Sombor index and degree-related properties of simplicial networks

Yilun Shang[1]

[1] *Department of Computer and Information Sciences, Northumbria University, Newcastle NE1 8ST, UK*
Email: `shylmath@hotmail.com; yilun.shang@northumbria.ac.uk`

Abstract

Many dynamical effects in biology, social and technological complex systems have recently revealed their relevance to group interactions beyond traditional dyadic relationships between individual units. In this paper, we propose a growing simplicial network to model the higher-order interactions represented by clique structures. We analytically study the degree distribution and clique distribution of the network model. As an important degree-based topological index, Sombor index of the model has been derived in an iterative manner and an approximation method with closed expression is proposed. Moreover, we observe power-law and small-word effect for the simplicial networks and examine the effectiveness of the approximation method for Sombor index through computational experiments. We discover the scaling constant for Sombor index with the evolution of the network when the initial seed network is modeled as an Erdős-Rényi random graph. Our findings suggest the relevance and potential applicability of simplicial networks in modelling higher-order interactions in complex networked systems.

Keywords:
Sombor index, degree distribution, clique distribution, distance, complex network, random graph.
2000 MSC: 05C07, 05C69, 05C82, 05C80.

1. Introduction

Networks provide a fundamental system-level description of complex interconnected systems made of interacting units through the edges in the networks. Despite the success of network presentation during the past decades, the strong limitation of a single type of pairwise or dyadic interactions falls short in effectively capturing many empirical systems [5, 6]. The significance of higher-order interactions has been highlighted recently in a variety of real-world systems in nature [22], biology [47] and technology [39], with examples ranging from scientific collaboration [48] to neuronal activity in brains [9], from social contagion [29] to competition and cooperation in ecosystems [30]. With higher-order interaction modeled at the the level of groups of nodes, it is found that collective behaviors in neuroscience can be more faithfully predicted [46, 21] and essential nonlinearity emerges in diffusion processes [34].

Statistical physics and network science methods are originally devised to describe pairwise interactions. These methods have been generalized to include building blocks like small subgraphs and motifs to account for higher-order organization of complex networked systems [8]. Higher-order structures are often encoded in mathematical frameworks of hypergraphs, Petri nets and simplicial complexes. Hypergraphs [7, 12] extend the standard networks by allowing group interactions through hyperedges of different sizes such as pairs (2-tuple edges), triples (3-tuple edges), quadruples (4-tuple edges) etc. Petri nets are also known as finite state machines, which carry additional tokens and can be viewed as a type of directed hypergraphs [43]. In the simplicial complex approach, a filled clique of $k+1$ nodes for $k \geq 1$ is called a k-simplex [38]. A simplicial complex is formed by binding simplexes along their faces of any dimension. Employing their geometric interpretation and algebraic topology tools, simplicial complexes have played an central role in topological data analysis [40]. These higher-order structures are found to be instrumental in shaping varied dynamical processes such as spreading [29], social dynamics [3, 43], synchronization [21, 46] and random walks [33].

In parallel to the development of organization of complex networks, there has been a lot of research on degree-based topological indices, which are capable of characterizing network structure and dynamics with interdisciplinary applications across mathematics, chemistry, informatics and physics [15, 17, 23]. A recent addition to the long list of topological indices is the Sombor index [24], which has prompted a wave of research enthusiasm in a very short time. Building on a geometric approach to interpreting degree-based topological indices, Sombor index is introduced as the sum of degree radii of all edges in a graph. Fundamental properties and bounds of Sombor indices have been studied for different graphs; see e.g. [2, 13, 14, 16, 18, 25, 31, 36, 37, 50]. It is unraveled in [20] that Sombor index can characterize physicochemical properties of polycyclic aromatic compounds. With a good correlation with the Shannon entropy, Sombor index is proposed in [1] as a complexity measure for random graphs.

Motivated by the above lines of research, we here study the Sombor index of a simplicial network model, which is proposed as a generative network displaying a rich structure of cliques. In general, simplicial networks represent the underlying graph structure of clique complexes from the perspective of topology [28] and are a powerful tool in characterizing higher-order interactions in neural networks and learning algorithms [10, 19]. We analytically derive the degree distribution and clique distribution of our simplicial network model, and reveal that its degree follows a power law regulated by the model parameter. The degree exponent is found to sit within the interval $[2, 3]$ resembling many real-life networks [49]. We establish recursively the Sombor index of the simplicial networks and develop an approximate computation formula with a closed expression. Finally, extensive computational experiments are performed to further illustrate the topological properties (including degree, cliques, and distance) over some random variants of the simplicial network model. The approximation of Sombor index is shown to be good in all considered cases. When the initial network is chosen as an Erdős-Rényi random graph, we determine the right scaling constant for Sombor index.

The remainder of the paper is organized as follows. Sections 2 and 3 are devoted to the network model and its topological properties. Section 4 deals with the Sombor index. Computational studies are performed in Section 5 with a conclusion drawn in Section 6.

2. Simplicial network model

We consider a simplicial network model constructed iteratively from an initial graph $G(0) = (V(G(0)), E(G(0)))$ at time step $t = 0$, where $V(G(0))$ and $E(G(0))$ are the vertex set and the edge set, respectively. Assume the number of vertices is $n(0) := |V(G(0))| \geq 2$ and the minimum degree in $G(0)$ is at least 1, i.e., $d_{\min}(G(0)) \geq 1$. Hence, the number of edges $m(0) := |E(G(0))| \geq 1$. The network sequence $\{G(t)\}_{t\geq 1}$ is built by the following process:

Model A. The initial graph is $G(0)$. Given $t \geq 1$, for each edge $e \in E(G(t-1))$, let K_{r_e} be a r_e-clique associated with e, where $r_e \geq 1$. $G(t)$ is obtained by joining the two end vertices of e to every vertex of K_{r_e}, where $e \in E(G(t-1))$.

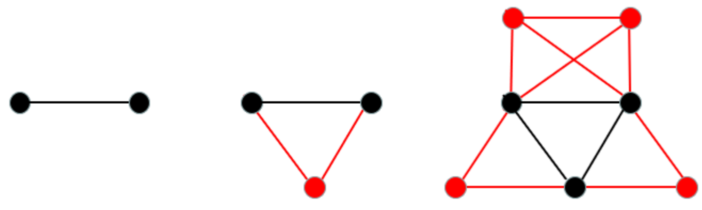

Figure 1: An example of Model A, where $G(0) = K_2$. In $G(t)$, the newly added vertices and edges at time t are coded in red. Here, $\tilde{n}(2) = 4$ and $\tilde{m}(2) = 9$.

