**ORIGINAL PAPER**

# On the Randić index and its variants of network data

Mingao Yuan[1]

## Abstract

Summary statistics play an important role in network data analysis. They can provide us with meaningful insight into the structure of a network. The Randić index is one of the most popular network statistics that has been widely used for quantifying information of biological networks, chemical networks, pharmacologic networks, etc. A topic of current interest is to find bounds or limits of the Randić index and its variants. A number of bounds of the indices are available in literature. Recently, there are several attempts to study the limits of the indices in the Erdős–Rényi random graph by simulation. In this paper, we shall derive the limits of the Randić index and its variants of an inhomogeneous Erdős–Rényi random graph. Our results charaterize how network heterogeneity affects the indices and provide new insights about the Randić index and its variants. Finally we apply the indices to several real-world networks.

## 1 Introduction

A network (graph) consists of a set of agents and a set of pairwise interactions among the agents. Networks are canonical models that capture relations within or between data

✉ Mingao Yuan
  mingao.yuan@ndsu.edu

[1] Department of Statistics, North Dakota State University, Fargo, USA

sets. Due to the increasing popularity of relational data, network data analysis has been a primary research topic in statistics, machine learning and many other scientific fields (Abbe 2018; Bickel and Sarkar 2016; Goldenberg et al. 2010; Kolaczyk 2009; Newman 2009). One of the fundamental problems in network data analysis is to understand the structural properties of a given network. The structure of a small network can be easily described by its visualization. However, larger networks can be difficult to envision and describe. It is thus important to have several summary statistics that provide us with meaningful insight into the structure of a network. Based on these statistics, we are able to compare networks or classify them according to properties that they exhibit. There are a wealth of descriptive statistics that measure some aspect of the structure or characteristics of a network. For example, the diameter of a network measures the maximum distance between two individuals; the global clustering coefficient measures the extent to which individuals in a graph tend to cluster together; the modularity is a measure of the strength of division of a network into subgroups.

Summary statistics of networks are sometimes termed topological indices, especially in chemical or pharmacological science (Ma et al. 2018). One of the most popular topological indices is the Randić index invented in Randić (1975). The Randić index measures the extent of branching of a network (Bonchev and Trinajstic 1978; Randić 1975). It was observed that the Randić index is strongly correlated with a variety of physico-chemical properties of alkanes (Randić 1975). The Randić index play a central role in understanding quantitative structure–property and structure–activity relations in chemistry and pharmocology (Randić 2008; Randic et al. 2016). In subsequent years, the Randić index finds countless applications. For instance, it is used to characterize and quantify the similarity between different networks or subgraphs of the same network (Fourches and Tropsha 2013), it serves as a quantitative characterization of network heterogeneity (Estrada 2010), and graph robustness can be easily estimated by the Randić index (De Meo et al. 2018; Dattola et al. 2021). Moreover, the Randić index possesses a wealth of non-trivial and interesting mathematical properties (Bollobás and Erdos 1998; Bollobás et al. 1999; Cavers et al. 2010; Das et al. 2017; Li and Shi 2008). Motivated by the Randić index, various Randić-type indices have been introduced and attracted great interest in the past years. Among them, the harmonic index is a well-known one (Fajtlowicz 1987; Favaron et al. 1993; RodrIguez and Sigarreta 2017; Zhong 2012).

One of the popular research topics in the study of topological indices is to derive bounds of the indices and study their asymptotic properties. Recently, Martinez-Martinez et al. (2020, 2021) performed numeric and analytic analyses of the Randić index and the harmonic index in the Erdős–Rényi random graph. Analytic upper and lower bounds of the two indices are obtained and simulation studies show that the indices converge to one half of the number of nodes. Additionally, De Meo et al. (2018), Doslic et al. (2020) and Li et al. (2021) ind the expectations of variants of the Randić index in the Erdős–Rényi random graph. However, these results only apply to the Erdős–Rényi random graph and the exact limits of the indices are not theoretically studied.

In this paper, we shall derive the limits of the general Randić index and the general sum-connectivity index in an inhomogeneous Erdős–Rényi random graph. The general Randić index and the general sum-connectivity index contain the Randić index and the harmonic index as a special case, respectively. Thus our results theoretically validate the empirical observations in Martinez-Martinez et al. (2020, 2021) that the indices of the Erdős–Rényi random graph converge to one half of the number of nodes. In addition, our results explicitly describe how network heterogeneity affects the indices. We also observe that the limits of the Randić index and the harmonic index do not depend on the sparsity of a network, while the limits of their variants do. In this sense, the Randić index and the harmonic index are more preferable than their variants as measures of network structure.

The structure of the article is as follows. In Sect. 2 we present the main results. Section 3 summarizes simulation results and real data application. The proof is deferred to Sect. 4.

Notations: Let $c_1, c_2$ be positive constants and $n_0$ be a positive integer. For two positive sequence $a_n, b_n$, denote $a_n \asymp b_n$ if $c_1 \leq \frac{a_n}{b_n} \leq c_2$ for $n \geq n_0$; denote $a_n = O(b_n)$ if $\frac{a_n}{b_n} \leq c_2$ for $n \geq n_0$; $a_n = o(b_n)$ if $\lim_{n \to \infty} \frac{a_n}{b_n} = 0$. Let $X_n$ be a sequence of random variables. $X_n = O_P(a_n)$ means $\frac{X_n}{a_n}$ is bounded in probability. $X_n = o_P(a_n)$ means $\frac{X_n}{a_n}$ converges to zero in probability. Denote $a_+ = \max\{a, 0\}$.

## 2 The Randić index and its variants

A graph is a mathematical model of network that consists of nodes (vertices) and edges. Let $\mathcal{V} = [n] := \{1, 2, \ldots, n\}$ for a given positive integer $n$. An *undirected* graph on $\mathcal{V}$ is a pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ in which $\mathcal{E}$ is a collection of subsets of $\mathcal{V}$ such that

$|e| = 2$ for every $e \in \mathcal{E}$. Elements in $\mathcal{E}$ are called edges. A graph can be conveniently represented as an adjacency matrix $A$, where $A_{ij} = 1$ if $\{i, j\}$ is an edge, $A_{ij} = 0$ otherwise and $A_{ii} = 0$. It is clear that $A$ is symmetric, since $\mathcal{G}$ is undirected. A graph is said to be random if $A_{ij} (1 \leq i < j \leq n)$ are random.

Let $f = (f_{ij})$, $(1 \leq i < j \leq n)$ be a vector of numbers between 0 and 1. The inhomogeneous Erdős–Rényi random graph $\mathcal{G}(n, p_n, f)$ is defined as

$$\mathbb{P}(A_{ij} = 1) = p_n f_{ij},$$

where $p_n \in [0, 1]$ and $A_{ij}$ $(1 \leq i < j \leq n)$ are independent. If all $f_{ij}$ are the same, then $\mathcal{G}(n, p_n, f)$ is the Erdős–Rényi random graph. For a non-constant vector $f$, $\mathcal{G}(n, p_n, f)$ is an inhomogeneous version of the Erdős–Rényi random graph. This model covers several random graphs that have been extensively studied in random graph theory and algorithm analysis (Chakrabarty et al. 2020a, b, 2021; Chiasserini et al. 2016; Yu et al. 2021).

Given a constant $\alpha$, the general Randić index of a graph $\mathcal{G}$ is defined as (Bollobás and Erdos 1998)

$$\mathcal{R}_\alpha = \sum_{\{i,j\} \in \mathcal{E}} d_i^\alpha d_j^\alpha, \tag{1}$$

where $d_k$ is the degree of node $k$, that is, $d_k = \sum_{j \neq k} A_{kj}$. The index $\mathcal{R}_\alpha$ generalizes the well-known Randić index $\mathcal{R}_{-\frac{1}{2}}$ invented in Randić (1975). When $\alpha = -1$, the index $\mathcal{R}_{-1}$ corresponds to the modified second Zagreb index (Cavers et al. 2010; Nikolic et al. 2003).

Another popular variant of the Randić index is the general sum-connectivity index (Zhou and Trinajstic 2009, 2010) defined as

$$\chi_\alpha = \sum_{\{i,j\} \in \mathcal{E}} (d_i + d_j)^\alpha. \tag{2}$$

An important special case is the harmonic index $\mathcal{H} = 2\chi_{-1}$ (Fajtlowicz 1987; Favaron et al. 1993; Zhong 2012).

Recently Martinez-Martinez et al. (2020, 2021) conduct a simulation study of the Randić index $\mathcal{R}_{-\frac{1}{2}}$ and the harmonic index $\mathcal{H} = 2\chi_{-1}$ in the Erdős–Rényi random graph and observe that the indices converge to $n/2$. Moreover, De Meo et al. (2018), Doslic et al. (2020) and Li et al. (2021) derive analytical expressions of the expectations for the indices $\mathcal{R}_{-1}, \chi_1, \chi_2$ of the Erdős–Rényi random graph. In this paper, we shall

derive the exact limits of the general Randić index $\mathcal{R}_\alpha$ and the general sum-connectivity index $\chi_\alpha$ in $\mathcal{G}(n, p_n, f)$. Our results significantly improve the results in De Meo et al. (2018), Doslic et al. (2020), Martinez-Martinez et al. (2020, 2021) and Li et al. (2021) and provide new insights about the Randić index and its variants.

**Theorem 2.1** *Let $\alpha$ be a fixed constant and $\mathcal{G}(n, p_n, f)$ be the inhomogeneous Erdős–Rényi random graph. Suppose $np_n \log 2 \geq \log n$ and $\min_{1 \leq i < j \leq n}\{f_{ij}\} > \epsilon$ for some positive constant $\epsilon \in (0, 1)$. Then*

$$\mathcal{R}_\alpha = \left[1 + O_P\left(\frac{(\log(np_n))^{4(1-\alpha)_+}}{\sqrt{np_n}}\right)\right] p_n^{2\alpha+1} \sum_{i<j} f_i^\alpha f_j^\alpha f_{ij}, \qquad (3)$$

$$\chi_\alpha = \left[1 + O_P\left(\frac{(\log(np_n))^{2(1-\alpha)_+}}{\sqrt{np_n}}\right)\right] p_n^{\alpha+1} \sum_{i<j} (f_i + f_j)^\alpha f_{ij}, \qquad (4)$$

*where $f_i = \sum_{j \neq i}^n f_{ij}$.*

The condition $\min_{1 \leq i < j \leq n}\{f_{ij}\} > \epsilon$ implies the minimum expected degree scales with $np_n$. The condition $np_n \log 2 \geq \log n$ means that the graph is relatively dense. A similar condition is assumed in Chakrabarty et al. (2020a) to study the maximum eigenvalue of the inhomogeneous random graph.

