Received: 23 July 2020 | Revised: 10 March 2021 | Accepted: 23 April 2021

DOI: 10.1002/wics.1566

OVERVIEW

WIREs COMPUTATIONAL STATISTICS WILEY

# Community detection in complex networks: From statistical foundations to data science applications

**Asim K. Dey**[1,2] | **Yahui Tian**[3] | **Yulia R. Gel**[2]

[1]Department of Electrical and Computer Engineering, Princeton University, Princeton, New Jersey, USA

[2]Department of Mathematical Sciences, University of Texas at Dallas, Richardson, Texas, USA

[3]Department of Biostatistics and Data Sciences, Boehringer Ingelheim, Ridgefield, Connecticut, USA

**Correspondence**
Yulia R. Gel, Department of Mathematical Sciences, University of Texas at Dallas, Richardson, TX 75080, USA.
Email: ygl@utdallas.edu

**Funding information**
National Science Foundation, Grant/Award Numbers: DMS 1925346, DMS 1736368, IIS 1633331

**Edited by:** James E. Gentle and David W. Scott, Editors-in-Chief

**Abstract**

Identifying and tracking community structures in complex networks are one of the cornerstones of network studies, spanning multiple disciplines, from statistics to machine learning to social sciences, and involving even a broader range of application areas, from biology to politics to blockchain. This survey paper aims to provide an overview of some most popular approaches in statistical network community detection as well as the newly emerging research directions such as community extraction with higher-order features and community discovery in multilayer and multiscale networks. Our goal is to offer a unified view at methodological interconnections and the wide spectrum of interdisciplinary data science applications of network community analysis.

This article is categorized under:

Data: Types and Structure > Graph and Network Data

Statistical Learning and Exploratory Methods of the Data Sciences > Clustering and Classification

**KEYWORDS**

classification, clustering, community detection, complex networks, multilayer and multiscale networks, network motifs

## 1 | INTRODUCTION

Community detection is one of the earliest and still most actively developing areas in complex network analysis, attempting to shed an important light into intrinsic network organization. While the first methods for network community detection go back to Kernighan and Lin (1970), Tarjan (1971), Bron and Kerbosch (1973), and Zachary (1977), nowadays the area continues to attract the attention of researchers and data scientists in a broad range of disciplines, from monitoring pre-election political communication, racial protest rhetoric on Twitter, and hidden trendsetters in online media (Helal et al., 2017; Matuszewski & Szabó, 2019; Tien et al., 2020; Weaver et al., 2018) to identifying parts of the brain potentially affected by marijuana use (Brumback et al., 2016; Filbey et al., 2014; Hurd et al., 2019), to tracking whale clubs, and grouping malicious addresses on blockchain (Akcora, Li, et al., 2020; Goldsmith et al., 2020). Such a very diverse set of methods and applications of network community detection has resulted in a wide variety of surveys, with some providing a broader overview while others focusing on a particular discipline, for example, social sciences or physics (see Table 1).

Given the current vast and constantly increasing knowledge on community detection, the current paper as well as any other survey cannot be considered as an exhaustive overview of the subject. The contributions of our paper are multifold and can be summarized as follows. First, there still exist noticeably fewer review papers addressing community

**TABLE 1** A (non-exhaustive) list of survey papers on network community

| Review paper | Journal | Research area |
|---|---|---|
| Fortunato (2010) | Physics Reports | A detail review of methods with special attention to the contributions made by physicists |
| Malliaros and Vazirgiannis (2013) | Physics Reports | Communities in directed networks |
| Fortunato and Hric (2016) | Physics Reports | A user guide of community detection in networks (methods mostly within statistical physics) |
| Bedi and Sharma (2016) | WIREs Data Mining & Knowledge Discovery | Communities in social networks |
| Zhao (2017) | WIREs: Computational Statistics | Theoretical advances of community detection in networks |
| Abbe (2018) | JMLR | Evolution of SBM for community detection |
| Rossetti and Cazabet (2018) | CSUR | Community discovery in dynamic networks |
| Lee and Wilkinson (2019) | Applied Network Science | SBM and extensions |
| Dakiche et al. (2019) | Information Processing & Management | Tracking community evolution in social networks |

detection from a statistical perspective which constitutes our primary focus. Second, our goal is to overview and highlight some emerging interdisciplinary research directions in community detection such as methods based on higher-order network structures, for example, *network motifs*, community detection in multilayer and multiscale networks, and extracting communities in dynamic networks. Finally, we strive to offer an up-to-date overview of the approaches and algorithms till 2020.

The remainder of this paper is organized as follows. In Section 2 we provide a background on complex networks and the notion of community. We then proceed to the most widely used benchmark models in statistical community detection, namely, stochastic block models (SBMs) (see Section 3) and Kronecker product graph model (KPGM) (see Section 4). We discuss such emerging directions in community detection as methods based on higher-order network structures and analysis of communities in multilayer and multiscale networks in Sections 5 and 6. Other popular community detection methods such as modularity-based, hierarchical, dynamic, and spectral clustering approaches are discussed in Section 7. Section 8 is devoted to community detection in dynamic networks. Section 9 describes future directions of network clustering. Finally, the paper is concluded by a discussion in Section 10.