Note that r_e is a function of time, i.e., r_e can be different when e is viewed as edges in $G(t_1)$ and $G(t_2)$ with $t_1 \neq t_2$. An example is shown in Fig. 1. Let $n(t) = |V(G(t))|$ and $m(t) = |E(G(t))|$ for $t \geq 0$. Moreover, the newly added numbers of vertices and edges are denoted by $\tilde{n}(t)$ and $\tilde{m}(t)$, respectively. Clearly, we have $\tilde{n}(0) = n(0)$ and $\tilde{m}(0) = m(0)$. For $t \geq 1$, we have

$$\tilde{n}(t) = \sum_{e \in E(G(t-1))} r_e \tag{1}$$

and

$$\tilde{m}(t) = \sum_{e \in E(G(t-1))} \left[\binom{r_e + 2}{2} - 1 \right] = \sum_{e \in E(G(t-1))} \frac{(r_e + 3) r_e}{2}. \tag{2}$$

It follows from (1) and (2) that

$$m(t) = \sum_{e \in E(G(t-1))} \binom{r_e + 2}{2} \quad \text{and} \quad n(t) = n(t-1) + \sum_{e \in E(G(t-1))} r_e. \tag{3}$$

The Model A belongs to a broad class of heterogenous network models [11, 45], which have been used, for example, in analyzing group interactions in social dynamics [44]. These models are highly flexible and versatile in that different attributes (such as weights, probabilities, types, actions etc.) can be attached to different edges. Given the model generality and the recursive formulas in (3), we will only be able to derive degree-related properties in the form of iteration. Therefore, we present a special type of model having a homogeneous (in terms of both time and space) expanding parameter r, which facilitates closed-form analytical expressions.

Model B. The initial graph is $G(0)$. Given $t \geq 1$ and $r \geq 1$, we take $m(t-1)$ copies of r-clique K_r. $G(t)$ is obtained by joining the two end vertices of the i-th edge in $G(t-1)$ to every vertex of the i-th copy of K_r, where $1 \leq i \leq m(t-1)$.

The network generation mechanism in Model B can be thought of as a edge corona product of the network and the complete graph K_r, namely, $G(t) = G(t-1) \diamond K_r$. Edge corona is an important graph operation that has been investigated intensively in spectral graph theory [26, 4, 51, 27, 32]. An example of Model B is shown in Fig. 2. It follows from (3) and $r_e \equiv r$ that $m(t) = \binom{r+2}{2} m(t-1)$ and $n(t) = n(t-1) + rm(t-1)$ for $t \geq 1$. Solving these recursive relationships, we obtain

$$m(t) = \binom{r+2}{2}^t m(0) \tag{4}$$

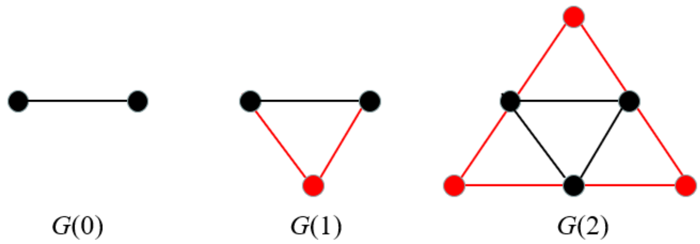

Figure 2: An example of Model B, where $G(0) = K_2$. In $G(t)$, the newly added vertices and edges at time t are coded in red. Here, $\tilde{n}(2) = 3$ and $\tilde{m}(2) = 6$.

and

$$
\begin{aligned}
n(t) =& n(0) + rm(0)\Big[1 + \binom{r+2}{2} + \cdots + \binom{r+2}{2}^{t-1}\Big] \\
=& n(0) + \frac{2m(0)}{r+3}\Big[\binom{r+2}{2}^{t} - 1\Big].
\end{aligned} \tag{5}
$$

The mean degree of $G(t)$ can be calculated as $\bar{d}(G(t)) := \frac{2m(t)}{n(t)}$, which tends to $r + 3$ as $t \to \infty$ by using (4) and (5). This indicates that Model B is sparse. Since the mean degree grows asymptotically linearly with the expanding parameter r, we know that Model A, squeezing between a minimum parameter $r_{\min} := \min_e r_e$ and a maximum parameter $r_{\max} := \max_e r_e$, is also sparse. Here, the edge e runs over all edge sets, i.e., $\cup_{t\geq 0} E(G(t))$.

3. Degree and clique distributions

In this section, we examine the distributions of vertex degrees and cliques for our models. For $t \geq 0$, denote by $d_v(t)$ the degree of vertex v at time t. Let t_v be the time step that the vertex v is born. By the construction of Model A, we have the following observation. For any vertex v and $t \geq t_v$, we have

$$
d_v(t+1) = \sum_{e \in \{e_1, e_2, \cdots, e_{d_v(t)}\}} (r_e + 1), \tag{6}
$$

where $\{e_1, e_2, \cdots, e_{d_v(t)}\}$ represents the set of edges incident to v at time t. For a vertex v with $t_v = 0$, namely, $v \in G(0)$, its initial degree is given by $d_v(0)$. On the other hand, for a vertex v with $t_v \geq 1$, its born degree is $d_v(t_v) = r_e + 1$, where e is associated with the clique K_{r_e} ($v \in K_{r_e}$) at time t_v. Combining these initial conditions with the recurrence relationship (6), the degree of any vertex v at time $t \geq t_v$ can be obtained.

In Model B, the relationship (6) reduces to $d_v(t+1) = d_v(t)(r+1)$ for $t \geq t_v$. Solving the geometric sequences, we derive

$$
d_v(t) = d_v(0)(r+1)^t, \qquad t \geq 0 \tag{7}
$$

for any vertex v with $t_v = 0$, and

$$
d_v(t) = (r+1)^{t-t_v+1}, \qquad t \geq t_v \tag{8}
$$

for any vertex v with $t_v \geq 1$.

Another important observation for the homogeneous Model B is that all vertices v with a fixed $T := t_v \geq 1$ share the same degree $d_v(t)$ given by (8) at any subsequent time $t \geq T$. The number of such vertices at time $t \geq T$ is $\tilde{n}(T)$, which is given by

$$\tilde{n}(T) = rm(T-1) = r\binom{r+2}{2}^{T-1} m(0) \tag{9}$$

employing (1) and (4).

