frac{1}{\sqrt{n}}(U^{(n)}-\mathbb{E}(V^{(n)})) \xrightarrow{\mathrm{d}} N(0,\sigma^2),$$

where $\sigma^2 = \Sigma_{1,1}^s$ and $\Sigma^s = \Sigma_{\widetilde{T},\widetilde{T}} - \Sigma_{\widetilde{T},\widetilde{M}}\Sigma_{\widetilde{M},\widetilde{M}}^{-1}\Sigma_{\widetilde{T},\widetilde{M}}^\top$ with $\Sigma_{\widetilde{T},\widetilde{T}}$, $\Sigma_{\widetilde{T},\widetilde{M}}$, $\Sigma_{\widetilde{M},\widetilde{M}}$ being the covariance matrices of $\widetilde{T}$, $\widetilde{M}$ given in Theorem 1.

Proof. Note that the random vector $(\widetilde{T}^{(n)},\widetilde{M}^{(n)})$ has joint asymptotic normality according to Theorem 1, and $T^{(n)}$ has the property of stochastic monotonicity with respect to $M^{(n)}$ by Lemma 2. By Corollary 2.5 of Ref. [26], we thus have

$$\widetilde{S}^{(n)} := (\widetilde{T}^{(n)} \mid \widetilde{M}^{(n)} = 0) \xrightarrow{\mathrm{d}} (\widetilde{T} \mid \widetilde{M} = 0).$$

Since $(\widetilde{T},\tilde{M})$ is jointly normally distributed, and it is well known that the conditional distribution is also normal, we thus have

$$n^{-1/2}(U^{(n)}-\mathbb{E}(V^{(n)})) = \widetilde{S}_1^{(n)} \xrightarrow{\mathrm{d}} N(0,\sigma^2),$$

which completes the proof of Theorem 2.

4 Discussion

In this paper, we introduce a random graph model built on the principle of exchangeability among community labels of vertices with identical degrees. Under this framework, we establish the asymptotic normality of the number of anomalous edges–edges connecting vertices assigned to different communities–in large graphs, provided several mild technical assumptions (e.g., bounded degree sequences, sparse graph density). These results provide a foundational statistical characterization for graphs generated under our null model, where community labels are randomly permuted among vertices of the same degree.

The implications of this characterization are profound for community detection validation. Specifically, our null model enables hypothesis testing: if the observed number of anomalous edges in a network significantly deviates from the asymptotic distribution predicted under exchangeability, it suggests the presence of non-random community structure. Formally, for a given network, a community detection algorithm's result is statistically justified if the observed count of anomalous edges falls outside the $(1-\alpha)\times 100\%$ confidence interval derived from our asymptotic framework. This provides a rigorous criterion to assess whether detected communities are statistically distinguishable from random fluctuations.

A critical assumption in our analysis is the restriction to graphs with two communities. While the theoretical framework can, in principle, be extended to multi-community settings, the technical complexity escalates significantly. For instance, expressions for covariance terms and limiting distributions become intricate due to the combinatorial explosion in cross-community interactions (as shown in the proof of Theorem 1). Given page constraints, we defer this generalization to future work, although our current results establish a viable pathway for such extensions.

This study bridges the gap between exchangeable graph

CSTR: 32290.14.JUSTC-2025-0013
DOI: 10.52396/JUSTC-2025-0013
JUSTC, **2025**, 55(X):

models and asymptotic statistical theory, offering a principled approach to evaluate community detection outcomes. By quantifying the stochastic behavior of anomalous edges under structural randomness, our findings contribute to the methodological toolkit for analyzing complex networks, particularly in distinguishing meaningful community structures from null hypotheses of random organization.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (11771418).

Conflict of interest

The authors declare that they have no conflicts of interest.

Biographies

Yaru Tian is a postgraduate student at the University of Science and Technology of China. Her research mainly focuses on random network analysis.

Qunqiang Feng is currently an Associate Professor at the University of Science and Technology of China (USTC). He received his Ph.D. degree from USTC in 2006. His research mainly focuses on applied probability, random network models, and network data analysis.

References

[1] Newman M E J. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America*, **2006**, *103* (23): 8577–8582.

[2] Schuetz P, Caflisch A. Multistep greedy algorithm identifies community structure in real-world and computer-generated networks. *Physical Review E*, **2008**, *78*: 026112.

[3] Guo X, Qiu Y, Zhang H, et al. Randomized spectral co-clustering for large-scale directed networks. arXiv: 2004.12164v2, **2020**.