Note that the expected total degree of $\mathcal{G}(n, p_n, f)$ has order $n^2 p_n$. Thus $p_n$ controls the sparsity of the network: a graph with smaller $p_n$ would have fewer edges. By (3) and (4), the limits of the Randić index $\mathcal{R}_{-\frac{1}{2}}$ and the harmonic $\chi_{-1}$ do not depend on $p_n$, while the limits of their variants do involve $p_n$. Asymptotically, the Randić index and the harmonic are uniquely determined by the network structure parametrized by $f$. In this sense, they are superior to their variants as measures of global structure of networks.

Now we present two examples of $\mathcal{G}(n, p_n, f)$. The simplest example is the Erdős–Rényi random graph, that is, $f_{ij} \equiv 1$. We denote the graph as $\mathcal{G}(n, p_n)$.

**Corollary 2.2** *Let $\alpha$ be a fixed constant. For the Erdős–Rényi random graph $\mathcal{G}(n, p_n)$ with $np_n \log 2 \geq \log n$, we have*

$$\mathcal{R}_\alpha = \frac{n^{2(1+\alpha)} p_n^{2\alpha+1}}{2} \left[1 + O_P\left(\frac{(\log(np_n))^{4(1-\alpha)_+}}{\sqrt{np_n}}\right)\right], \qquad (5)$$

$$\chi_\alpha = 2^{\alpha-1} n^{\alpha+2} p_n^{\alpha+1} \left[1 + O_P\left(\frac{(\log(np_n))^{2(1-\alpha)_+}}{\sqrt{np_n}}\right)\right]. \qquad (6)$$

*Especially, the Randić index $\mathcal{R}_{-\frac{1}{2}}$ is equal to*

$$\mathcal{R}_{-\frac{1}{2}} = \frac{n}{2}\left[1 + O_P\left(\frac{(\log(np_n))^{4(1-\alpha)_+}}{\sqrt{np_n}}\right)\right],$$

*the modified second Zagreb index $\mathcal{R}_{-1}$ is equal to*

$$\mathcal{R}_{-1} = \frac{1}{2p_n}\left[1 + O_P\left(\frac{(\log(np_n))^{4(1-\alpha)_+}}{\sqrt{np_n}}\right)\right],$$

*and the harmonic index $\mathcal{H}$ is equal to*

$$\mathcal{H} = \frac{n}{2}\left[1 + O_P\left(\frac{(\log(np_n))^{2(1-\alpha)_+}}{\sqrt{np_n}}\right)\right].$$

According to Corollary 2.2, the ratio $\frac{2}{n}\mathcal{R}_{-\frac{1}{2}}$ or $\frac{2}{n}\mathcal{H}$ converges in probability to 1 when $np_n \log 2 \geq \log n$. This theoretically confirms the empirical observation in Martinez-Martinez et al. (2020, 2021) that the Randić index $\mathcal{R}_{-\frac{1}{2}}$ or the harmonic index $\mathcal{H}$ is approximately equal to $\frac{n}{2}$. The expectation of the indices $\mathcal{R}_{-1}, \chi_1, \chi_2$ are derived in De Meo et al. (2018), Doslic et al. (2020) and Li et al. (2021). Our results show the indices are asymptotically equal to their expectations. Moreover, Corollary 2.2 clearly quantifies how $p_n$ affects the convergence rates: the larger $p_n$ is, the faster the convergence rates are.

In addition, (5) and (6) explicitly characterize how the leading terms of $\mathcal{R}_\alpha$ and $\chi_\alpha$ depend on $\alpha$. Note that

$$\frac{n^{2(1+\alpha)}p_n^{2\alpha+1}}{2} = \frac{n}{2}(np_n)^{2\alpha+1},$$
$$2^{\alpha-1}n^{\alpha+2}p_n^{\alpha+1} = 2^{\alpha-1}n(np_n)^{\alpha+1}.$$

For given $n$, $p_n$ such that $np_n \log 2 \geq \log n$, the leading terms are increasing functions of $\alpha$. The indices would be extremely large or small for large $|\alpha|$ and large $n$. In this sense, it is preferable to use $\mathcal{R}_\alpha$ or $\chi_\alpha$ with small $|\alpha|$ (for instance, $|\alpha| \leq 1$).

Next, we provide a non-trivial example. Let $f_{ij} = e^{-\kappa\frac{i}{n}}e^{-\kappa\frac{j}{n}}$ with a positive constant $\kappa$. Then $e^{-2\kappa} \leq f_{ij} \leq 1$ for $0 \leq i < j \leq n$. In this case, $\min_{1 \leq i < j \leq n}\{f_{ij}\} > \epsilon$ holds with $\epsilon = e^{-2\kappa}$. Straightforward calculation yields $f_i = ne^{-\kappa\frac{i}{n}}\frac{(1-e^{-\kappa})}{\kappa}(1 +$

$o(1))$ and

$$\sum_{i<j} f_i^{-1} f_j^{-1} f_{ij} = \frac{\kappa^2}{2(1-e^{-\kappa})^2} + o(1),$$

$$\sum_{i<j} f_i^{\alpha} f_j^{\alpha} f_{ij} = \frac{n^{2(\alpha+1)} \left(1-e^{-\kappa}\right)^{2\alpha} \left(1-e^{-(1+\alpha)\kappa}\right)^2}{2(1+\alpha)^2 \kappa^{2(\alpha+1)}}(1+o(1)), \quad \alpha \neq -1,$$

$$\sum_{i<j}(f_i + f_j)^{\alpha} f_{ij} = \frac{n^{\alpha+2}}{2} \left(\frac{1-e^{-\kappa}}{\kappa}\right)^{\alpha} \int_0^1 \int_0^1 \frac{\left(e^{-\kappa x} + e^{-\kappa y}\right)^{\alpha}}{e^{\kappa(x+y)}} dx dy + o(1).$$

Then

$$\mathcal{R}_{-1} = \left[1 + O_P\left(\frac{(\log(np_n))^2}{\sqrt{np_n}}\right)\right] \frac{1}{2p_n} \frac{\kappa^2}{(1-e^{-\kappa})^2}, \tag{7}$$

$$\mathcal{R}_{\alpha} = \left[1 + O_P\left(\frac{(\log(np_n))^2}{\sqrt{np_n}}\right)\right] \frac{n^{2(\alpha+1)} p_n^{2\alpha+1}}{2} \frac{\left(1-e^{-\kappa}\right)^{2\alpha} \left(1-e^{-(1+\alpha)\kappa}\right)^2}{(1+\alpha)^2 \kappa^{2(\alpha+1)}},$$

$$\alpha \neq -1, \tag{8}$$

$$\chi_{\alpha} = \left[1 + O_P\left(\frac{(\log(np_n))^2}{\sqrt{np_n}}\right)\right] \frac{n^{\alpha+2} p_n^{\alpha+1}}{2} \left(\frac{1-e^{-\kappa}}{\kappa}\right)^{\alpha} \int_0^1$$

$$\int_0^1 \frac{\left(e^{-\kappa x} + e^{-\kappa y}\right)^{\alpha}}{e^{\kappa(x+y)}} dx dy. \tag{9}$$

Since larger $\kappa$ makes the expected degrees more heterogeneous, the parameter $\kappa$ can be considered as heterogeneity level of the graph. As $\kappa$ increases, $\mathcal{R}_{\alpha}$ or $\chi_{\alpha}$ decreases if $\alpha > -1$, and $\mathcal{R}_{\alpha}$ or $\chi_{\alpha}$ increases if $\alpha \leq -1$. This shows the effect of heterogeneity on $\mathcal{R}_{\alpha}$ or $\chi_{\alpha}$. The indices could be used as indicators whether a network follows the Erdős–Rényi random graph model.

## 3 Real data application

In this section, we apply the general Randić index and the general sum index to the following real-world networks: 'karate', 'macaque', 'UKfaculty', 'enron', 'USairports', 'immuno', 'yeast'. These networks are available in the 'igraphdata' package of R.

For each network, the indices $\mathcal{R}_{-\frac{1}{2}}$, $\mathcal{R}_{-1}$, $\chi_{-\frac{1}{2}}$, $\chi_{-1}$ and the bound $\log n/(n \log 2)$ are calculated. Here, $\log n/(n \log 2)$ is the sparsity lower bound required by Theorem 2.1 and Corollary 2.2. In addition, we also compute several descriptive statistics: the number of nodes ($n$), the edge density, the maximum degree ($d_{max}$), the median degree ($d_{mean}$) and the minimum degree ($d_{min}$). These results are summerized in Table

**Table 1** The Randić index and harmonic index of real networks

| Network | $n$ | $\log n/(n \log 2)$ | Density | $d_{max}$ | $d_{median}$ | $d_{min}$ | $\mathcal{R}_{-\frac{1}{2}}$ | $\mathcal{R}_{-1}$ | $\chi_{-\frac{1}{2}}$ | $\chi_{-1}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| karate | 34 | 0.149 | 0.134 | 17 | 5 | 3 | 13.970 | 2.866 | 21.001 | 5.927 |
| macaque | 45 | 0.122 | 0.251 | 22 | 11 | 4 | 21.576 | 2.092 | 50.702 | 10.374 |
| UKfaculty | 81 | 0.078 | 0.175 | 41 | 13 | 2 | 37.728 | 2.957 | 99.101 | 17.738 |
| enron | 184 | 0.040 | 0.130 | 111 | 31 | 21 | 80.876 | 4.063 | 276.792 | 37.672 |
| USairports | 755 | 0.012 | 0.016 | 168 | 11 | 5 | 262.836 | 41.776 | 602.894 | 106.592 |
| immuno | 1316 | 0.0078 | 0.0072 | 17 | 10 | 3 | 648.820 | 70.951 | 1410.842 | 320.022 |
| yeast | 2617 | 0.004 | 0.003 | 118 | 10 | 4 | 1076.274 | 285.491 | 2034.479 | 469.020 |