Let $p(d,t)$ be the degree distribution of $G(t)$ in Model B, namely, the probability that a randomly chosen vertex in $G(t)$ is adjacent to d vertices. Define $F(d,t) = \sum_{s=d}^{\infty} p(s,t)$ to be the complementary cumulative degree distribution of $G(t)$.

Theorem 1. *For Model B,*

$$F(d,t) = \frac{-2m(0) + n(0)(r+3) + m(0)d^{1-\alpha}2^{-t}[(r+1)(r+2)]^{t+1}}{-2m(0) + n(0)(r+3) + m(0)2^{1-t}[(r+1)(r+2)]^{t}}. \tag{10}$$

where $\alpha := 2 + \frac{\ln(r+2)}{\ln(r+1)} - \frac{\ln 2}{\ln(r+1)}$. *Hence,* $p(d,\infty) \sim d^{-\alpha}$.

Proof. Define a number x satisfying $\min_{v \in V(G(0))} d_v(0) = (r+1)^x$. Hence,

$$x = \frac{\ln(\min_{v \in V(G(0))} d_v(0))}{\ln(r+1)}. \tag{11}$$

By our model assumption $d_{\min}(G(0)) \geq 1$ and $r \geq 1$, we obtain $0 \leq x < \infty$.

For any $v \in V(G(0))$, we have $d_v(t) \geq (r+1)^{t+x}$ for $t \geq 0$ in view of (7). Noting that the degrees of vertices in $G(t)$ are discrete, we set

$$d = (r+1)^{t-t_v+\min\{1,\lceil x \rceil\}}. \tag{12}$$

By the comments above Theorem 1 and considering two cases $x \geq 1$ and $0 \leq x < 1$, we obtain

$$F(d,t) = \frac{\sum_{s=0}^{t_v+1-\min\{1,\lceil x \rceil\}} \tilde{n}(s)}{n(t)}, \tag{13}$$

where d is given by (12). Using (9), the numerator in the expression (13) becomes

$$\begin{aligned} & n(0) + m(0) \sum_{s=1}^{t_v+1-\min\{1,\lceil x \rceil\}} r\binom{r+2}{2}^{s-1} \\ =& n(0) + \frac{2m(0)}{r+3}\Big[\binom{r+2}{2}^{t_v+1-\min\{1,\lceil x \rceil\}} - 1\Big]. \end{aligned} \tag{14}$$

Combining (14) with (5), (12) and (13), we obtain

$$F(d,t) = \frac{n(0)(r+3) + m(0)d^{1-\alpha}2^{-t}[(r+1)(r+2)]^{t+1} - 2m(0)}{n(0)(r+3) + m(0)2^{1-t}[(r+1)(r+2)]^{t} - 2m(0)}. \tag{15}$$

where $\alpha := 2 + \frac{\ln(r+2)}{\ln(r+1)} - \frac{\ln 2}{\ln(r+1)}$. Let t go to infinity and we derive $\lim_{t\to\infty} F(d,t) = \frac{(r+1)(r+2)}{2} d^{1-\alpha}$. Hence, $p(d,\infty) = F(d,\infty) - F(d+1,\infty) \sim d^{-\alpha}$. □

The degree of Model B follows asymptotically a power-law distribution with degree exponent α, which is between 2 and 3. This is consistent with many real-life scale-free networks [49]. Noting that α is an increasing function with respect to the parameter r, the construction of our networks indicates that Model A also follows a power-law distribution. An example is shown in Fig. 3 in Section 5 below.

Next, we investigate the clique distribution of the simplicial network and start with Model B. For $t \geq 0$, let $\omega_k(G(t))$ be the number of k-clique $G(t)$. Obviously $\omega_1(G(t)) = n(t)$ and $\omega_2(G(t)) = m(t)$, which are given by (5) and (4), respectively.

Theorem 2. *For Model B,*

$$\omega_k(G(t)) = \omega_k(G(0)) + \frac{2m(0)}{r(r+3)}\binom{r+2}{k}\Big[\binom{r+2}{2}^t - 1\Big], \quad k \geq 3. \tag{16}$$

Proof. Note that the combinatorics number $\binom{r+2}{k} = 0$ when $k > r+2$. Hence (16) holds for $k > r+2$ and $t \geq 0$. When $t = 0$, (16) is true by the model construction. What remains to show is the case for $t \geq 1$ and $3 \leq k \leq r+2$.

Fix any k satisfying $3 \leq k \leq r+2$. Based on the model construction, it is easy to see that any k-clique at step $t-1$ will always be present at time t and that each edge in $G(t-1)$ will generate a $(r+2)$-clique, in which $\binom{r+2}{k}$ new k-cliques will appear. Therefore,

$$\omega_k(G(t)) = \omega_k(G(t-1)) + m(t-1)\binom{r+2}{k} \tag{17}$$

for $t \geq 1$. Solving (17) in view of (4), we obtain

$$\omega_k(G(t)) = \omega_k(G(0)) + m(0)\binom{r+2}{k}\frac{2}{r(r+3)}\Big[\binom{r+2}{2}^t - 1\Big], \tag{18}$$

which concludes the proof. □

Denote by $\omega(G(t))$ the clique number of $G(t)$. Then $\omega(G(t)) = \max_{\omega_k(G(t))>0} k$, where $\omega_k(G(t))$ is given by (16). It is easy to see that $\omega(G(t)) = \max\{r+2, \omega(G(0))\}$.

Turning to Model A, a similar argument leads to the following recurrence relationship analogous to (17):

$$\omega_k(G(t)) = \omega_k(G(t-1)) + \sum_{e \in E(G(t-1))}\binom{r_e+2}{k}, \quad t \geq 1, \tag{19}$$

where the initial condition is the same $\omega_k(G(0))$. The clique number of $G(t)$ in Model A is

$$\omega(G(t)) = \max\{\max_{e \in \cup_{s=0}^{t} E(G(s))} r_e + 2, \omega(G(0))\}. \tag{20}$$

4. Sombor index

Recall that Sombor index is a degree-based graph invariant defined as [24]

$$SO(G(t)) = \sum_{e=\{u,v\} \in E(G(t))} \sqrt{d_u(t)^2 + d_v(t)^2}. \tag{21}$$

The initial Sombor index is $SO(G(0))$. The following result determines the Sombor index for Model B.