[4] Newman M E J. Fast algorithm for detecting community structure in networks. *Physical Review E*, **2004**, *69*: 066133.

[5] Sun H, Guyon I. Modularity in deep learning: A survey. In: Intelligent Computing. Cham, Switzerland: Springer Nature, **2023**: 561–595.

[6] Abbe E. Community detection and stochastic block models: Recent developments. *Journal of Machine Learning Research*, **2018**, *18* (177): 6446–6531.

[7] Holland P W, Laskey K B, Leinhardt S. Stochastic blockmodels: First steps. *Social Networks*, **1983**, *5* (2): 109–137.

[8] Rohe K, Chatterjee S, Yu B. Spectral clustering and the high-dimensional stochastic blockmodel. *The Annals of Statistics*, **2011**, *39* (4): 1878–1915.

[9] Abbe E, Bandeira A S, Hall G. Exact recovery in the stochastic block model. *IEEE Transactions on Information Theory*, **2016**, *62* (1): 471–487.

[10] Hu J, Zhang J, Qin H, et al. Using maximum entry-wise deviation to test the goodness of fit for stochastic block models. *Journal of the American Statistical Association*, **2021**, *116* (535): 1373–1382.

[11] Fortunato S, Hric D. Community detection in networks: A user guide. *Physics Reports*, **2016**, *659*: 1–44.

[12] Guimerà R, Danon L, Díaz-Guilera A, et al. Self-similar community structure in a network of human interactions. *Physical Review E*, **2003**, *68*: 065103.

[13] Berahmand K, Bouyer A. LP-LPA: A link influence-based label propagation algorithm for discovering community structures in networks. *International Journal of Modern Physics B*, **2018**, *32* (06): 1850062.

[14] Andersen R, Lang K J. Communities from seed sets. In: WWW ’06: Proceedings of the 15th International Conference on World Wide Web. New York: Association for Computing Machinery, **2006**: 223–232.

[15] Leskovec J, Lang K J, Dasgupta A, et al. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics*, **2009**, *6* (1): 29–123.

[16] Liu X, Song W, Musial K, et al. Stochastic block models for complex network analysis: A survey. *ACM Transactions on Knowledge Discovery from Data*, **2025**, *19* (3): No. 55.

[17] Lauritzen S, Rinaldo A, Sadeghi K. On exchangeability in network models. arXiv: 1709.03885, **2018**.

[18] Caron F, Fox E B. Sparse graphs using exchangeable random measures. *Journal of the Royal Statistical Society Series B*, **2017**, *79*: 1295–1366.

[19] Todeschini A, Miscouridou X, Caron F. Exchangeable random measures for sparse and modular graphs with overlapping communities. *Journal of the Royal Statistical Society Series B*, **2020**, *82*: 487–520.

[20] Lüdtke S, Bartelt C, Stuckenschmidt H. Exchangeability-aware sumproduct networks. arXiv: 2110.05165, **2022**.

[21] Gallo A, Garlaschelli D, Lambiotte R, et al. Testing structural balance theories in heterogeneous signed networks. *Communications Physics*, **2024**, *7* (1): 154.

[22] Gallo A, Garlaschelli D, Squartini T. Assessing frustration in real-world signed networks: A statistical theory of balance. *Physical Review Research*, **2024**, *6*: L042065.

[23] Apicella C L, Marlowe F W, Fowler J H, et al. Social networks and cooperation in hunter-gatherers. *Nature*, **2012**, *481* (7382): 497–501.

[24] De Finetti B. Funzione caratteristica di un fenomeno aleatorio. *Memorie della Accademia Nazionale dei Lincei, Classe di Scienze Fisiche, Matematiche e Naturali*, **1930**, *4*: 251–299.

[25] Nerman O. Stochastic monotonicity and conditioning in the limit. *Scandinavian Journal of Statistics*, **1998**, *25* (3): 569–572.

[26] Janson S. Monotonicity, asymptotic normality and vertex degrees in random graphs. *Bernoulli*, **2007**, *13* (4): 952–965.

[27] Van der Hofstad R. Random Graphs and Complex Networks. Cambridge, UK: Cambridge University Press, **2016**.

[28] Zheng G. A Peccati-Tudor type theorem for Rademacher chaoses. *ESAIM Probability and Statistics*, **2019**, *23*: 874–892.

CSTR: 32290.14.JUSTC-2025-0013
DOI: 10.52396/JUSTC-2025-0013
JUSTC, **2025**, 55(X):