1. The edge densities of networks 'macaque', 'UKfaculty', 'enron' and 'USairports' are greater than $\log n/(n \log 2)$, which indicates our theoretical results are applicable. The Randić indices $\mathcal{R}_{-\frac{1}{2}}$ and the harmonic indices $2\chi_{-1}$ of 'enron' and 'USairports' are much smaller than $\frac{n}{2}$, the indices of the Erdős–Rényi random graph. Thus the Erdős–Rényi random graph may not be a good model for these two networks. The networks 'macaque' and 'UKfaculty' have the indices close to $\frac{n}{2}$. In this sense, they can be considered as samples from the Erdős–Rényi random graph model. For the networks 'karate', 'immuno' and 'yeast', the edge densities are slightly smaller than the bound $\log n/(n \log 2)$. Note that the condition $p_n > \log n/(n \log 2)$ is a sufficient condition for Theorem 2.1 and Corollary 2.2 to hold and can not be relaxed based on the current proof technique. We conjecture that Theorem 2.1 and Corollary 2.2 still hold if $np_n \to \infty$. Currently, we are not clear whether our theoretical results can be applied to the networks 'karate', 'immuno' and 'yeast' or not. For sparse networks, that is, $np_n = O(1)$, the Randić index $\mathcal{R}_{-\frac{1}{2}}$ could assume any value between 0 and $\frac{n}{2}$, which is empirically verified in Martinez-Martinez et al. (2020). Therefore, the Randić index $\mathcal{R}_{-\frac{1}{2}}$ far less than $\frac{n}{2}$ does not necessarily imply the network are not generated from the Erdős–Rényi random graph model. We point out that a statistical hypothesis testing is needed to test whether the Randić index is equal to some number. Based on our knowledge, there is no such test available in literature. It is an interesting future topic to propose a test for the Randić index.

## 4 Proof of main results

In this section, we provide the detailed proofs of the main results. Recall that $A_{ij} = 1$ if and only if $\{i, j\}$ is an edge. Then the general Randić index in (1) and the general sum-connectivity index in (2) can be written as

$$\mathcal{R}_\alpha = \sum_{1 \le i < j \le n} A_{ij} d_i^\alpha d_j^\alpha,$$

$$\chi_\alpha = \sum_{1 \le i < j \le n} A_{ij} (d_i + d_j)^\alpha.$$

Note that the degrees $d_i$ are not independently and identically distributed. Moreover, $\mathcal{R}_\alpha$ and $\chi_\alpha$ are non-linear functions of $d_i$. These facts make it a non-trivial task to derive the limits of $\mathcal{R}_\alpha$ and $\chi_\alpha$ for general $\alpha$. The proof strategy is as follows: (a) use the Taylor expansion to expand $\mathcal{R}_\alpha$ or $\chi_\alpha$ as a sum of leading term and reminder terms; (b) find the order of the leading term and the reminder terms.

**Proof of Theorem 2.1** (I) We prove the result of the general Randić index first. For convenience, let

$$\mathcal{R}_{-\alpha} = \sum_{1 \le i < j \le n} A_{ij} d_i^{-\alpha} d_j^{-\alpha}. \tag{10}$$

We provide the proof in two cases: $\alpha > -1$ and $\alpha \le -1$. Denote $\mu_i = \mathbb{E}(d_i) = p_n f_i$.

Let $\alpha > -1$. Applying the mean value theorem to the mapping $x \to x^{-\alpha}$, we have

$$\frac{1}{d_i^\alpha} = \frac{1}{\mu_i^\alpha} - \alpha \frac{d_i - \mu_i}{X_i^{\alpha+1}},$$

where $d_i \le X_i \le \mu_i$ or $\mu_i \le X_i \le d_i$. Since $A_{ii} = 0$ $(i = 1, 2, \ldots, n)$ and the adjacency matrix $A$ is symmetric, by (10) one has

$$\begin{aligned}
\mathcal{R}_{-\alpha} &= \frac{1}{2} \sum_{1 \le i, j \le n} \frac{A_{ij}}{d_i^\alpha d_j^\alpha} \\
&= \frac{1}{2} \sum_{1 \le i, j \le n} \frac{A_{ij}}{\mu_i^\alpha \mu_j^\alpha} - \frac{\alpha}{2} \sum_{1 \le i, j \le n} \frac{A_{ij}(d_i - \mu_i)}{X_i^{\alpha+1} \mu_j^\alpha} - \frac{\alpha}{2} \sum_{1 \le i, j \le n} \frac{A_{ij}(d_j - \mu_j)}{X_j^{\alpha+1} \mu_i^\alpha} \\
&\quad + \frac{\alpha^2}{2} \sum_{1 \le i, j \le n} \frac{A_{ij}(d_i - \mu_i)(d_j - \mu_j)}{X_i^{\alpha+1} X_j^{\alpha+1}}.
\end{aligned} \tag{11}$$

Next we show the first term in (11) is the leading term. To this end, we will find the exact order of the first term and show the remaining terms are of smaller order.

Firstly, we show the first term in (11) is asymptotically equal to its expectation. By the assumption $\min_{1\le i,j\le n}\{f_{ij}\} > \epsilon$, it is clear that $np_n\epsilon \le \mu_i \le np_n$ for all $i \in [n]$ and $\epsilon n^2 \le \sum_{1\le i,j\le n} f_{ij} \le n^2$. Note that $A_{ij}$ ($1 \le i < j \le n$) are independent and $\mathbb{E}(A_{ij}) = p_n f_{ij}$. Then

$$\mathbb{E}\left[\sum_{1\le i<j\le n}\frac{A_{ij}-p_n f_{ij}}{\mu_i^\alpha \mu_j^\alpha}\right]^2 = \sum_{1\le i<j\le n}\mathbb{E}\left[\frac{A_{ij}-p_n f_{ij}}{\mu_i^\alpha \mu_j^\alpha}\right]^2 = O\left(\frac{n^2 p_n}{(np_n)^{4\alpha}}\right).$$

By the Markov's inequality, it follows that

$$\left|\sum_{1\le i<j\le n}\frac{A_{ij}}{\mu_i^\alpha \mu_j^\alpha} - \sum_{1\le i<j\le n}\frac{p_n f_{ij}}{\mu_i^\alpha \mu_j^\alpha}\right| = \left|\sum_{1\le i<j\le n}\frac{A_{ij}-p_n f_{ij}}{\mu_i^\alpha \mu_j^\alpha}\right| = O_P\left(\frac{\sqrt{n}\sqrt{np_n}}{(np_n)^{2\alpha}}\right).$$

Then we get

$$\sum_{1\le i<j\le n}\frac{A_{ij}}{\mu_i^\alpha \mu_j^\alpha} = \sum_{1\le i<j\le n}\frac{p_n f_{ij}}{\mu_i^\alpha \mu_j^\alpha} + O_P\left(\frac{\sqrt{n}\sqrt{np_n}}{(np_n)^{2\alpha}}\right)$$
$$= \sum_{1\le i<j\le n}\frac{p_n f_{ij}}{\mu_i^\alpha \mu_j^\alpha}\left(1 + O_P\left(\frac{1}{\sqrt{n}\sqrt{np_n}}\right)\right). \tag{12}$$

Now we find a bound of the second term in (11). The idea is to find an upper bound of the expectation of its absolute value and then apply the Markov's inequality to get a bound. Note that

$$\mathbb{E}\left[\left|\sum_{1\le i,j\le n}\frac{A_{ij}(d_i-\mu_i)}{X_i^{\alpha+1}\mu_j^\alpha}\right|\right] = \mathbb{E}\left[\left|\sum_{1\le i\le n}\left(\sum_{1\le j\le n}\frac{A_{ij}}{\mu_j^\alpha}\right)\frac{(d_i-\mu_i)}{X_i^{\alpha+1}}\right|\right]$$
$$\le \mathbb{E}\left[\sum_{1\le i\le n}\left(\sum_{1\le j\le n}\frac{A_{ij}}{\mu_j^\alpha}\right)\frac{|d_i-\mu_i|}{X_i^{\alpha+1}}\right]. \tag{13}$$

Let $\delta_n = [\log(np_n)]^{-2}$. Recall that $X_i$ is between $d_i$ and $\mu_i$. If $X_i < \delta_n\mu_i$ and $X_i < d_i$, then $X_i < d_i$ and $X_i < \mu_i$. In this case, $X_i$ can not be between $d_i$ and $\mu_i$. Therefore, $X_i < \delta_n\mu_i$ implies $d_i \le X_i$. Then $I[X_i < \delta_n\mu_i] \le I[d_i \le X_i < \delta_n\mu_i] \le I[X_i < \delta_n\mu_i]$. Note that $np_n\epsilon \le \mu_i \le np_n$ for all $i \in [n]$, then we have

$$\mathbb{E}\left[\left|\sum_{1\le i,j\le n}\frac{A_{ij}(d_i-\mu_i)}{X_i^{\alpha+1}\mu_j^\alpha}\right|\right] \le O\left(\frac{1}{(np_n)^\alpha}\right)\sum_{1\le i\le n}\mathbb{E}\left[\frac{d_i|d_i-\mu_i|}{X_i^{\alpha+1}}\right]$$

$$= O\left(\frac{1}{(np_n)^\alpha}\right) \sum_{1 \le i \le n} \mathbb{E}\left[\frac{d_i|d_i - \mu_i|}{X_i^{\alpha+1}} I[\delta_n \mu_i \le X_i]\right]$$

$$+ O\left(\frac{1}{(np_n)^\alpha}\right) \sum_{1 \le i \le n} \mathbb{E}\left[\frac{d_i|d_i - \mu_i|}{X_i^{\alpha+1}} I[\delta_n \mu_i > X_i]\right],$$

$$= O\left(\frac{1}{(np_n)^\alpha}\right) \sum_{1 \le i \le n} \mathbb{E}\left[\frac{d_i|d_i - \mu_i|}{X_i^{\alpha+1}} I[\delta_n \mu_i \le X_i]\right]$$

$$+ O\left(\frac{1}{(np_n)^\alpha}\right) \sum_{1 \le i \le n} \mathbb{E}\left[\frac{d_i|d_i - \mu_i|}{X_i^{\alpha+1}} I[d_i \le X_i < \delta_n \mu_i]\right].$$

$$(14)$$

Note that $\alpha > -1$. If $\delta_n \mu_i \le X_i$, then

$$\frac{1}{X_i^{\alpha+1}} \le \frac{1}{(\delta_n \mu_i)^{\alpha+1}} = O\left(\frac{1}{(\delta_n np_n)^{\alpha+1}}\right).$$