Theorem 3. *For Model B,*

$$
\begin{aligned}
SO(G(t)) = & \sqrt{2}m(0)\binom{r+2}{2}^{t-1}\binom{r}{2}(r+1) + SO(G(t-1))(r+1) \\
& + 2r(r+1)\Big[\sum_{v\in V(G(0))}\sqrt{1+d_v(0)^2(r+1)^{2(t-1)}} \\
& + \sum_{v\in V(G(t-1))\setminus V(G(0))}\sqrt{1+(r+1)^{2(t-t_v)}}\Big], \quad t\ge 1.
\end{aligned}
\tag{22}
$$

Proof. Fix time step $t \ge 1$. In each K_{r+2} created at time t, there is one edge (say $\{u,v\}$) which belongs to $G(t-1)$, and $2r$ new edges containing only one old vertex (i.e., u or v), and the remaining $\binom{r}{2}$ new edges with two newly added end vertices. Summing these contribution together, we derive

$$
\begin{aligned}
SO(G(t)) = & \sum_{\{u,v\}\in E(G(t-1))}\Big[\binom{r}{2}\sqrt{(r+1)^2+(r+1)^2} \\
& + \sqrt{(r+1)^2 d_u(t-1)^2 + (r+1)^2 d_v(t-1)^2} \\
& + r\sqrt{(r+1)^2+(r+1)^2 d_u(t-1)^2} + r\sqrt{(r+1)^2+(r+1)^2 d_v(t-1)^2}\Big] \\
= & m(t-1)\binom{r}{2}\sqrt{2(r+1)^2} + (r+1)\sum_{\{u,v\}\in E(G(t-1))}\sqrt{d_u(t-1)^2+d_v(t-1)^2} \\
& + r(r+1)\sum_{\{u,v\}\in E(G(t-1))}\Big[\sqrt{1+d_u(t-1)^2} + \sqrt{1+d_v(t-1)^2}\Big] \\
= & \sqrt{2}m(0)\binom{r+2}{2}^{t-1}\binom{r}{2}(r+1) + (r+1)SO(G(t-1)) \\
& + 2r(r+1)\sum_{v\in V(G(t-1))}\sqrt{1+d_v(t-1)^2},
\end{aligned}
\tag{23}
$$

where we have applied (4) in the last equality. Thanks to (7) and (8), the last term in (23) can be further decomposed as

$$
\begin{aligned}
\sum_{v\in V(G(t-1))}\sqrt{1+d_v(t-1)^2} = & \sum_{v\in V(G(0))}\sqrt{1+d_v(t-1)^2} + \sum_{v\in V(G(t-1))\setminus V(G(0))}\sqrt{1+d_v(t-1)^2} \\
= & \sum_{v\in V(G(0))}\sqrt{1+d_v(0)^2(r+1)^{2(t-1)}} \\
& + \sum_{v\in V(G(t-1))\setminus V(G(0))}\sqrt{1+(r+1)^{2(t-t_v)}}.
\end{aligned}
\tag{24}
$$

Feeding (24) into (23), we complete the proof. □

For Model A, using a similar idea, we can derive the Somber index by recursively invoking the following formula for $t \geq 1$:

$$
\begin{aligned}
SO(G(t)) = \sum_{e=\{u,v\}\in E(G(t-1))} & \left\{ \binom{r_e}{2} \sqrt{2(r_e+1)^2} \right. \\
& + \sqrt{\Big[\sum_{\tilde{e}=\{u,\tilde{u}\}\in E(G(t-1))} (r_{\tilde{e}}+1) \Big]^2 + \Big[\sum_{\tilde{e}=\{v,\tilde{v}\}\in E(G(t-1))} (r_{\tilde{e}}+1) \Big]^2} \\
& + r_e \sqrt{(r_e+1)^2 + \Big[\sum_{\tilde{e}=\{u,\tilde{u}\}\in E(G(t-1))} (r_{\tilde{e}}+1) \Big]^2} \\
& \left. + r_e \sqrt{(r_e+1)^2 + \Big[\sum_{\tilde{e}=\{v,\tilde{v}\}\in E(G(t-1))} (r_{\tilde{e}}+1) \Big]^2} \right\}. \qquad (25)
\end{aligned}
$$

Although the above formulas (22) and (25) produce the Sombor index for any $G(t)$, they do not give an straightforward estimate of the magnitude of Sombor index. Here, we present an approximation calculation method by underestimating the degrees of each old vertex in $G(t-1)$. In Model A, we simply assume the two end vertices of an edge $e = \{u, v\} \in E(G(t-1))$ have degree $q_e + 1$ in $G(t)$. This means all edges incident to u or v in $G(t-1)$ are neglected in the calculation of $SO(G(t))$. Since the network grows exponentially, we expect the approximation is close. This is illustrated in the simulations below. For Model A, we derive

$$
SO(G(t)) = \sum_{e\in E(G(t-1))} SO(K_{r_e+2}) = \sum_{e\in E(G(t-1))} \frac{(r_e+2)(r_e+1)^2}{\sqrt{2}}, \qquad (26)
$$

which is equivalent to the Sombor index of the graph $\cup_{e\in E(G(t-1))} K_{r_e+2}$. For Model B, in the light of (4), this reduces to

$$
SO(G(t)) = m(t-1) SO(K_{r+2}) = 2^{-t}\sqrt{2} m(0)(r+1)^{t+1}(r+2)^t. \qquad (27)
$$

To appreciate the expression (27) in the context of existing results of Sombor index, some remarks are in order.

Remark 1. It follows from (5) that

$$
n(t) - n(0) \sim \frac{2m(0)}{r+3} \binom{r+2}{2}^t. \qquad (28)
$$

Hence, (27) yields

$$
SO(G(t)) \sim \frac{(n(t)-n(0))(r+1)(r+3)}{\sqrt{2}} \sim n(t)\frac{(r+1)(r+3)}{\sqrt{2}} \qquad (29)
$$

for large t. It is shown in [16, Theorem 1] that a graph G over n vertices with minimum degree $d_{\min}$ and maximum degree $d_{\max}$ has the Sombor index

$$
\frac{n d_{\min}^2}{\sqrt{2}} \leq SO(G) \leq \frac{n d_{\max}^2}{\sqrt{2}}. \qquad (30)
$$

For large t, recall that $d_{\min}(G(t)) = r + 1$, mean degree $\bar{d}(G(t)) \sim r + 3$ and $G(t)$ is scale-free with a large maximum degree. The estimate (29) is in line with (30).