## 2 | BACKGROUND ON GRAPHS AND NETWORK COMMUNITY DETECTION

Let $G = (V, E, \omega)$ be an undirected graph,[1] where $V$ is the set of nodes and $E \subset V \times V$ is the set of edges. The total number of nodes in $G$ is, $n = |V|$. Here $\omega : V \times V \mapsto \mathbb{R}$ is an *edge weight* function such that each edge $e_{uv} \in E$ has a weight $\omega_{uv}$. The corresponding $n \times n$ adjacency matrix $A$ is defined by

$$A_{ij} = \begin{cases} \omega_{ij}, & \text{if } (i,j) \in E \\ 0, & \text{otherwise.} \end{cases}$$

For an unweighted graph, the $n \times n$ adjacency matrix $A$ is defined by $A_{ij} = 1$ if $(i, j) \in E$ and 0 otherwise. Let $D$ be a diagonal matrix of (weighted) node degrees, that is, $d_i = D_{ii} = \sum_{j=1}^{n} A_{ij}$.

While there exists no unique, uniformly defined notion of the network community, *community*, or *cluster*, is conventionally understood as a subset of nodes which tend to be more "similar" to each other (e.g., more densely connected within a group) as compared to the rest of the network. Broadly speaking, under scenarios when no training

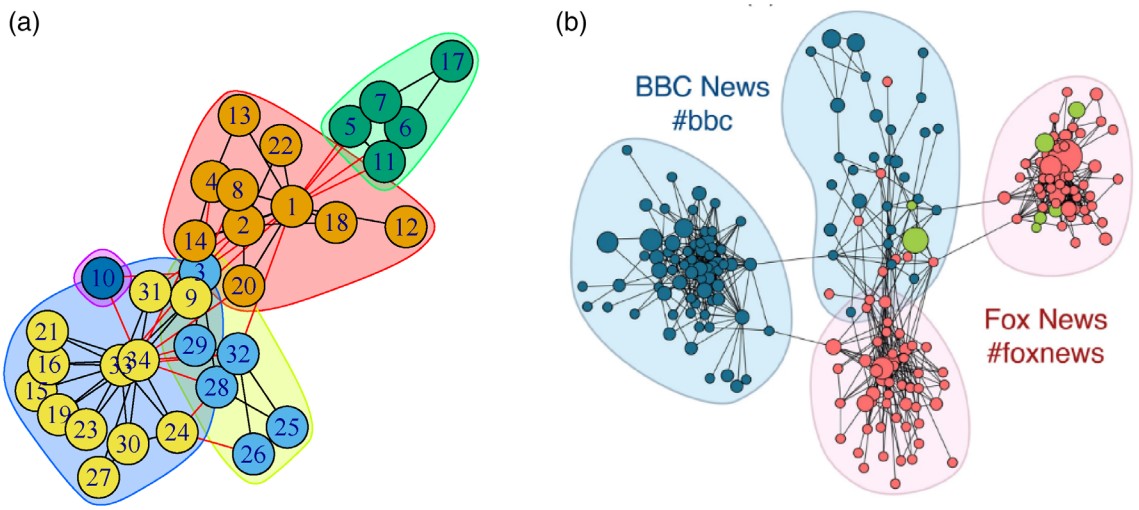



**FIGURE 1** Communities in social interaction networks. (a) Community structure in Zachary's Karate Club network. (b) Community structure among Twitter users sharing the hashtags #BBC and #FoxNews. Reprinted (Twitter) figure with permission from Weng et al. (2013)

data are available (i.e., in unsupervised settings), the process of community detection is primarily focused on partitioning nodes into groups (i.e., clustering), while a number of clusters may be either given a priori or be unknown and, hence, to be estimated. In turn, community detection in supervised settings (i.e., in cases where there exists some training data with labeled ground truth), the process of community detection primarily targets classifying (i.e., predicting) community membership of the remaining nodes.

Figure 1 shows community structures in one of the most widely used benchmark datasets in complex network analysis, that is, Zachary's Karate Club network (Zachary, 1977), as well as the Twitter network of media user preferences (Weng et al., 2013). Zachary's Karate Club describes the relationships in a university karate club. Figure 1a demonstrates an example of community structure in Zachary's Karate Club with five communities. Communities among Twitter users (Figure 1b) share the hashtags #BBC and #FoxNews. Blue nodes are #BBC users, red nodes are #FoxNews users, and users who have used both hashtags are green. Node size is proportional to tweet activity, links represent mutual following relations. Communities in both networks tend to exhibit denser internal links and noticeably sparser connections in-between communities.