Hence we have

$$\frac{1}{(np_n)^\alpha} \sum_{1 \le i \le n} \mathbb{E}\left[\frac{d_i|d_i - \mu_i|}{X_i^{\alpha+1}} I[\delta_n \mu_i \le X_i]\right]$$

$$\le O\left(\frac{1}{(\delta_n np_n)^{\alpha+1}(np_n)^\alpha}\right) \sum_{1 \le i \le n} \mathbb{E}\left[d_i|d_i - \mu_i|I[\delta_n \mu_i \le X_i]\right]$$

$$\le O\left(\frac{1}{(\delta_n np_n)^{\alpha+1}(np_n)^\alpha}\right) \sum_{1 \le i \le n} \mathbb{E}\left[d_i|d_i - \mu_i|\right]. \qquad (15)$$

By definition, the second moment of degree $d_i$ is equal to

$$\mathbb{E}[d_i^2] = \mathbb{E}\left[\sum_{j \ne k} A_{ij} A_{ik} + \sum_j A_{ij}\right] = p_n^2 \sum_{j \ne k} f_{ij} f_{ik} + p_n \sum_j f_{ij},$$

and $Var(d_i) = \sum_{j \ne i} p_n f_{ij}(1 - p_n f_{ij})$, then by the Cauchy-Schwarz inequality, one has

$$\sum_{1 \le i \le n} \mathbb{E}\left[d_i|d_i - \mu_i|\right] \le \sum_{1 \le i \le n} \sqrt{\mathbb{E}[d_i^2]\mathbb{E}[(d_i - \mu_i)^2]}$$

$$= \sum_{1 \le i \le n} \sqrt{\left(p_n^2 \sum_{j \ne k} f_{ij} f_{ik} + p_n \sum_j f_{ij}\right) \sum_j p_n f_{ij}(1 - p_n f_{ij})}$$

$$= O\left(n\sqrt{n^3 p_n^3}\right). \qquad (16)$$

Combining (15) and (16) yields

$$\frac{1}{(np_n)^\alpha} \sum_{1 \le i \le n} \mathbb{E}\left[ \frac{d_i|d_i - \mu_i|}{X_i^{\alpha+1}} I[\delta_n \mu_i \le X_i] \right] = O\left( \frac{n\sqrt{n^3 p_n^3}}{(\delta_n np_n)^{\alpha+1}(np_n)^\alpha} \right)$$

$$= \frac{n^2 p_n}{(np_n)^{2\alpha}} O\left( \frac{1}{\delta_n^{\alpha+1}\sqrt{np_n}} \right)$$

$$= \frac{n^2 p_n}{(np_n)^{2\alpha}} O\left( \frac{(\log(np_n))^{2(\alpha+1)}}{\sqrt{np_n}} \right).$$

$$(17)$$

Now we bound the second term of (14). Note that if $d_i \le X_i < \delta_n \mu_i$, then $d_i < \mu_i$ and $\frac{d_i}{X_i^{\alpha+1}} \le \frac{1}{d_i^\alpha}$. Since $d_i$ is the degree of node $i$, it can only take integer value between 0 and $n - 1$. Moreover, $d_i = 0$ implies $A_{ij} = 0$ for any $j \in [n]$. By the definition of the Randić index (1), these terms with $d_i = 0$ are zero in (10) and (11). Therefore, we only consider the terms with $d_i \ge 1$ and $d_j \ge 1$. Then the second term of (14) can be bounded by

$$\frac{1}{(np_n)^\alpha} \sum_{1 \le i \le n} \mathbb{E}\left[ \frac{d_i|d_i - \mu_i|}{X_i^{\alpha+1}} I[d_i \le X_i < \delta_n \mu_i] \right]$$

$$\le \frac{1}{(np_n)^\alpha} \sum_{1 \le i \le n} \mathbb{E}\left[ \frac{\mu_i - d_i}{d_i^\alpha} I[d_i < \delta_n \mu_i] \right]$$

$$= \frac{1}{(np_n)^\alpha} \sum_{1 \le i \le n} \sum_{k=1}^{\delta_n \mu_i} \frac{\mu_i - k}{k^\alpha} \mathbb{P}(d_i = k).$$

$$(18)$$

Next we obtain an upper bound of $\mathbb{P}(d_i = k)$. Note that the degree $d_i$ follows the Poisson-Binomial distribution $PB(p_n f_{i1}, p_n f_{i2}, \ldots, p_n f_{in})$. Then

$$\mathbb{P}(d_i = k) = \sum_{S \subset [n]\setminus\{i\}, |S|=k} \prod_{j \in S} p_n f_{ij} \prod_{j \in S^C\setminus\{i\}} (1 - p_n f_{ij})$$

$$\le \sum_{S \subset [n]\setminus\{i\}, |S|=k} \prod_{j \in S} p_n \prod_{j \in S^C\setminus\{i\}} (1 - p_n \epsilon)$$

$$= \binom{n}{k} p_n^k (1 - p_n \epsilon)^{n-k}.$$

$$(19)$$

Note that $\binom{n}{k} \leq e^{k \log n - k \log k + k}$ and $(1 - p_n\epsilon)^{n-k} = e^{(n-k) \log(1 - p_n\epsilon)}$. Then by (19) we get

$$\mathbb{P}(d_i = k) \leq \exp\left(k \log(np_n) - k \log k + k + (n - k) \log(1 - p_n\epsilon)\right). \quad (20)$$

Let $g(k) = k \log(np_n) - k \log k + k + (n - k) \log(1 - p_n\epsilon)$. Then

$$g'(k) = \log\left(\frac{np_n}{1 - p_n\epsilon}\right) - \log k.$$

For $k < \frac{np_n}{1 - p_n\epsilon}$, $g'(k) < 0$. For $k > \frac{np_n}{1 - p_n\epsilon}$, $g'(k) > 0$. Hence $g(k)$ achieves its maximum at $k = \frac{np_n}{1 - p_n\epsilon}$. For $k \leq \delta_n np_n$, $g(k) \leq g(\delta_n np_n)$. Hence

$$\begin{aligned}
\mathbb{P}(d_i = k) &\leq \exp\left(\delta_n np_n \log \frac{1}{\delta_n(1 - p_n\epsilon)} + \delta_n np_n + n \log(1 - p_n\epsilon)\right) \\
&\leq \exp\left(-np_n\epsilon(1 + o(1))\right).
\end{aligned}$$

Note that $\mu_i \leq np_n$. Then for $k \leq \delta_n\mu_i \leq \delta_n np_n$, by (18), (19), (20), one has

$$\begin{aligned}
\mathbb{E}&\left[\frac{d_i|d_i - \mu_i|}{X_i^{\alpha+1}} I[d_i \leq X_i < \delta_n\mu_i]\right] \\
&\leq \exp\left(\log(\delta_n np_n)\right) \exp\left(\log(np_n)\right) \exp\left(-np_n\epsilon(1 + o(1))\right) \\
&= \exp\left(-np_n\epsilon(1 + o(1))\right). \quad (21)
\end{aligned}$$

Hence, we get

$$\begin{aligned}
\frac{1}{(np_n)^\alpha} \sum_{1 \leq i \leq n} \mathbb{E}\left[\frac{d_i|d_i - \mu_i|}{X_i^{\alpha+1}} I[d_i \leq X_i < \delta_n\mu_i]\right] &= \frac{1}{(np_n)^\alpha} n e^{-\epsilon np_n(1+o(1))} \\
&= \frac{n^2 p_n}{(np_n)^{2\alpha}} e^{-\epsilon np_n(1+o(1))}. \quad (22)
\end{aligned}$$

Recall that $np_n \log 2 \geq \log n$. Then $\frac{(\log(np_n))^s}{(np_n)^k} e^{-\epsilon np_n(1+o(1))} = o(1)$ for any fixed positive constants $k, s, \epsilon$. By (13), (14), (17), (22) and the Markov's inequality, one has

$$\sum_{1 \leq i,j \leq n} \frac{A_{ij}(d_i - \mu_i)}{X_i^{\alpha+1} \mu_j^\alpha} = O_P\left(\frac{n^2 p_n}{(np_n)^{2\alpha}} \frac{(\log(np_n))^{2(\alpha+1)}}{\sqrt{np_n}}\right). \quad (23)$$

The third term in (11) can be similarly bounded as the second term. Now we consider the last term in (11). Note that

$$
\sum_{1 \le i,j \le n} \frac{A_{ij}(d_i - \mu_i)(d_j - \mu_j)}{X_i^{\alpha+1} X_j^{\alpha+1}}
$$

$$
= \sum_{1 \le i,j \le n} \frac{A_{ij}(d_i - \mu_i)(d_j - \mu_j)}{X_i^{\alpha+1} X_j^{\alpha+1}} I[X_i \ge \delta_n \mu_i, \, X_j \ge \delta_n \mu_j]
$$

$$
+ \sum_{1 \le i,j \le n} \frac{A_{ij}(d_i - \mu_i)(d_j - \mu_j)}{X_i^{\alpha+1} X_j^{\alpha+1}} I[X_i < \delta_n \mu_i, \, X_j \ge \delta_n \mu_j]
$$

$$
+ \sum_{1 \le i,j \le n} \frac{A_{ij}(d_i - \mu_i)(d_j - \mu_j)}{X_i^{\alpha+1} X_j^{\alpha+1}} I[X_i \ge \delta_n \mu_i, \, X_j < \delta_n \mu_j]
$$

$$
+ \sum_{1 \le i,j \le n} \frac{A_{ij}(d_i - \mu_i)(d_j - \mu_j)}{X_i^{\alpha+1} X_j^{\alpha+1}} I[X_i < \delta_n \mu_i, \, X_j < \delta_n \mu_j]. \tag{24}
$$