Remark 2. It is shown in [24, Theorem 2] that for any connected graph G over $n \geq 3$ vertices:

$$2\sqrt{5} + (2n-6)\sqrt{2} = SO(P_n) \leq SO(G) \leq SO(K_n) = \frac{n(n-1)^2}{\sqrt{2}}, \tag{31}$$

where P_n is a path over n vertices. By (29) we have the following limit

$$\lim_{t\to\infty} \frac{SO(G(t))}{SO(P_{n(t)})} = \frac{(r+1)(r+3)}{4}. \tag{32}$$

This suggests the structure of $G(t)$ is much closer to that of $P_{n(t)}$ as compared to the other extreme $K_{n(t)}$ of Sombor index $O(n(t)^3)$. This is in line with the sparse construction of $G(t)$.

On the other hand, by (32) we know

$$\lim_{t\to\infty} \frac{SO(G(t))}{SO(P_{n(t)})} = \lim_{t\to\infty} \frac{(r+1)\frac{2m(t)}{n(t)}}{4} = \frac{r+1}{2} \lim_{t\to\infty} \frac{m(t)}{|E(P_{n(t)})|}. \tag{33}$$

Note that the mean degree $\bar{d}(\cup_{m(t-1)\ \text{copies}} K_{r+2}) = r+1$ and $\bar{d}(P_{n(t)}) = 2$. Since the average Sombor index from a mean-field theory perspective [1] can be calculated as $\langle SO(G)\rangle = \sqrt{2}|E(G)|\bar{d}(G)$, the equation (33) can be reproduced by considering the average Sombor index.

Remark 3. It has been shown in Section 3 that the clique number of $G(t)$ in Model B is $\omega(G(t)) = \max\{r+2, \omega(G(0))\}$. With a sparse initial graph $G(0)$, we may assume $\omega(G(t)) = r+2$. In this case, using (27) we obtain

$$SO(G(t)) = \binom{r+2}{2}^t \sqrt{2}m(0)(r+1) = \binom{\omega(G(t))}{2}^t \sqrt{2}m(0)(\omega(G(t))-1). \tag{34}$$

It is shown in [18, Theorem 1] that any graph G over n vertices with clique number ω has Sombor index lower bounded by

$$SO(G) \geq SO(Ki_{n,\omega}) \sim \binom{\omega-1}{2}\sqrt{2}(\omega-1) + 2\sqrt{2}n + \sqrt{2}\omega^2. \tag{35}$$

where $Ki_{n,\omega}$ is a long kite graph over n vertices and with clique number ω as depicted in [18, Figure 1]. In view of (34), (5) and recall $\omega(G(t)) = r+2$, we have

$$\begin{aligned} SO(G(t)) =& \binom{\omega(G(t))}{2}\sqrt{2}m(0)(\omega(G(t))-1) \\ &+ \left[\binom{\omega(G(t))}{2}^t - \binom{\omega(G(t))}{2}\right]\sqrt{2}m(0)(\omega(G(t))-1) \\ \geq& \binom{\omega(G(t))-1}{2}\sqrt{2}(\omega(G(t))-1) + \frac{4\sqrt{2}m(0)}{r+3}\binom{r+2}{2}^t + \sqrt{2}(r+2)^2 \\ \sim& SO(Ki_{n(t),\omega(G(t))}) \end{aligned} \tag{36}$$

holds for $t \geq 3$ by direct calculations. Therefore, this agrees with [18, Theorem 1]. The inequality (35) is violated for the cases $t = 1$ and 2, which is due to the fact that (27) is only an approximation for $SO(G(t))$ from below.

Remark 4. Finally, as pointed out in [23, 24], there is a whole list of degree-based topological indices that can be represented in the form of $TI(G) = \sum_{\{u,v\}\in E(G)} \phi(d_u, d_v)$, where ϕ is a symmetric function. In the case of Sombor index, $\phi(\cdot,\cdot)$ is taken as $\phi(x, y) = \sqrt{x^2 + y^2}$. Although the exact calculation of $TI(G(t))$ may be not straightforward, the approximation method used in (27) can be applied analogously. For example, the first Zagreb index $Zg(G(t))$ of $G(t)$ in Model B may be estimated as follows:

$$Zg(G(t)) \sim m(t-1)Zg(K_{r+2}) = 2m(0)\binom{r+2}{2}^t (r+1). \tag{37}$$

5. Computational examples

We perform numerical computations in this section to illustrate the properties of our simplicial network models.