## 3 | STOCHASTIC BLOCK MODELS

Community detection methods developed in different disciplines, for example, social sciences, physics, biology, engineering, and computer science often rely on disparate domain knowledge interpretation and varying validation metrics. However, ideally, a "good" community detection algorithm should be not only validated empirically but also demonstrate a reliable theoretical basis. Statistical methods for extracting communities in complex networks aim to address this question by revealing the latent structure of the network based on model fitting and the associated statistical inference (Bickel & Sarkar, 2016; He et al., 2020; Lancichinetti et al., 2011; Saldaña et al., 2017; Sanna Passino & Heard, 2020).

In this context, SBMs and their different extensions are arguably the most popular statistical models in network community detection, going back to Holland et al. (1983) and Bui et al. (1984). SBM is a generative model that assigns a probability value to each node pair $i, j$ in the network and encodes specific assumptions on different latent or unknown parameters and their interaction to create edges. SBM parameters can be interpreted with respect to certain hypotheses about the network structure, for example, assortativity, or tendency of nodes to connect to similar nodes as the case in many social interaction networks. Furthermore, the SBM framework allows to use likelihood scores to compare different parameterizations or different models. The SBM framework also enables to estimate missing or future structures, based on a partial or past observation of network. Figure 2 shows a schematic relationship between the observed network and SBM framework.

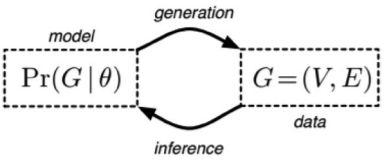



**FIGURE 2** Stochastic block model (SBM) defines a probability distribution over networks $Pr(G|\theta)$, where $\theta$ is the SBM's parameters. If the values of $\theta$ are chosen one can draw instances from the corresponding distribution. In turn, for a given network $G$, the values of $\theta$ can be estimated using statistical inference that best explain or reproduce the observed pattern of connectivity. Reprinted figure with permission from Clauset (2013)

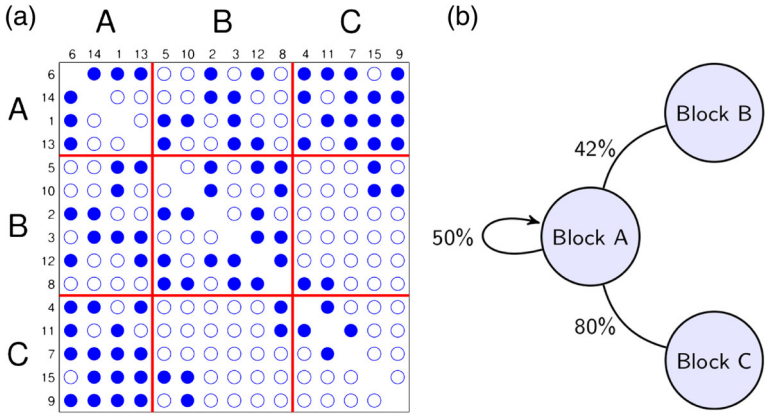



**FIGURE 3** (a) Block decomposition of an undirected network on 15 nodes (numbered from 1 to 15), where the blue circles mark the presence of connection, the empty circles mark the absence of connection. (b) Stochastic representation of block A in the network's block decomposition. In block A, each node having a 50% chance to form an edge with another node in Block A (stochastically equivalence), a 42% chance to form an edge with another node in block B, and an 80% chance to form an edge with another node in block C. Reprinted figure with permission from Pavlović et al. (2014)

Conventionally, as input SBM takes a set of parameters $\theta = \{\mathbf{k}, Z, M\}$ defined as follows:

- **k:** Each node $i \in V$ of a graph $G = (V, E)$ is assigned one of $k$ communities, say $\{C_1, C_2, ..., C_k\}$.
- $Z$: The class of each node is defined by a discrete latent variable $Z = \{z_1, z_2, ..., z_n\}$, a $n \times 1$ vector, where $z_i = C_l$ if node $i$ belongs to class $C_l$, where, $l = 1, 2, ..., k$.
- $M$: A $k \times k$-(stochastic) block matrix, where $M_{\ell r}$ is the probability that a node in group $\ell$ is connected to a node in group $r$.

Any two nodes $i \in C_\ell$ and $j \in C_r$ are connected independently by an edge following Bernoulli distribution with probability $M_{\ell r}$, $i, j \in \{1, 2, ..., n\}$, $\ell, r \in \{1, 2, ..., k\}$. That is,

$$A_{ij} \big| \left(z_i = C_l z_j = C_r\right) \sim Bernoulli(M_{\ell r}). \tag{1}$$

Each node in a given group connects to all other nodes in the same way. Hence, nodes with the same label are called *stochastically equivalent*. That is, the SBM generates *Erdös–Ŕenyi* random graphs within each community $C_\ell$, with an internal probability given by $M_{\ell\ell} = M_\ell$. If $M_{\ell r} \equiv p$ for all nodes $i, j \in V$, then SBM reduces to the Erdös–Ŕenyi random graph model $G(n, p)$. Figure 3 depicts an example of the relationship between three different groups (blocks).