We shall bound each term in (24). The first term can be bounded as follows.

$$
\mathbb{E}\left[\left| \sum_{1 \le i,j \le n} \frac{A_{ij}(d_i - \mu_i)(d_j - \mu_j)}{X_i^{\alpha+1} X_j^{\alpha+1}} I[X_i \ge \delta_n \mu_i, \, X_j \ge \delta_n \mu_j] \right|\right]
$$

$$
\le \frac{1}{\delta_n^{2(\alpha+1)}} \sum_{1 \le i,j \le n} \mathbb{E}\left[ \frac{A_{ij}|d_i - \mu_i||d_j - \mu_j|}{\mu_i^{\alpha+1} \mu_j^{\alpha+1}} I[X_i \ge \delta_n \mu_i, \, X_j \ge \delta_n \mu_j] \right]
$$

$$
\le \frac{1}{\delta_n^{2(\alpha+1)}} O\left( \frac{1}{(np_n)^{2(\alpha+1)}} \right) \sum_{1 \le i,j \le n} \mathbb{E}\left[ A_{ij}|d_i - \mu_i||d_j - \mu_j| \right]. \tag{25}
$$

Denote $\tilde{d}_i = \sum_{k \ne j,i} A_{ik}$, $\tilde{d}_j = \sum_{k \ne j,i} A_{jk}$, $\tilde{\mu}_i = \mathbb{E}(\tilde{d}_i)$ and $\tilde{\mu}_j = \mathbb{E}(\tilde{d}_j)$. Then $\tilde{d}_i$ and $\tilde{d}_j$ are independent, $d_i = \tilde{d}_i + A_{ij}$ and $d_j = \tilde{d}_j + A_{ij}$. It is easy to get that

$$
|d_i - \mu_i| = |\tilde{d}_i - \tilde{\mu}_i + A_{ij} - p_n f_{ij}| \le |\tilde{d}_i - \tilde{\mu}_i| + |A_{ij} - p_n f_{ij}| \le |\tilde{d}_i - \tilde{\mu}_i| + 1,
$$

$$
\mathbb{E}[|\tilde{d}_i - \tilde{\mu}_i|] \le \sqrt{\mathbb{E}[(\tilde{d}_i - \tilde{\mu}_i)^2]} = \sqrt{\sum_{k \ne j,i} p_n f_{ik}(1 - p_n f_{ik})} = O(\sqrt{np_n}).
$$

Similarly, $|d_j - \mu_j| \le |\tilde{d}_j - \tilde{\mu}_j| + 1$ and $\mathbb{E}[|\tilde{d}_j - \tilde{\mu}_j|] = O(\sqrt{np_n})$. Then we have

$$
\mathbb{E}\left[ A_{ij}|d_i - \mu_i||d_j - \mu_j| \right] \le \mathbb{E}[A_{ij}] + \mathbb{E}[A_{ij}|\tilde{d}_i - \tilde{\mu}_i||\tilde{d}_j - \tilde{\mu}_j|]
$$

$$
+ \mathbb{E}[A_{ij}|\tilde{d}_i - \tilde{\mu}_i|] + \mathbb{E}[A_{ij}|\tilde{d}_j - \tilde{\mu}_j|]
$$

$$
= p_n f_{ij} + p_n f_{ij} \mathbb{E}[|\tilde{d}_i - \tilde{\mu}_i|] \mathbb{E}[|\tilde{d}_j - \tilde{\mu}_j|]
$$

$$+ p_n f_{ij} \mathbb{E}[|\tilde{d}_i - \tilde{\mu}_i|] + p_n f_{ij} \mathbb{E}[|\tilde{d}_j - \tilde{\mu}_j|]$$
$$= O\left(np_n^2\right). \tag{26}$$

Combining (25) and (26) yields

$$\mathbb{E}\left[\left|\sum_{1 \le i,j \le n} \frac{A_{ij}(d_i - \mu_i)(d_j - \mu_j)}{X_i^{\alpha+1} X_j^{\alpha+1}} I[X_i \ge \delta_n \mu_i, X_j \ge \delta_n \mu_j]\right|\right]$$

$$\le \frac{1}{\delta_n^{2(\alpha+1)}} O\left(\frac{n^3 p_n^2}{(np_n)^{2(\alpha+1)}}\right)$$

$$= \frac{n^2 p_n}{(np_n)^{2\alpha}} O\left(\frac{1}{\delta_n^{2(\alpha+1)} np_n}\right)$$

$$= \frac{n^2 p_n}{(np_n)^{2\alpha}} O\left(\frac{(\log(np_n))^{4(\alpha+1)}}{np_n}\right), \tag{27}$$

The second term in (24) can be bounded as follows.

$$\mathbb{E}\left[\left|\sum_{1 \le i,j \le n} \frac{A_{ij}(d_i - \mu_i)(d_j - \mu_j)}{X_i^{\alpha+1} X_j^{\alpha+1}} I[X_i \ge \delta_n \mu_i, d_j \le X_j < \delta_n \mu_j]\right|\right]$$

$$\le \frac{1}{\delta_n^{\alpha+1}} \sum_{1 \le i,j \le n} \mathbb{E}\left[\frac{A_{ij}|d_i - \mu_i||d_j - \mu_j|}{\mu_i^{\alpha+1} d_j^{\alpha+1}} I[X_i \ge \delta_n \mu_i, d_j \le X_j < \delta_n \mu_j]\right]$$

$$\le \frac{1}{\delta_n^{\alpha+1}(np_n)^{\alpha+1}} \sum_{1 \le i,j \le n} \mathbb{E}\left[\frac{A_{ij}|d_i - \mu_i||d_j - \mu_j|}{d_j^{\alpha+1}} I[d_j < \delta_n \mu_j]\right]. \tag{28}$$

Recall that

$$|d_i - \mu_i| = |\tilde{d}_i - \tilde{\mu}_i + A_{ij} - p_n f_{ij}|, \quad |d_j - \mu_j| = |\tilde{d}_j - \tilde{\mu}_j + A_{ij} - p_n f_{ij}|.$$

Moreover, $d_j < \delta_n \mu_j$ implies $\tilde{d}_j < \delta_n \mu_j$. Then we have

$$\mathbb{E}\left[\frac{A_{ij}|d_i - \mu_i||d_j - \mu_j|}{d_j^{\alpha+1}} I[d_j < \delta_n \mu_j]\right]$$

$$= \mathbb{E}\left[\frac{A_{ij}|\tilde{d}_i - \tilde{\mu}_i + A_{ij} - p_n f_{ij}||\tilde{d}_j - \tilde{\mu}_j + A_{ij} - p_n f_{ij}|}{d_j^{\alpha+1}}\right.$$

$$\left. I[d_j < \delta_n \mu_j]\Big| A_{ij} = 1\right] \mathbb{P}(A_{ij} = 1)$$

$$\le p_n \mathbb{E}\left[\frac{|\tilde{d}_i - \tilde{\mu}_i + 1 - p_n f_{ij}||\tilde{d}_j - \tilde{\mu}_j + 1 - p_n f_{ij}|}{(\tilde{d}_j + 1)^{\alpha+1}} I[\tilde{d}_j < \delta_n \mu_j]\right]. \tag{29}$$

Since $\tilde{d}_i, \tilde{d}_j$ are independent and $\mathbb{E}[|\tilde{d}_j - \tilde{\mu}_j|] = O(\sqrt{np_n})$, then by a similar argument as in (18)–(22), it follows that

$$
\begin{aligned}
p_n \mathbb{E} & \left[ \frac{|\tilde{d}_i - \tilde{\mu}_i + 1 - p_n f_{ij}||\tilde{d}_j - \tilde{\mu}_j + 1 - p_n f_{ij}|}{(\tilde{d}_j + 1)^{\alpha+1}} I[\tilde{d}_j < \delta_n \mu_j] \right] \\
& \leq p_n \sqrt{np_n} \mathbb{E} \left[ \frac{|\tilde{d}_j - \tilde{\mu}_j + 1 - p_n f_{ij}|}{(\tilde{d}_j + 1)^{\alpha+1}} I[\tilde{d}_j < \delta_n \mu_j] \right] \\
& \leq p_n \sqrt{np_n} e^{-\epsilon n p_n (1 + o(1))}.
\end{aligned}
\tag{30}
$$

Combining (21), (29) and (30) yields

$$
\begin{aligned}
\mathbb{E} & \left[ \left| \sum_{1 \leq i, j \leq n} \frac{A_{ij}(d_i - \mu_i)(d_j - \mu_j)}{X_i^{\alpha+1} X_j^{\alpha+1}} I[X_i \geq \delta_n \mu_i, d_j \leq X_j < \delta_n \mu_j] \right| \right] \\
& \leq \frac{p_n \sqrt{np_n}}{\delta_n^{\alpha+1} (np_n)^{\alpha+1}} n^2 e^{-\epsilon n p_n (1 + o(1))} \\
& = \frac{n^2 p_n}{(np_n)^{2\alpha}} e^{-\epsilon n p_n (1 + o(1))}.
\end{aligned}
\tag{31}
$$