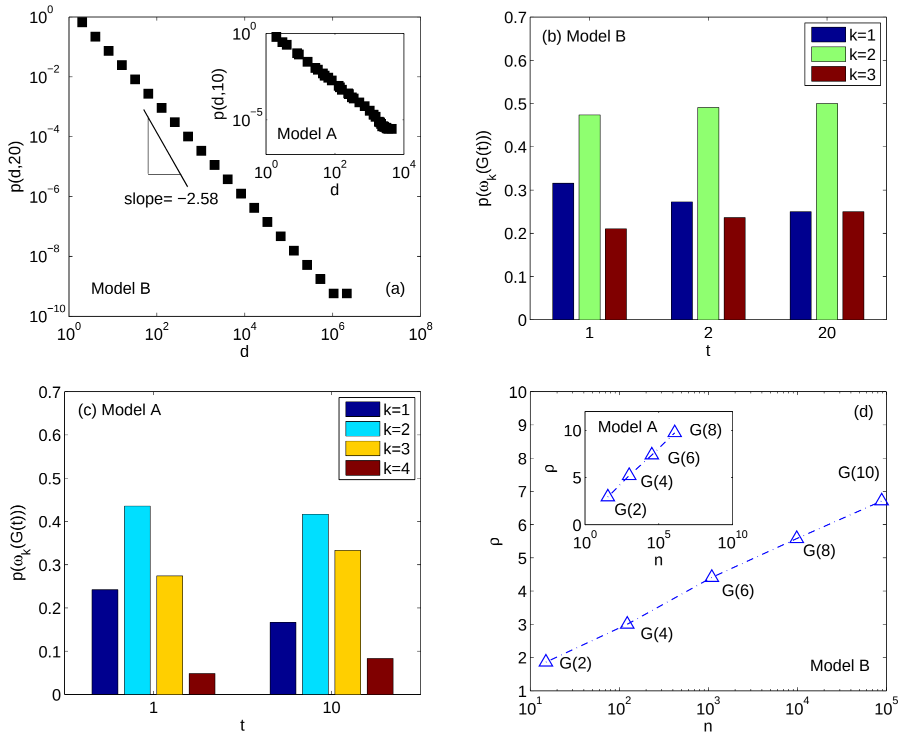

Figure 3: (a) The degree distribution of $G(t)$ for Model B in Example 1, where, $t = 20$, $G(0) = K_3$ and $r = 1$. The degree exponent as $t \to \infty$ is $\alpha = 2.58$. Inset: the degree distribution of $G(t)$ for Model A in Example 1, where $t = 10$, $G(0) = K_3$ and r_e is chosen from $\{1, 2\}$ uniformly at random at each step. (b) The distribution of number of k-cliques for the above Model B with $t = 1, 2$ and 20. (c) The distribution of number of k-cliques for the above Model A with different $t = 1$ and 10. The results are averaged over 20 independent implementations. (d) Average path length of $G(t)$ for the above Model B as a function of number of vertices. The inset shows the results for the above Model A. All data points are averaged over 50 random runs.

Example 1. First, we consider Model B with the initial network being $G(0) = K_3$, namely, a triangle. If we choose $r = 1$, then the model resembles a triangular tiling or tessellation in the

Euclidean plane [51]. In the main panel of Fig. 3(a) we show the degree distribution of the network for $G(20)$. The limit degree exponent is calculated as $\alpha = 2.58$ by Theorem 1. We then consider Model A by choosing the expanding parameter r_e from the set $\{1, 2\}$ at random independently for each edge e. In the inset of Fig. 3(a), we show the degree distribution for a randomly generated network $G(10)$, which again follows a power-law.

The corresponding distributions of numbers of k-cliques for these two models are shown in Fig. 3(b) and Fig. 3(c). This clique distribution is defined as

$$p(\omega_k(G(t))) = \frac{\omega_k(G(t))}{\sum_{k=1}^{\omega(G(t))} \omega_k(G(t))}. \tag{38}$$

The clique number for Model B is 3 and that for Model A is 4 as one would expect. We observe from Fig. 3(b) and Fig. 3(c) that the percentage of smaller cliques tends to decline and that of larger cliques tends to increase as the network grows. This feature of our simplicial network model is pertinent to the model selection for higher-order interactions in practical applications. It is revealed in [52] that only an appropriate profile of higher-order structures commensurate with the difficulty of the task would facilitate a solution in large-scale complex systems. In the context of public goods games for instance, the effect of different profiles of clique size have been investigated in [3].

We show in Fig. 3(d) the average path length $\rho(G(t))$ of $G(t)$ for the above Model A and Model B. The average path length of a graph is the average of shortest path lengths over the graph, which is a key performance metric for network topology [35]. We observe from Fig. 3(d) that ρ scales logarithmic with respect to the number of vertices for both models, which indicates a small-world effect of the network. Similar phenomenon has been observed for a related growing network model [41], where only new edges give birth to vertices.

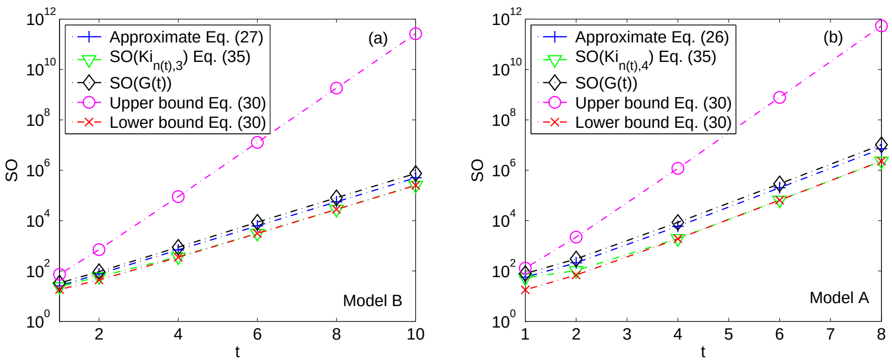

Figure 4: Sombor indices for (a) Model B and (b) Model A in Example 1. Blue pluses are approximation for the Sombor index using (26) and (27), green triangles are the Sombor indices for $Ki_{n(t),\omega(G(t))}$ using (35), black diamonds are the calculated Sombor index of $G(t)$ using (22) and (25), magenta circles and red crosses are upper and lower bounds, respectively, using (30). The data points are averaged over 20 random networks.

In Fig. 4 we present the numerical calculations for the Sombor index of $G(t)$ for the above setting of Model A and Model B. For both models we observe that the Sombor index for $G(t)$ is close to the lower bound in (30) and that of the minimum Sombor index for graphs with a given clique number (35). This is in line with the previous comments since $G(t)$ follows approximately a power-law distribution with the mean degree close to the minimum degree. It is worth noting that the approximate calculation approach proposed in Section 4 offers a good estimate of the actual Sombor index in all situations considered here.

Example 2. In this example we examine the influence of initial network for Sombor index. In contrast to Example 1, we here deal with large $G(0)$ by choosing $G(0) \sim G_{n(0),p}$ following the Erdős-Rényi random graph model [11]. Each pair of vertices in $G(0)$ are connected independently with link probability p. Erdős-Rényi random graphs are often used as the null model for studying structure and dynamics of complex networks [35]. They also form a vibrant branch of study in combinatorial probability.