For a given $k$ and an observed network $G$, SBM can be used to infer the latent community assignments $Z$ and stochastic block matrix $M$. Under the Bernoulli assumption (Equation (1)) and given $Z$ and $M$, the likelihood of the data can be written as

$$\mathscr{L}(G|M,Z) = \prod_{(i,j)\in E} M_{\ell r} \prod_{(i,j)\notin E} (1 - M_{\ell r}). \tag{2}$$

Sometime right-hand side of Equation (2) is written as $\prod_{(i,j)\in E} M_{z_i,z_j} \prod_{(i,j)\notin E} \left(1 - M_{z_i,z_j}\right)$. Parameters of SBM and its variants

are typically estimated either using a class of spectral methods (see Section 7.4) or various maximum likelihood (Amini

et al., 2013; Celisse et al., 2012; Nowicki & Snijders, 2001; Wang & Bickel, 2017) and modularity optimization techniques such as semidefinite programming (SDP) (Amini & Levina, 2018; Cai & Li, 2015; Chen et al., 2018; Guédon & Vershynin, 2016; Makarychev et al., 2016; Moitra et al., 2016; Qian et al., 2019; Zhao et al., 2012).

There have been developed many SBM extensions, falling under the general framework of the SBM family. One such notable and widely used SBM modification is the class of degree-corrected stochastic block models (DC-SBM) (Karrer & Newman, 2011). In the standard SBM, given $Z$ and $M$, the expected degree for all nodes in the group is the same. Hence, if the network has a skewed degree distribution, the maximum likelihood partitioning under the SBM framework tends to group vertices by degree. The DC-SBM approach modifies the model in a way that allows nodes to have arbitrary degrees, without having to force the combination of $z$ and $M$ to produce them. Furthermore, DC-SBM adds a new parameter $\gamma_j$ to each node that controls the expected degree of node $j$. The number of edges between a pair of nodes $i$ and $j$ in the DC-SBM follows a Poisson distribution with mean $\gamma_i\gamma_j M_{\ell r}$. Hence, the likelihood function becomes

$$\mathscr{L}(G|\gamma,M,Z)=\prod_{i,j}Poisson\left(\gamma_i\gamma_j M_{\ell r}\right).$$

Figure 4a shows a standard SBM places the five highest-degree nodes in one group and all other nodes in the other group. However, DC-SBM (see Figure 4b) yields an inferred division that appears to be noticeably closer to the truth.

Another popular extension of SBM is a weighted stochastic block model (WSBM). WSBM is a generalization of the standard SBM which allows for learning both from edge-existence and edge-weight information. In particular, WSBM accounts for the adjacency matrix $A=\{A_{ij}\}\in\{0,1\}$ as well as for $d$-dimensional edge covariates $X=\{x_{ij}\}$ on the edges, where $X$ follows an exponential family distribution. The likelihood function of WSBM has a form of an exponential family

$$\mathscr{L}(G|M,Z)\propto\exp\left(\sum_{ij}T\left(A_{ij}\right).\eta(M_{\ell r})\right),$$

where $T$, $\eta$ are fixed mappings, and $M_{\ell r}$ is the exponential family parameter (Aicher et al., 2014; Peixoto, 2018; Vu et al., 2013). The WSBM framework has found its place in a broad range of diverse applications, from biology (Faskowitz et al., 2018) to social sciences (Peixoto, 2018; Yenerdag, 2016).

To model network structures where nodes possibly belong to multiple communities, we can employ such SBM modifications as, for instance, mixed membership SBM of Airoldi et al. (2008) and Li et al. (2016)) and overlapping stochastic block model of Latouche et al. (2011), Anandkumar et al. (2014), and Peixoto (2015). Finally, Chien et al. (2018) propose a d-wise hypergraph stochastic block model which extends SBM from graphs to d-uniform hypergraphs.

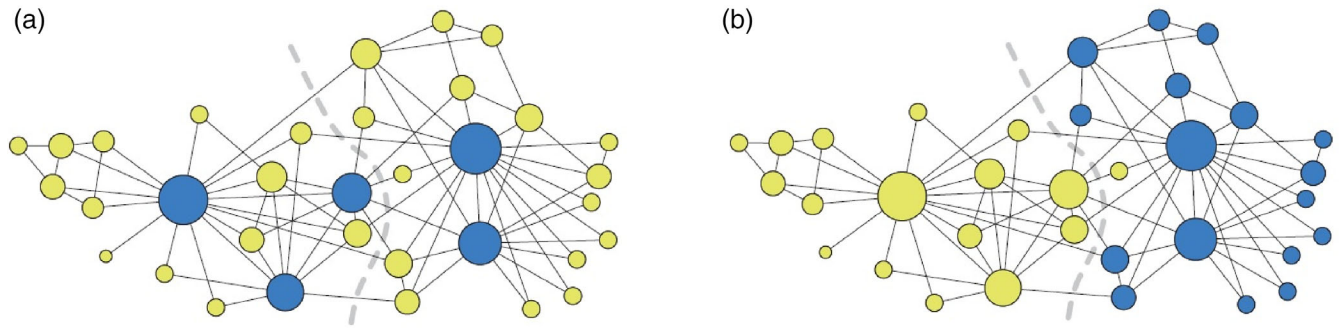



**FIGURE 4** Divisions of the karate club network based on estimated stochastic block model (SBM) with $r=2$. (a) Uncorrected SBM and (b) corrected SBM. The size of each node is proportional to its degree. The node color reflects inferred group membership and the dashed line indicates the split observed in real life (truth). Reprinted figure with permission from Karrer and Newman (2011)