The third term in (24) can be similarly bounded as the second term. Now we consider the last term in (24). By a similar argument as in (21)–(31), one gets

$$
\begin{aligned}
\mathbb{E} & \left[ \left| \sum_{1 \leq i, j \leq n} \frac{A_{ij}(d_i - \mu_i)(d_j - \mu_j)}{X_i^{\alpha+1} X_j^{\alpha+1}} I[d_i \leq X_i < \delta_n \mu_i, d_j \leq X_j < \delta_n \mu_j] \right| \right] \\
& \leq \sum_{1 \leq i, j \leq n} \mathbb{E} \left[ \frac{A_{ij}|d_i - \mu_i||d_j - \mu_j|}{d_i^{\alpha+1} d_j^{\alpha+1}} I[d_i \leq \delta_n \mu_i, d_j \leq \delta_n \mu_j] \right] \\
& \leq \sum_{1 \leq i, j \leq n} \mathbb{E} \left[ \frac{A_{ij}|\tilde{d}_i - \tilde{\mu}_i + A_{ij} - p_n f_{ij}||\tilde{d}_j - \tilde{\mu}_j + A_{ij} - p_n f_{ij}|}{(\tilde{d}_j + A_{ij})^{\alpha+1}(\tilde{d}_j + A_{ij})^{\alpha+1}} \right. \\
& \qquad \left. I[\tilde{d}_i \leq \delta_n \mu_i, \tilde{d}_j \leq \delta_n \mu_j] \right] \\
& \leq p_n \sum_{1 \leq i, j \leq n} \mathbb{E} \left[ \frac{(|\tilde{d}_i - \tilde{\mu}_i| + 1)(|\tilde{d}_j - \tilde{\mu}_j| + 1)}{(\tilde{d}_j + 1)^{\alpha+1}(\tilde{d}_j + 1)^{\alpha+1}} I[\tilde{d}_i \leq \delta_n \mu_i, \tilde{d}_j \leq \delta_n \mu_j] \right] \\
& = p_n \left( \sum_{1 \leq i \leq n} \mathbb{E} \left[ \frac{(|\tilde{d}_i - \tilde{\mu}_i| + 1)}{(\tilde{d}_i + 1)^{\alpha+1}} I[\tilde{d}_i \leq \delta_n \mu_i] \right] \right)^2 \\
& \leq p_n n^2 e^{-2\epsilon n p_n (1 + o(1))} = \frac{n^2 p_n}{(np_n)^{2\alpha}} e^{-2\epsilon n p_n (1 + o(1))}.
\end{aligned}
\tag{32}
$$

By (24)–(32) and the Markov's inequality, it follows that

$$\sum_{1 \le i,j \le n} \frac{A_{ij}(d_i - \mu_i)(d_j - \mu_j)}{X_i^{\alpha+1} X_j^{\alpha+1}} = O_P \left( \frac{n^2 p_n}{(np_n)^{2\alpha}} \frac{(\log(np_n))^{4(\alpha+1)}}{np_n} \right). \quad (33)$$

It is easy to verify that $\sum_{1 \le i < j \le n} \frac{p_n f_{ij}}{\mu_i^\alpha \mu_j^\alpha} \ge \frac{\epsilon n(n-1)p_n}{2(np_n)^{2\alpha}}$. Then combining (11), (12), (23) and (33) yields the limit of $\mathcal{R}_{-\alpha}$ with $\alpha > -1$.

Next, we consider $\mathcal{R}_{-\alpha}$ for $\alpha \le -1$. In this case, we rewrite the general Randić index as

$$\mathcal{R}_\alpha = \sum_{1 \le i < j \le n} A_{ij} d_i^\alpha d_j^\alpha, \alpha \ge 1. \quad (34)$$

By the Taylor expansion, we have

$$d_i^\alpha = \mu_i^\alpha + \alpha X_i^{\alpha-1}(d_i - \mu_i),$$

where $X_i$ is between $d_i$ and $\mu_i$. Then

$$\begin{aligned}
\mathcal{R}_\alpha &= \frac{1}{2} \sum_{1 \le i,j \le n} A_{ij} d_i^\alpha d_j^\alpha \\
&= \frac{1}{2} \sum_{1 \le i,j \le n} A_{ij} \mu_i^\alpha \mu_j^\alpha \\
&\quad + \frac{\alpha}{2} \sum_{1 \le i,j \le n} A_{ij}(d_i - \mu_i) X_i^{\alpha-1} \mu_j^\alpha + \frac{\alpha}{2} \sum_{1 \le i,j \le n} A_{ij}(d_j - \mu_j) X_j^{\alpha-1} \mu_i^\alpha \\
&\quad + \frac{\alpha^2}{2} \sum_{1 \le i,j \le n} A_{ij}(d_i - \mu_i)(d_j - \mu_j) X_i^{\alpha-1} X_j^{\alpha-1}. \quad (35)
\end{aligned}$$

We shall show that the first term in (35) is the leading term and the remaining terms are of smaller order. Similar to (12), it is easy to get

$$\sum_{1 \le i < j \le n} A_{ij} \mu_i^\alpha \mu_j^\alpha = \sum_{1 \le i < j \le n} p_n f_{ij} \mu_i^\alpha \mu_j^\alpha \left( 1 + O_P \left( \frac{1}{\sqrt{n} \sqrt{np_n}} \right) \right). \quad (36)$$

Since the second term and the third term in (35) have the same order, we only need to bound the second term and the last term. Let $M = \frac{4}{\epsilon(1-p_n\epsilon)}$. Clearly $M$ is bounded and $M > 4$. The expectation of the absolute value of the second term in (35) can be bounded by

$$\mathbb{E}\left[\left|\sum_{1 \le i,j \le n} A_{ij}(d_i - \mu_i)X_i^{\alpha-1}\mu_j^{\alpha}\right|\right]$$

$$\le \mathbb{E}\left[\sum_{1 \le i,j \le n} A_{ij}|d_i - \mu_i| X_i^{\alpha-1}\mu_j^{\alpha} I[M\mu_i \le X_i \le d_i]\right]$$

$$+\mathbb{E}\left[\sum_{1 \le i,j \le n} A_{ij}|d_i - \mu_i| X_i^{\alpha-1}\mu_j^{\alpha} I[X_i \le M\mu_i]\right]. \tag{37}$$

Note that

$$\mathbb{E}\left[\sum_{1 \le i,j \le n} A_{ij}|d_i - \mu_i| X_i^{\alpha-1}\mu_j^{\alpha} I[X_i \le M\mu_i]\right]$$

$$\le M^{\alpha-1}(np_n)^{2\alpha-1} \sum_{1 \le i,j \le n} \mathbb{E}\left[A_{ij}\left|\tilde{d}_i - \mu_i + A_{ij}\right|\right]$$

$$= (np_n)^{2\alpha} n^2 p_n O\left(\frac{1}{\sqrt{np_n}}\right), \tag{38}$$

and

$$\mathbb{E}\left[\sum_{1 \le i,j \le n} A_{ij}|d_i - \mu_i| X_i^{\alpha-1}\mu_j^{\alpha} I[M\mu_i \le X_i \le d_i]\right]$$

$$\le O((np_n)^{\alpha})\mathbb{E}\left[\sum_{1 \le i,j \le n} A_{ij}|d_i - \mu_i| d_i^{\alpha-1} I[M\mu_i \le d_i]\right]$$

$$= O((np_n)^{\alpha} p_n) \sum_{1 \le i,j \le n} \mathbb{E}\left[\left|\tilde{d}_i - \tilde{\mu}_i + 1 - p_n f_{ij}\right| \tilde{d}_i^{\alpha-1} I[M\mu_i - 1 \le \tilde{d}_i]\right]$$

$$= O((np_n)^{\alpha} p_n) \sum_{1 \le i,j \le n} \sum_{k=M\mu_i-1}^{n-2} k^{\alpha-1}(k - \tilde{\mu}_i + 1 - p_n f_{ij})\mathbb{P}(\tilde{d}_i = k). \tag{39}$$

By a similar argument as in (20), it follows that

$$\sum_{k=M\mu_i-1}^{n-2} k^{\alpha-1}(k - \tilde{\mu}_i + 1 - p_n f_{ij})\mathbb{P}(\tilde{d}_i = k) \le \sum_{k=M\mu_i-1}^{n-2} k^{\alpha}\binom{n}{k} p_n^k (1 - p_n \epsilon)^{n-k}$$

$$\le \sum_{k=M\mu_i-1}^{n-2} \exp\left(\alpha \log k + g(k)\right). \tag{40}$$

Let $h(k) = \alpha \log k + g(k)$. Then

$$h'(k) = \frac{\alpha}{k} + \log\left(\frac{np_n}{1 - p_n\epsilon}\right) - \log k.$$

Hence $h(k)$ is decreasing for $k > \frac{1.1np_n}{1-p_n\epsilon}$ and large $n$. Since $k \geq M\mu_i - 1 \geq M\epsilon np_n - 1 \geq \frac{2np_n}{1-p_n\epsilon}$ for large $n$, then

$$\begin{aligned} h(k) \ &\leq h\left(\frac{2np_n}{1 - p_n\epsilon}\right) = \alpha \log\left(\frac{2np_n}{1 - p_n\epsilon}\right) - \frac{2np_n \log 2}{1 - p_n\epsilon} \\ &+ n\log(1 - p_n\epsilon) \leq -\frac{np_n \log 2}{1 - p_n\epsilon} - \epsilon np_n. \end{aligned}$$

By the assumption $np_n \log 2 \geq \log n$, it is easy to get $\log n - \frac{np_n \log 2}{1-p_n\epsilon} < 0$. Then

$$\sum_{k=M\mu_i-1}^{n-2} k^{\alpha-1}(k - \tilde{\mu}_i + 1 - p_n f_{ij})\mathbb{P}(\tilde{d}_i = k) \leq n \exp\left(-\frac{np_n \log 2}{1 - p_n\epsilon} - \epsilon np_n\right)$$

$$\leq \exp\left(-\epsilon np_n(1 + o(1))\right). \quad (41)$$

Hence (37) is bounded by $(np_n)^{2\alpha}n^2 p_n O\left(\frac{1}{\sqrt{np_n}}\right)$.