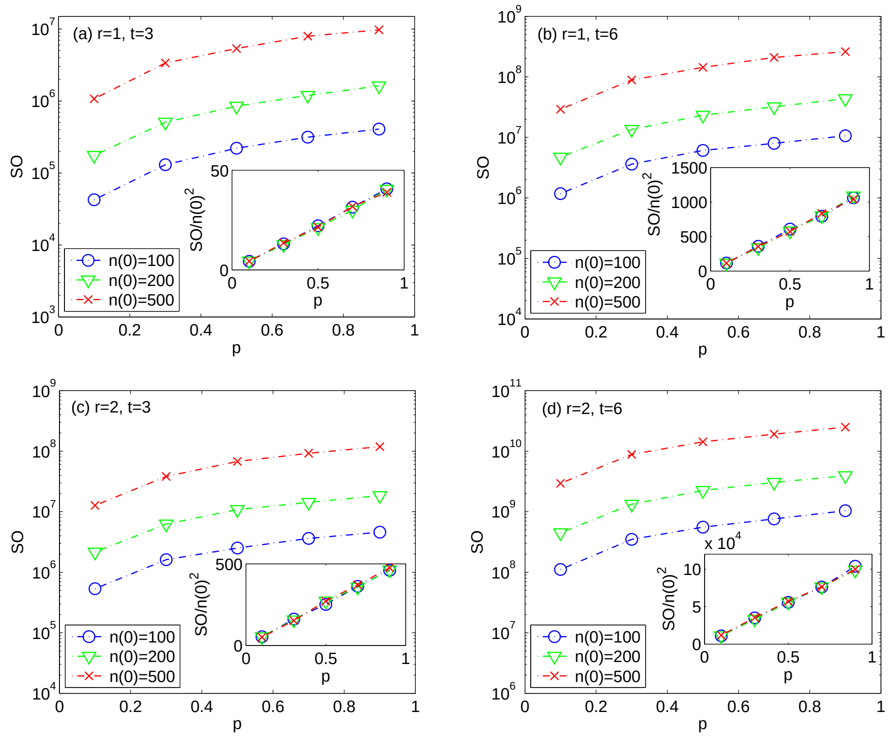

Figure 5: Sombor index of $G(t)$ in Example 2 for (a) $r = 1$ and $t = 3$, (b) $r = 1$ and $t = 6$, (c) $r = 2$ and $t = 3$, (d) $r = 2$ and $t = 6$. The initial network is an Erdős-Rényi random graph $G_{n(0),p}$. Rescaled Sombor index is shown in the insets. All data points are averaged over 50 random implementations.

We show in Fig. 5 the Sombor index of $G(t)$ in Model B with different expanding parameter r and time step t. The initial network $G(0)$ has size ranging from 100 to 500. In the insets of Fig. 5, we display the Sombor index scaled by the square of the order of $G(0)$, $n(0)^2$, for our simplicial network models. Remarkably, this normalization seems to work very well for different network sizes, link probabilities and expanding parameters. This phenomenon is reminiscent of the recent study in [1], where the order of the graph is found to be the correct normalization constant for several random graph models. The fact that the square works here is rooted in the edge-initiated expansion mechanism of our simplicial network model, where $n(0)^2p$ indicates the seed average edge number at the outset.

6. Conclusion

In this paper we have introduced a growing simplicial network model $G(t)$ featuring power-law distributions, small-world effect and rich clique structure that can be used for studying higher-order interactions in complex networked systems. We have analytically investigated the degree and clique distributions of $G(t)$ and the Sombor index $SO(G(t))$. A simple explicit calculation is proposed to estimate the Sombor index of $G(t)$ with good approximation. Through computational studies, it is found that $SO(G(t))$ scales well with the square of the order of $G(0)$ in the case of Erdős-Rényi random graphs. As directed networks often predict complex dynamical behaviors more faithfully than undirected networks, it is desirable to bring directedness to higher-order structures. An interesting research direction would be extending the simplicial network model to accommodate directed cliques [42].

References

[1] R. Aguilar-Sanchez, J. A. Mendez-Bermudez, J. M. Rodriguez, J. M. Sigarreta, Normalized Sombor indices as complexity measures of random networks. *Entropy*, 23(2021) 976

[2] S. Alikhani, N. Ghanbari, Sombor index of polymers. *MATCH Commun. Math. Comput. Chem.*, 86(2021) 715–728

[3] U. Alvarez-Rodriguez, F. Battiston, G. F. de Arruda, Y. Moreno, M. Perc, V. Latora, Evolutionary dynamics of higher-order interactions in social networks. *Nat. Hum. Behav.*, 5(2021) 586–595

[4] S. Barik, G. Sahoo, On the Laplacian spectra of some variants of corona. *Linear Algebra Appl.*, 512(2017) 32–47

[5] F. Battiston, E. Amico, A. Barrat, G. Bianconi, G. F. de Arruda, B. Franceschiello, I. Iacopini, S. Kéfi, V. Latora, Y. Moreno, M. M. Murray, T. P. Peixoto, F. Vaccarino, G. Petri, The physics of higher-order interactions in complex systems. *Nat. Physics*, 17(2021) 1093–1098

[6] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, G. Petri, Networks beyond pairwise interactions: structure and dynamics. *Phys. Rep.*, 874(2020) 1–92

[7] A. R. Benson, Three hypergraph eigenvector centralities. *SIAM J. Math. Data Sci.*, 1(2019) 293–312

[8] A. R. Benson, D. F. Gleich, J. Leskovec, Higher-order organization of complex networks. *Science*, 353(6295)(2016) 163–166

[9] J. Billings, M. Saggar, J. Hlinka, S. Keilholz, G. Petri, Simplicial and topological descriptions of human brain dynamics. *Netw. Neurosci.*, 5(2021) 549–568

[10] C. Bodnar, F. Frasca, Y. G. Wang, N. Otter, G. Montufar, P. Lió, M. M. Bronstein, Weisfeiler and Lehman go topological: Message passing simplicial networks. *Proc. 38th Int. Conf. on Machine Learning*, PMLR vol. 139, 2021, pp. 1026–1037

[11] B. Bollobás, S. Janson, O. Riordan, The phase transition in inhomogeneous random graphs. *Random Struct. Alg.*, 31(2007) 3–122

[12] P. S. Chodrow, Configuration models of random hypergraphs. *J. Complex Netw.*, 8(2020) cnaa018

[13] R. Cruz, J. Rada, Extremal values of the Sombor index in unicyclic and bicyclic graphs. *J. Math. Chem.*, 59(2021) 1098–1116.