The SBM and its extensions are arguably the primary choices for benchmark model for validating new community detection methods (Funke & Becker, 2019; Ghasemian et al., 2019; Lancichinetti & Fortunato, 2009). The SBM family also remains a very actively developing research area in network science and network applications, with some most recent additions including SBM for multilayer complex networks (see Section 6), hierarchical SBM (see Section 7.2), and dynamic SBMs (see Section 8).

## 4 | KRONECKER PRODUCT GRAPH MODEL

KPGM is a generative model, which uses *Kronecker product* operation and creates self-similar graphs, recursively (Leskovec et al., 2005; Leskovec & Faloutsos, 2007). The KPGM produces networks where different properties of real networks, for example, small diameters, heavy tails degree distribution, heavy tails for the eigenvalues and eigenvectors, *naturally emerge*. In addition, KPGM is a mathematically tractable model that allows for rigorous analysis of network properties (Leskovec, 2009; Leskovec et al., 2010).

The algorithm begins with an initiator graph $K_1$, with $N_1$ nodes and $E_1$ edges, and then recursively produces larger graphs $K_2$, $K_3$, ... such that the $k$th graph $K_k$ is on $N^k = N$ nodes. If $G_1$ and $G_2$ are graphs with adjacency matrices $A(G_1)$ and $A(G_2)$ respectively, then the Kronecker product $G_1 \otimes G_2$ is defined as the graph with adjacency matrix $A(G_1) \otimes A(G_2)$ (Weichsel, 1962). Edges in Kronecker-product graph can be defined as Edge $(X_{ij}, X_{kl}) \in G_1 \otimes G_2$ iff $(X_i, X_k) \in G_1$ and $(X_j, X_l) \in G_2$, where $X_{ij}$, $X_{kl}$ are nodes in $G_1 \otimes G_2$, and $X_i$, $X_j$, $X_k$, and $X_l$ are the corresponding nodes in $G_1$ and $G_2$. Figure 5 depicts an example of Kronecker graph formation.

The stochastic version of KPGM is the *stochastic Kronecker graph model*, where each entry of the (stochastic) adjacency matrix gives the probability of that particular edge appearing in the graph (Leskovec et al., 2010; Seshadhri et al., 2013). Extensions of KPGM also include, for instance, a tied KPGM which accounts for variance in network populations (Moreno et al., 2010), a block-KPGM capturing a fractal structure observed in many real world networks (Moreno et al., 2013), an enhanced version of the KPGM which uses random noise for smoothing (Seshadhri et al., 2013), and a mixture-model based KPGM (Mahapatra & Chandola, 2015). Recently, Hao et al. (2019) study the controllability of KPGM, while Ma et al. (2019) investigate the existence of *restricted edge connectivity*, that is, the minimum number of edges whose deletion disconnects the graph, of KPGM.

Furthermore, Kim and Leskovec (2010, 2012) propose a generalization of KPGM, called multiplicative attribute graph model (MAGM). As shown by Kim and Leskovec (2011), MAGM tends to reflect the power-law degree distribution and some higher-order graph statistics of real-world networks more accurately than KPGM. Pfeiffer et al. (2014) introduce an extension of MAGM which models a network structure with correlated attributes. In turn, Radcliffe and Young (2014) study theoretical properties, in particular, asymptotic bounds on the spectra of the adjacency matrix and the normalized Laplacian matrix of MAGM. Finally, recently Qu and Makowski (2019) explore the existence or absence of isolated nodes in MAGM.

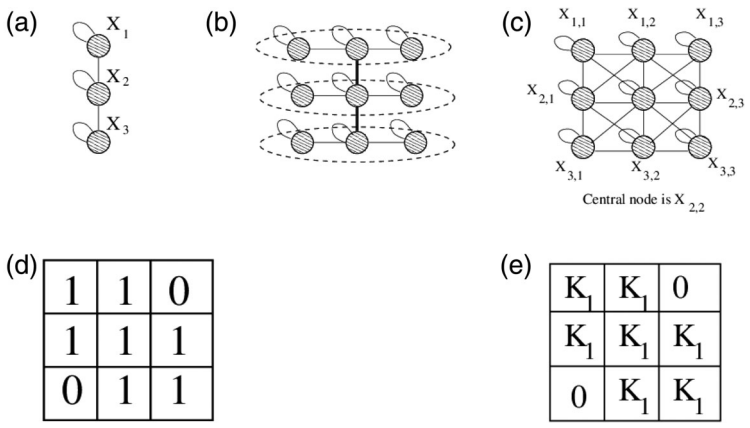



**FIGURE 5** Top row shows an initiator graph ($K_1$) and its Kronecker product with itself. (a) Graph $K_1$. (b) Intermediate stage. (c) Graph $K_2 = K_1 \otimes K_1$. Bottom row represents corresponding adjacency matrices. (d) Adjacency matrix of $K_1$. (d) Adjacency matrix of $K_2 = K_1 \otimes K_1$. Reprinted figure with permission from Leskovec et al. (2010)