Now we bound the last term in (35). Note that

$$\begin{aligned} &\sum_{1\leq i,j\leq n} \left|A_{ij}(d_i - \mu_i)(d_j - \mu_j)X_i^{\alpha-1}X_j^{\alpha-1}\right| \\ &= \sum_{1\leq i,j\leq n} \left|A_{ij}(d_i - \mu_i)(d_j - \mu_j)X_i^{\alpha-1}X_j^{\alpha-1}\right| I\left[X_i \leq M\mu_i, X_j \leq M\mu_j\right] \\ &+ \sum_{1\leq i,j\leq n} \left|A_{ij}(d_i - \mu_i)(d_j - \mu_j)X_i^{\alpha-1}X_j^{\alpha-1}\right| I\left[X_i \leq M\mu_i, X_j \geq M\mu_j\right] \\ &+ \sum_{1\leq i,j\leq n} \left|A_{ij}(d_i - \mu_i)(d_j - \mu_j)X_i^{\alpha-1}X_j^{\alpha-1}\right| I\left[X_i \geq M\mu_i, X_j \leq M\mu_j\right] \\ &+ \sum_{1\leq i,j\leq n} \left|A_{ij}(d_i - \mu_i)(d_j - \mu_j)X_i^{\alpha-1}X_j^{\alpha-1}\right| I\left[X_i \geq M\mu_i, X_j \geq M\mu_j\right]. \end{aligned}$$

$$(42)$$

Since $X_i$ is between $d_i$ and $\mu_i$, then $X_i \leq M\mu_i$ implies $d_i \leq X_i \leq M\mu_i$, and $X_i \geq M\mu_i$ implies $d_i \geq X_i \geq M\mu_i$. Similar results hold for $X_j$. Then by (42) we have

$$\sum_{1\leq i,j\leq n} \left|A_{ij}(d_i - \mu_i)(d_j - \mu_j)X_i^{\alpha-1}X_j^{\alpha-1}\right|$$

$$\leq \sum_{1 \leq i,j \leq n} \left|A_{ij}(d_i - \mu_i)(d_j - \mu_j)X_i^{\alpha-1}X_j^{\alpha-1}\right| I\left[X_i \leq M\mu_i, X_j \leq M\mu_j\right]$$

$$+ \sum_{1 \leq i,j \leq n} \left|A_{ij}(d_i - \mu_i)(d_j - \mu_j)X_i^{\alpha-1}X_j^{\alpha-1}\right| I\left[X_i \leq M\mu_i, d_j \geq X_j \geq M\mu_j\right]$$

$$+ \sum_{1 \leq i,j \leq n} \left|A_{ij}(d_i - \mu_i)(d_j - \mu_j)X_i^{\alpha-1}X_j^{\alpha-1}\right| I\left[d_i \geq X_i \geq M\mu_i, X_j \leq M\mu_j\right]$$

$$+ \sum_{1 \leq i,j \leq n} \left|A_{ij}(d_i - \mu_i)(d_j - \mu_j)X_i^{\alpha-1}X_j^{\alpha-1}\right.$$

$$\left. I\left[d_i \geq X_i \geq M\mu_i, d_j \geq X_j \geq M\mu_j\right]. \right. \tag{43}$$

Now we bound the expectation of each term in (43). Since the second term and the third term have the same order, it suffices to bound the first term, second term and the last term. By a similar argument as in (39) and (41), it is easy to get the following results.

$$\mathbb{E}\left[\sum_{1 \leq i,j \leq n} A_{ij}|d_i - \mu_i||d_j - \mu_j|X_i^{\alpha-1}X_j^{\alpha-1} I\left[X_i \leq M\mu_i, X_j \leq M\mu_j\right]\right]$$

$$\leq O\left((np_n)^{2(\alpha-1)}p_n\right) \sum_{1 \leq i,j \leq n} \mathbb{E}|\tilde{d}_i - \tilde{\mu}_i + 1 - p_n f_{ij}||\tilde{d}_j - \tilde{\mu}_j + 1 - p_n f_{ij}|$$

$$= O\left((np_n)^{2(\alpha-1)}p_n n^2 np_n\right)$$

$$= (np_n)^{2\alpha} n^2 p_n O\left(\frac{1}{np_n}\right), \tag{44}$$

$$\mathbb{E}\left[\sum_{1 \leq i,j \leq n} A_{ij}|d_i - \mu_i||d_j - \mu_j|X_i^{\alpha-1}X_j^{\alpha-1} I\left[d_i \geq X_i \geq M\mu_i, d_j \geq X_j \geq M\mu_j\right]\right]$$

$$\leq \mathbb{E}\left[\sum_{1 \leq i,j \leq n} A_{ij}d_i^{\alpha}d_j^{\alpha} I\left[d_i \geq M\mu_i, d_j \geq M\mu_j\right]\right]$$

$$\leq p_n \sum_{1 \leq i,j \leq n} \mathbb{E}\left[(\tilde{d}_i + 1)^{\alpha}(\tilde{d}_j + 1)^{\alpha} I\left[\tilde{d}_i \geq M\mu_i - 1, \tilde{d}_j \geq M\mu_j - 1\right]\right]$$

$$= p_n \left(\sum_{1 \leq i \leq n} \mathbb{E}\left[(\tilde{d}_i + 1)^{\alpha} I[\tilde{d}_i \geq M\mu_i - 1]\right]\right)^2$$

$$= O\left(n^2 p_n\right) \exp\left(-2\epsilon np_n(1 + o(1))\right), \tag{45}$$

and

$$\mathbb{E}\left[\sum_{1 \leq i,j \leq n} A_{ij}|d_i - \mu_i||d_j - \mu_j|X_i^{\alpha-1}X_j^{\alpha-1} I[X_i \leq M\mu_i, d_j \geq X_j \geq M\mu_j]\right]$$

$$\leq O\left((np_n)^{\alpha-1}\right)\mathbb{E}\left[\sum_{1 \leq i,j \leq n} A_{ij}|d_i - \mu_i|d_j^{\alpha} I[d_j \geq M\mu_j]\right]$$

$$\leq O\left((np_n)^{\alpha-1}p_n\right)\sum_{1\leq i,j\leq n}\mathbb{E}\left[|\tilde{d}_i-\tilde{\mu}_i+1-p_nf_{ij}|(\tilde{d}_j+1)^{\alpha}I[\tilde{d}_j\geq M\mu_j-1]\right]$$

$$= O\left((np_n)^{\alpha-1}p_nn^2\sqrt{np_n}\right)\exp\left(-\epsilon np_n(1+o(1))\right). \tag{46}$$

Combining (35)–(46) yields the desired result. Then the proof of the result of the general Randić index is complete.

(II). Now we prove the result of the general sum-connectivity index. We provide the proof in two cases: $\alpha < 1$ and $\alpha \geq 1$.

Firstly we work on $\chi_{-\alpha}$ with $\alpha > -1$. By Taylor expansion or the mean value theorem, we have

$$\chi_{-\alpha} = \frac{1}{2}\sum_{1\leq i,j\leq n}\frac{A_{ij}}{(d_i+d_j)^{\alpha}} = \frac{1}{2}\sum_{1\leq i,j\leq n}\frac{A_{ij}}{(\mu_i+\mu_j)^{\alpha}}$$

$$-\frac{\alpha}{2}\sum_{1\leq i,j\leq n}\frac{A_{ij}}{X_{ij}^{\alpha+1}}(d_i-\mu_i+d_j-\mu_j), \tag{47}$$

where $X_{ij}$ is between $\mu_i+\mu_j$ and $d_i+d_j$. We shall prove the first term is the leading term and the second term has smaller order than the first term.

By a similar argument as in (12), it is easy to get

$$\sum_{i<j}\frac{A_{ij}}{(\mu_i+\mu_j)^{\alpha}} = \sum_{i<j}\frac{p_nf_{ij}}{(\mu_i+\mu_j)^{\alpha}}\left(1+O_P\left(\frac{1}{\sqrt{n^2p_n}}\right)\right). \tag{48}$$

Hence the first term of (47) is asymptotically equal to $\sum_{i<j}\frac{p_nf_{ij}}{(\mu_i+\mu_j)^{\alpha}}$.

Let $\delta_n = [\log(np_n)]^{-2}$. Since $X_{ij}$ is between $\mu_i+\mu_j$ and $d_i+d_j$, $X_{ij}\leq\delta_n(\mu_i+\mu_j)$ implies $d_i+d_j\leq X_{ij}\leq\delta_n(\mu_i+\mu_j)$. Then

$$\sum_{i,j}\left|\frac{A_{ij}}{X_{ij}^{\alpha+1}}(d_i-\mu_i+d_j-\mu_j)\right|$$

$$\leq\sum_{i,j}\left|\frac{A_{ij}}{X_{ij}^{\alpha+1}}(d_i-\mu_i+d_j-\mu_j)\right|I\left[d_i+d_j\leq X_{ij}\leq\delta_n(\mu_i+\mu_j)\right]$$

$$+\sum_{i,j}\left|\frac{A_{ij}}{X_{ij}^{\alpha+1}}(d_i-\mu_i+d_j-\mu_j)\right|I\left[X_{ij}\geq\delta_n(\mu_i+\mu_j)\right]. \tag{49}$$

Next we bound the expectation of each term in (45). For the second term, the expectation can be bounded as follows.

$$\mathbb{E}\left[\sum_{i,j}\frac{A_{ij}}{X_{ij}^{\alpha+1}}(|d_i-\mu_i|+|d_j-\mu_j|)I\left[X_{ij}\geq\delta_n(\mu_i+\mu_j)\right]\right]$$

$$\leq O\left(\frac{1}{\delta_n^{\alpha+1}(np_n)^{\alpha+1}}\right)\sum_{i,j}\mathbb{E}\left[A_{ij}(|\tilde{d}_i-\mu_i+A_{ij}|+|\tilde{d}_j-\mu_i+A_{ij}|)\right]$$

$$= O\left(\frac{n^2 p_n\sqrt{np_n}}{\delta_n^{\alpha+1}(np_n)^{\alpha+1}}\right) = \frac{n^2 p_n}{(np_n)^\alpha}O\left(\frac{[\log(np_n)]^{2(\alpha+1)}}{\sqrt{np_n}}\right). \tag{50}$$

Next we focus on the first term in (45). It is clear that

$$\mathbb{E}\left[\sum_{i,j}\frac{A_{ij}}{X_{ij}^{\alpha+1}}(|d_i-\mu_i|+|d_j-\mu_j|)I[d_i+d_j\leq X_{ij}<\delta_n(\mu_i+\mu_j)]\right]$$

$$\leq \mathbb{E}\left[\sum_{i,j}\frac{A_{ij}(|d_i-\mu_i|+|d_j-\mu_j|)}{(d_i+d_j)^{\alpha+1}}I\left[d_i+d_j<\delta_n(\mu_i+\mu_j)\right]\right].$$