[14] R. Cruz, J. Rada, J. M. Sigarreta, Sombor index of trees with at most three branch vertices. *Appl. Math. Comput.*, 409(2021) 126414

[15] R. Cruz, A. D. Santamaria-Galvis, J. Rada, Extremal values of vertex-degree-based topological indices of coronoid systems. *Int. J. Quantum Chem.*, 121(2021) e26536

[16] K. C. Das, A. S. Çevik, I. N. Cangul, Y. Shang, On Somber index. *Symmetry*, 13(2021) 140

[17] K. C. Das, S. Elumalai, S. Balachandran, Open problems on the exponential vextex-degree-based topological indices of graphs. *Disc. Appl. Math.*, 293(2021) 38–49

[18] K. C. Das, Y. Shang, Some extremal graphs with respect to Sombor index. *Mathematics*, 9(2021) 1202

[19] S. Ebli, M. Defferrard, G. Spreemann, Simplicial neural networks. *NeurIPS Workshop on Topological Data Analysis and Beyond*, Vancouver, Canada, 2020, arXiv:2010.03633

[20] X. Fang, L. You, H. Liu, The expected values of Sombor indices in random hexagonal chains, phenylene chains and Sombor indices of some chemical graphs. *Int. J. Quantum Chem.*, 121(2021) e26740

[21] L. V. Gambuzza, F. Di Patti, L. Gallo, S. Lepri, M. Romance, R. Criado, M. Frasca, V. Latora, S. Boccaletti, Stability of synchronization in simplicial complexes. *Nat. Commun.*, 12(2021) 1255

[22] J. Grilli, G. Barabás, M. J. Michalska-Smith, S. Allesina, Higher-order interactions stabilize dynamics in competitive network models. *Nature*, 548(2017) 210–213

[23] I. Gutman, Degree-based topological indices. *Croat. Chem. Acta*, 86(2013) 351–361

[24] I. Gutman, Geometric approach to degree-based topological indices: Sombor indices. *MATCH Commun. Math. Comput. Chem.*, 86(2021) 11–16

[25] B. Horoldagva, C. Xu, On Sombor index of graphs. *MATCH Commun. Math. Comput. Chem.*, 86(2021) 793–713

[26] Y. Hou, W.-C. Shiu, The spectrum of the edge corona of two graphs. *Electron. J. Linear Algebra*, 20(2010) 586–594

[27] S. Klavžar, M. Tavakoli, Dominated and dominator colorings over (edge) corona and hierarchical products. *Appl. Math. Comput.*, 390(2021) 125647

[28] W. Kook, K.-J. Lee, Simplicial networks and effective resistance. *Adv. Appl. Math.*, 100(2018) 71–86

[29] N. W. Landry, J. G. Restrepo, The effect of heterogeneity on hypergraph contagion models. *Chaos*, 30(2020) 103117

[30] J. M. Levine, J. Bascompte, P. B. Alder, S. Allesina, Beyond pairwise mechanisms of species coexistence in complex communities. *Nature*, 546(2017) 56–64

[31] H. Liu, L. You, Z. Tang, J. B. Liu, On the reduced Sombor index and its applications. *MATCH Commun. Math. Comput. Chem.*, 86(2021) 729–753

[32] A. Mahanta, I. J. Gogoi, A. Bharali, A note on the generalized adjacency, Laplacian and signless Laplacian spectra of the weighted edge corona networks. *Physica A*, 581(2021) 126241

[33] S. Mukherjee, J. Steenbergen, Random walks on simplicial complexes and harmonics. *Random Struct. Alg.*, 49(2016) 379–405

[34] L. Neuhäuser, A. Mellor, R. Lambiotte, Multibody interactions and nonlinear consensus dynamics on networked systems. *Phys. Rev. E*, 101(2020) 032310

[35] M. Newman, *Networks, 2nd Ed.*, Oxford University Press, Oxford, 2018

[36] J. Rada, J. M. Rodríguez, J. M. Sigarreta, General properties on Sombor indices. *Disc. Appl. Math.*, 299(2021) 87–97

[37] I. Redžepović, Chemical applicability of Sombor indices: Survey. *J. Serb. Chem. Soc.*, 86(2021) 445–457

[38] V. Salnikov, D. Cassese, R. Lambiotte, Simplicial complexes and complex systems. *European J. Phys.*, 40(2019) 014001

[39] M. T. Schaub, Y. Zhu, J.-B. Seby, T. M. Roddenberry, S. Segarra, Signal processing on higher-order networks: Livin' on the edge ... and beyond. *Signal Process.*, 187(2021) 108–149

[40] D. H. Serrano, D. S. Gómez, Centrality measures in simplicial complexes: Applications of topological data analysis to network science. *Appl. Math. Comput.*, 382(2020) 125331

[41] Y. Shang, Distinct clusterings and characteristic path lengths in dynamic small-world networks with identical limit degree distribution. *J. Stat. Phys.*, 149(2012) 505–518

[42] Y. Shang, Large dicliques in a directed inhomogeneous random graph. *Int. J. Comput. Math.*, 90(3)(2013) 445–456

[43] Y. Shang, Consensus formation in networks with neighbor-dependent synergy and observer effect. *Commun. Nonlinear Sci. Numer. Simulat.*, 95(2021) 105632

[44] Y. Shang, A note on the majority dynamics in inhomogeneous random graphs. *Results Math.*, 76(2021) 119

[45] C. Shi, Y. Li, J. Zhang, Y. Sun, P. S. Yu, A survey of heterogeneous information network analysis. *IEEE Trans. Know. Data Engin.*, 29(2017) 17–37

[46] P. S. Skardal, A. Arenas, Higher order interactions in complex networks of phase oscillators promote abrupt synchronization switching. *Commun. Phys.*, 3(2020) 218

[47] A. Tozzi, The multidimensional brain. *Phys. Life Rev.*, 31(2019) 86–103

[48] E. Vasilyeva, A. Kozlov, K. Alfano-Bittner, D. Musatov, A. M. Raigorodskii, M. Perc, S. Boccaletti, Multilayer representation of collaboration networks with higher-order interactions. *Sci. Rep.*, 11(2021) 5666

[49] I. Voitalov, P. van der Hoorn, Remco van der Hofstad, D. Krioukov, Scale-free networks well done. *Phys. Rev. Research*, 1(2019) 033034

[50] Z. Wang, Y. Mao, Y. Li, B. Furtula, On relations between Sombor and other degree-based indices. *J. Appl. Math. Comput.*, (2021) doi: 10.1007/s12190-021-01516-x

[51] Y. Wang, Y. Yi, W. Xu, Z. Zhang, Modeling higher-order interactions in complex networks by edge product of graphs. *Computer J.*, (2021) doi: 10.1093/comjnl/bxab070

[52] S. Yoon, H. Song, K. Shin, Y. Yi, How much and when do we need higher-order information in hypergraphs? A case study on hyperedge prediction. *Proc. The Web Conference (WWW'20)*, Taipei, Taiwan, 2020, pp. 2627–2633