## 5 | CLUSTERING BASED ON HIGHER-ORDER NETWORK STRUCTURES

Most currently available methods for the analysis of complex network organization and functionality tend to be based on lower-order pairwise connectivity and consider the level of individual nodes and edges. However, recent studies indicate that higher-order network substructures, for example, network *motifs* tend to play a critical role in understanding organization and functionality of complex networks (Ahmed et al., 2016; Li & Milenkovic, 2017; Milo et al., 2002; Pržulj, 2007; Stone et al., 2019). A *motif* is broadly defined as a recurrent multi-node subgraph pattern in a network (Milo et al., 2002), and in case of induced subgraphs motifs are often referred to as *graphlets* (Pržulj, 2007). Both motifs and graphlets have been originally studied in conjunction with cellular networks in bioinformatics (Milo et al., 2002; Pržulj, 2007; Shen-Orr et al., 2002). Later the role of motif signatures has been assessed in the context of many disparate network systems, including resilience quantification of power grid networks (Dey et al., 2019; Menck et al., 2014; Schultz et al., 2014), transportation systems (Agasse-Duval & Lawford, 2018), and protein interactomes Zitnik et al. (2019). Most recently, network motifs have demonstrated a high utility in describing hidden mechanisms behind cryptocurrency trading dynamics (Abay et al., 2019; Dey et al., 2020) and money laundering on blockchain (Ranshous et al., 2017). Figure 6 shows all 13 connected three-node directed motifs.

As noted by Watts and Strogatz (1998), higher-order network structures may often demonstrate a stronger signal of the community existence than edges alone. Rather than using only network edge information, Benson et al. (2016) and Tsourakakis et al. (2017) introduce community detection methods based on network motifs within a spectral clustering framework (see Section 7.4). For a motif $M$ of interest, the method first forms a motif adjacency matrix $W_M$, whose $(i, j)$ entries, $W_M ij$, are the co-occurrence counts of nodes $i$ and $j$ in the motif $M$, that is, the number of instances of $M$ that contain nodes $i$ and $j$. The next step is to compute the spectral ordering ($\sigma$) of the nodes from the normalized Laplacian of $W_M$. For a given motif $M$ the aim is then to detect a cluster defined by $S$ that minimizes the ratio as

$$\phi_M(S) = \frac{cut_M(S, \bar{S})}{\min\{vol_M(S), vol_M(\bar{S})\}}, \tag{3}$$

where $\bar{S}$ is the complement of $S$, $cut_M(S, \bar{S})$ is the number of instances of motif $M$ with at least one node in $S$ and one in $\bar{S}$, and $vol_M(S)$ is the number of nodes in instances of $M$ that reside in $S$ (Benson et al., 2016). Figure 7 shows a toy example of the clustering method based on directed motif $M_7$.

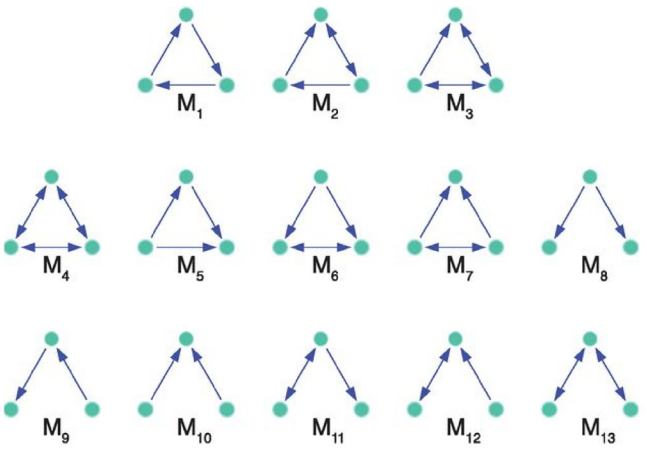



**FIGURE 6** All 13 connected three-node directed motifs in a network. Reprinted figure with permission from Benson et al. (2016)

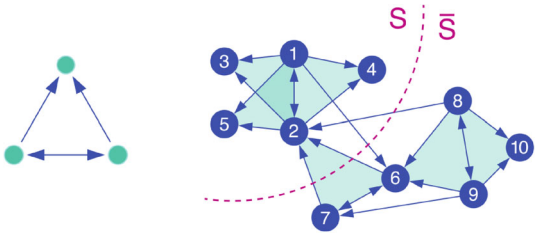



**FIGURE 7** Clustering of a network based on directed motif $M_7$. Reprinted figure with permission from Benson et al. (2016)