Note that $d_i+d_j<\delta_n(\mu_i+\mu_j)$ implies $d_i<\delta_n(\mu_i+\mu_j)$ and $d_j<\delta_n(\mu_i+\mu_j)$, and

$$\frac{|d_i-\mu_i|+|d_j-\mu_j|}{(d_i+d_j)^{\alpha+1}}=\frac{|d_i-\mu_i|}{(d_i+d_j)^{\alpha+1}}+\frac{|d_j-\mu_j|}{(d_i+d_j)^{\alpha+1}}\leq\frac{|d_i-\mu_i|}{d_i^{\alpha+1}}+\frac{|d_j-\mu_j|}{d_j^{\alpha+1}}.$$

Then we have

$$\mathbb{E}\left[\sum_{i,j}\frac{A_{ij}}{X_{ij}^{\alpha+1}}(|d_i-\mu_i|+|d_j-\mu_j|)I\left[d_i+d_j\leq X_{ij}<\delta_n(\mu_i+\mu_j)\right]\right]$$

$$\leq \mathbb{E}\left[\sum_{i,j}\frac{A_{ij}|d_i-\mu_i|}{d_i^{\alpha+1}}I\left[d_i<\delta_n(\mu_i+\mu_j)\right]\right]$$

$$+\mathbb{E}\left[\sum_{i,j}\frac{A_{ij}|d_j-\mu_j|}{d_j^{\alpha+1}}I\left[d_j<\delta_n(\mu_i+\mu_j)\right]\right]$$

$$\leq 2 p_n \mathbb{E}\left[\sum_{i,j} \frac{|\tilde{d}_i - \mu_i + 1|}{(\tilde{d}_i + 1)^{\alpha+1}} I\left[\tilde{d}_i < \delta_n(\mu_i + \mu_j)\right]\right]$$

$$= n^2 p_n e^{-\epsilon n p_n(1+o(1))} = \frac{n^2 p_n}{(np_n)^\alpha} e^{-\epsilon n p_n(1+o(1))}. \tag{51}$$

Combining (47)–(51) yields

$$\chi_{-\alpha} = p_n^{1-\alpha} \sum_{i<j} \frac{f_{ij}}{(f_i + f_j)^\alpha} \left(1 + O_P\left(\frac{[\log(np_n)]^{2(\alpha+1)}}{\sqrt{n^2 p_n}}\right)\right), \alpha > -1.$$

Now we work on $\chi_\alpha$ with $\alpha \geq 1$. When $\alpha = 1$, the proof is trivial. We will focus on $\alpha > 1$. By the mean value theorem, one has

$$\chi_\alpha = \frac{1}{2} \sum_{i,j} A_{ij}(d_i + d_j)^\alpha = \frac{1}{2} \sum_{i,j} A_{ij}(\mu_i + \mu_j)^\alpha$$

$$+ \frac{\alpha}{2} \sum_{i,j} A_{ij} X_{ij}^{\alpha-1}(d_i - \mu_i + d_j - \mu_j), \tag{52}$$

where $X_{ij}$ is between $\mu_i + \mu_j$ and $d_i + d_j$.

The remaining proof is similar to the proof of the case $\alpha < 1$. Let $M = \frac{4}{\epsilon(1-p_n\epsilon)}$. It is clear $M$ is bounded and $M > 4$. Note that

$$\sum_{i,j} \mathbb{E}\left[A_{ij} X_{ij}^{\alpha-1}(|d_i - \mu_i| + |d_j - \mu_j|) I[X_{ij} \leq M(\mu_i + \mu_j)]\right]$$

$$= (np_n)^\alpha n^2 p_n O\left(\frac{1}{\sqrt{np_n}}\right), \tag{53}$$

and

$$\sum_{i,j} \mathbb{E}\left[A_{ij} X_{ij}^{\alpha-1}(|d_i - \mu_i + d_j - \mu_j|) I[d_i + d_j \geq X_{ij} > M(\mu_i + \mu_j)]\right]$$

$$\leq O(1) \sum_{i,j} \mathbb{E}\left[A_{ij}(\tilde{d}_i + \tilde{d}_j + 2A_{ij})^{\alpha-1}(|\tilde{d}_i + \tilde{d}_j - \mu_i - \mu_j + 2A_{ij}|)\right.$$

$$\left. I[\tilde{d}_i + \tilde{d}_j > M(\mu_i + \mu_j - 1)]\right]$$

$$\leq O(1) p_n \sum_{i,j} \mathbb{E}\left[(\tilde{d}_i + \tilde{d}_j + 2)^{\alpha-1}(\tilde{d}_i + \tilde{d}_j) I[\tilde{d}_i + \tilde{d}_j > M(\mu_i + \mu_j - 1)]\right]$$

$$\leq O(1)p_n \sum_{i,j} \sum_{k=M(\mu_i+\mu_j-1)}^{2(n-2)} (k+2)^{\alpha-1} k \mathbb{P}(\tilde{d}_i + \tilde{d}_j = k)$$

$$= n^2 p_n n e^{-\epsilon n p_n(1+o(1))} = (np_n)^\alpha n^2 p_n e^{-\epsilon n p_n(1+o(1))}, \tag{54}$$

where the second last step follows from a similar argument as in (41) by noting that $\tilde{d}_i + \tilde{d}_j$ follows the Poisson–Binomial distribution.

Combining (52), (53) and (54) yields

$$\chi_\alpha = \left(1 + O_P\left(\frac{1}{\sqrt{np_n}}\right)\right) p_n^{\alpha+1} \sum_{i<j} (f_i + f_j)^\alpha f_{ij}, \alpha \geq 1.$$

Then the proof is complete. □

## Declarations

**Conflict of interest** The author has no conflict of interest to disclose.

## References

Abbe E (2018) Community detection and stochastic block models: recent developments. J Mach Learn Res 18:1–86

Bickel PJ, Sarkar P (2016) Hypothesis testing for automated community detection in networks. J R Stat Soc Ser B 78:253–273

Bollobás B, Erdos P (1998) Graphs of extremal weights. Ars Comb 50:225–233

Bollobás B, Erdős P, Sarkar A (1999) Extremal graphs for weights. Discrete Math 200:5–19

Bonchev D, Trinajstic N (1978) On topological characterization of molecular branching. Int J Quantum Chem Quantum Chem Symp 12:293–303

Cavers M, Fallat S, Kirkland S (2010) On the normalized Laplacian energy and general Randić index $r_1$ of graphs. Linear Algebra Appl 433:172–190

Chakrabarty A, Hazra SR, Hollander FD, Sfragara M (2020) Large deviation principle for the maximal eigenvalue of inhomogeneous Erdős-Rényi random graphs. J Theor Probab. https://doi.org/10.1007/s10959-021-01138-w

Chakrabarty A, Chakrabarty S, Hazra RS (2020) Eigenvalues outside the bulk of of inhomogeneous Erdős-Rényi random graphs. J Stat Phys 181:1746–1780

Chakrabarty A, Hazra SR, Hollander FD, Sfragara M (2021) Spectra of adjacency and Laplacian matrices of inhomogeneous Erdős-Rényi random graphs. Random Matrices Theory Appl 10(1):215009

Chiasserini CF, Garetto M, Leonardi E (2016) Social network de-anonymization under scale-free user relations. IEEE/ACM Trans Netw 24(6):3756–3769

Das KC, Sun S, Gutman I (2017) Normalized Laplacian eigenvalues and Randić energy of graphs. MATCH Commun Math Comput Chem 77:45–59

Dattola S et al (2021) Testing graph robustness indexes for EEG analysis in Alzheimer's disease diagnosis. Electronics 10:1440

De Meo P et al (2018) Estimating graph robustness through the Randic index. IEEE Trans Cybern 48(11):3232–3242

Doslic T et al (2020) On generalized Zagreb indices of random graphs. MATCH Commun Math Comput Chem 84:499–511

Estrada E (2010) Quantifying network heterogeneity. Phys Rev E 82:066102

Fajtlowicz S (1987) On conjectures of Graffiti-II. Congr Numer 60:187–197

Favaron O, Mahéo M, Saclé JF (1993) Some eigenvalue properties in graphs (conjectures of Graffiti-II). Discrete Math 111:197–220

Fourches D, Tropsha A (2013) Using graph indices for the analysis and comparison of chemical datasets. Mol Inform 32(9–10):827–842

Goldenberg A et al (2010) A survey of statistical network models. Found Trends Mach Learn 2(2):129–233

Kolaczyk E (2009) Statistical analysis of network data. Springer, Berlin

Li X, Shi Y (2008) A survey on the Randic index. MATCH Commun Math Comput Chem 59:127–156

Li S, Shi L, Gao W (2021) Two modified Zagreb indices for random structures. Main Group Met Chem 44:150–156

Ma YD et al (2018) From the connectivity index to various Randic-type descriptors. MATCH Commun Math Comput Chem 80:85–106

Martinez-Martinez CT, Mendez-Bermudez JA, Rodriguez J, Sigarreta J (2020) Computational and analytical studies of the Randić index in Erdős-Rényi models. Appl Math Comput 377:125137

Martinez-Martinez CT, Mendez-Bermudez JA, Rodriguez J, Sigarreta J (2021) Computational and analytical studies of the harmonic index on Erdős-Rényi models. MATCH Commun Math Comput Chem 85:395–426

Newman M (2009) Networks: an introduction. Oxford University Press, Oxford

Nikolic S et al (2003) The Zagreb indices 30 years after. Croat Chem Acta 76:113–124

Randić M (1975) Characterization of molecular branching. J Am Chem Soc 97(23):6609–6615

Randić M (2008) On history of the Randić index and emerging hostility toward chemical graph theory. MATCH Commun Math Comput Chem 59:5–124

Randic M, Novi CM, Plavsic D (2016) Solved and unsolved problems in structural chemistry. CRC Press, Boca Raton

RodrIguez JM, Sigarreta JM (2017) New results on the harmonic index and its generalizations. MATCH Commun Math Comput Chem 78:387–404

Yu L, Xu J, Lin X (2021) The power of D-hops in matching power-law graphs. Proc ACM Meas Anal Comput Syst 5(2):1–43

Zhong L (2012) The harmonic index for graphs. Appl Math Lett 25:561–566

Zhou B, Trinajstic N (2009) On a novel connectivity index. J Math Chem 46:1252–1270

Zhou B, Trinajstic N (2010) On general sum-connectivity index. J Math Chem 47:210–218