Nowadays, community detection with higher-order graph structures is one of the most actively evolving research directions (Chodrow & Mellor, 2020; Ribeiro et al., 2019), largely due to its high applicability in such emerging areas as tracking trading patterns and e-crime detection on blockchain (Akcora, Li, et al., 2020; Jourdan et al., 2018; Ranshous et al., 2017), assessment of organization and functionality of brain connectome (Battiston et al., 2017; Märtens et al., 2017; Martinet et al., 2020; Meier et al., 2017), and analysis of the topology of transcriptional regulatory networks (Roy et al., 2020). More recent methodological developments on community detection with higher-order network structures include, for instance, Li et al. (2017) who analyze motif based clustering for social networks and Li and Milenkovic (2017) who discuss inhomogeneous hypergraph clustering for learning hidden high-order network organization patterns. In turn, Huang et al. (2019) introduce a community detection method for multilayer networks based on multilayer motifs (M-motif). Finally, Li, Huang, et al. (2020) propose a motif-based label propagation method for community detection, while Mémoli and Pinto (2020), armed with the notion of motifs, advance hierarchical clustering to directed networks. Other recent examples of motif-based community detection in network include Arenas, Fernández, Fortunato, and Gomez (2008), Tsourakakis et al. (2017), Li, Huang, et al. (2019), and Huang et al. (2019).

## 6 | COMMUNITY DETECTION IN MULTILAYER, MULTISCALE AND HYPERGRAPH NETWORKS

Many man-made systems, for example, critical infrastructures integrating power grids, transportation, telecommunication, and other vital societal functions, as well as natural phenomena such as socio-environmental ecosystems and virus–host interactome exhibit a sophisticated, highly interdependent structure (Bachmann et al., 2020; De Domenico et al., 2016; Messina et al., 2020; Pilosof et al., 2017; Tang et al., 2015). Such interdependency can be, in turn, modeled as a multilayer graph, leading to an increasing spike of interest in complex multilayer networks.

A multilayer network, $\mathcal{G}$, consists of $m$ non-overlapping layers $\mathcal{G}=\{G_1, G_2, ..., G_m\}$, where each layer $G_i$ is modeled with a single-layer $G_k = (V_k, E_k)$, $k = 1, ..., m$. Here, $|V(G_k)|$ and $|E(G_k)|$ are the numbers of nodes and edges of layer $k$, respectively. The edges $E(G_k)$ refers to *within-layers* or *intralayer* links. The multilayer network $\mathcal{G}$ consists of a total of $\sum_{k=1}^{m}|V(G_k)|$ nodes. In addition, $\mathcal{G}$ has a set $E_I(\mathcal{G})$ of *interlayer* links which connect nodes across layers, that is, for each edge $(u,v)\in E_I(\mathcal{G})$ we have $u \in V(G_k)$ and $v \in V(G_l)$ for $k \neq l$ (Kivelä et al., 2014; Takes et al., 2018; Wider et al., 2016). The adjacency matrix of the multilayer network $\mathcal{G}$, which is referred to as *supra-adjacency matrix*, can be represented with a block-matrix structure, where diagonal elements represent within-layer links and off-diagonal elements represent cross-layer links. Figure 8 represents a simple illustration of a three-layer multilayer network with uniform interlayer weights.

Similar to the case of unilayer networks, the objective of community detection in a multilayer network is to unveil meaningful patterns of node groupings and to partition nodes into communities. However, several new challenges arise for community detection in multi-layer graphs. Every single layer represents distinct meaningful information from its own perspective. For example, in a case of critical infrastructures the first layer may corresponds to a power system, the second layer may represent a transportation system, while the third layer may be a telecommunication sector, and most importantly, all these layers are also tightly interconnected (Fügenschuh et al., 2021; Yang et al., 2021). Hence, we now deal with community detection not in a single network but in a *network of networks*, requiring to exploit and fuse multiple aspects of disparate, yet interdependent sources of information. While some community detection methods discussed in the previous sections can be extended to the analysis of multilayer networks, extracting multilayer communities poses many new challenges due to nontrivial heterogeneous interlayer and intralayer dependencies and yet remains a substantially less developed area in complex network analysis (Amelio et al., 2020; Contisciani et al., 2020; Yuvaraj et al., 2021). Figure 9 shows a schematic representation of communities in a multilayer network.

According to Tagarelli et al. (2017), we can classify the existing community detection methods for multilayer networks into three main categories: *flattening methods*, *aggregation methods*, and *direct methods*. Flattening methods convert the multilayer network to a single-layer network (flatten) and then apply any conventional community detection method (Berlingerio et al., 2011; Rocklin & Pinar, 2013). Aggregation methods, first, detect a community structure for each layer separately, after that an aggregation mechanism is used to combine each layer community structure (Berlingerio et al., 2013; Burgess et al., 2016). Direct methods work on the input multilayer network directly and evaluate a community structure by optimizing community-quality assessment criteria (De Domenico et al., 2015; Kuncheva & Montana, 2015; Mucha et al., 2010).

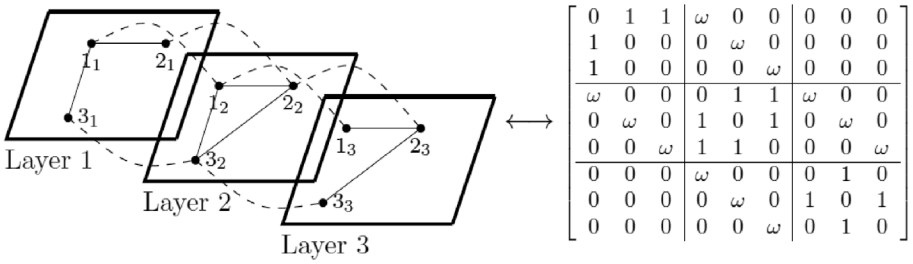
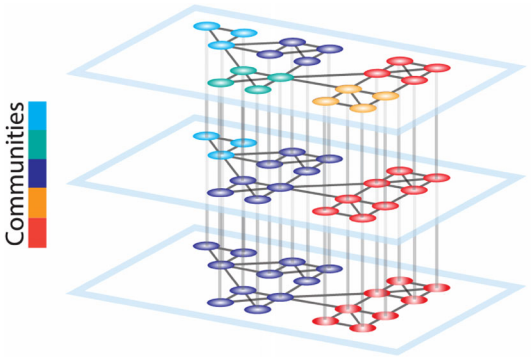



**FIGURE 8** A multilayer network (left) with (unweighted) intralayer connections (solid lines) and (weighted, $\omega$) interlayer connections (dashed curves) and its "supra-adjacency matrix" adjacency matrix (right). Reprinted figure with permission from Bazzi et al. (2016) (Copyright ©2016 Society for Industrial and Applied Mathematics. All rights reserved)

**FIGURE 9** Communities in a multilayer network. Reprinted figure with permission from Ashourvan et al. (2019)




To detect communities in multiplex networks, that is, a particular case of multilayer networks where nodes cannot be connected with other nodes in other layers, Arinik et al. (2020) consider a signed graph and employ the concepts of spectral clustering (see also Section 7.4). In turn, Hidalgo and Ma (2018) develop *multilayer NCut* clustering approach, with a particular focus on applications to multilayer omics data.

Vallès-Català et al. (2016), Paul and Chen (2016a), and Barbillon et al. (2017) extend the SBM framework to multilayer networks and refer the proposed extended model as *multilayer stochastic block model*. Similarly to the standard SBM model for unilayer networks (Equation (1)), for each $m$, conditioned on the community assignments of the nodes $z_i$ and $z_j$, the edges are formed independently following a Bernoulli distribution with probability $P(i, j)$, where, $P(i,j)\in G_m \left| \left(z_i = C_\ell, z_j = C_k\right) = M^m_{lk}, i, j \in \{1, 2, ..., |V(G_m)|\}, \ell, k \in \{1, 2, ..., r\}\right.$.

*Null models* based community detection in multilayer networks are described in Bazzi et al. (2016), Paul and Chen (2016b), and Pamfil et al. (2019). Kumar et al. (2010), DeFord and Pauls (2019), Znidi et al. (2019), Mercado et al. (2019), and Paul and Chen (2020) employ spectral methods (see Section 7.4) for clustering multilayer networks. Extracting overlapping communities in multiplex networks is considered by Afsarmanesh and Magnani (2016). Finally, recent results of Huang et al. (2019) extend the motif based approaches described in Section 5 to *M-motif* methodology for community detection in multilayer networks.

*Multiscale network* analysis is an emerging field in network sciences which investigates functionality and interdependency among different network components at various scales, for example, topological, temporal, and spatial scales. Such multiscale extensions of complex network appear in many domains of knowledge, including brain science (Ashourvan et al., 2019; Betzel et al., 2016; Betzel & Bassett, 2017) (e.g., Figure 10 shows brain network across three different scales), omics studies (Li et al., 2008), engineering (Gao et al., 2016), and environmental sciences (Agarwal et al., 2017). For instance, Mucha et al. (2010), Betzel et al. (2013), Garcia et al. (2018), and Peel et al. (2018) discuss applications of multiscale complex networks and assess evolution of community structures over a range of different scales.

Besides the methods discussed above, some other most recent developments in the community detection methodology for multilayer and multiscale networks include, for instance, community extraction in critical infrastructure networks (Li, Dong, & Mostafavi, 2019; Yildirimoglu & Kim, 2017) and temporal multilayer networks (Al-sharoa et al., 2019; Bazzi et al., 2016; Vajdi et al., 2020).

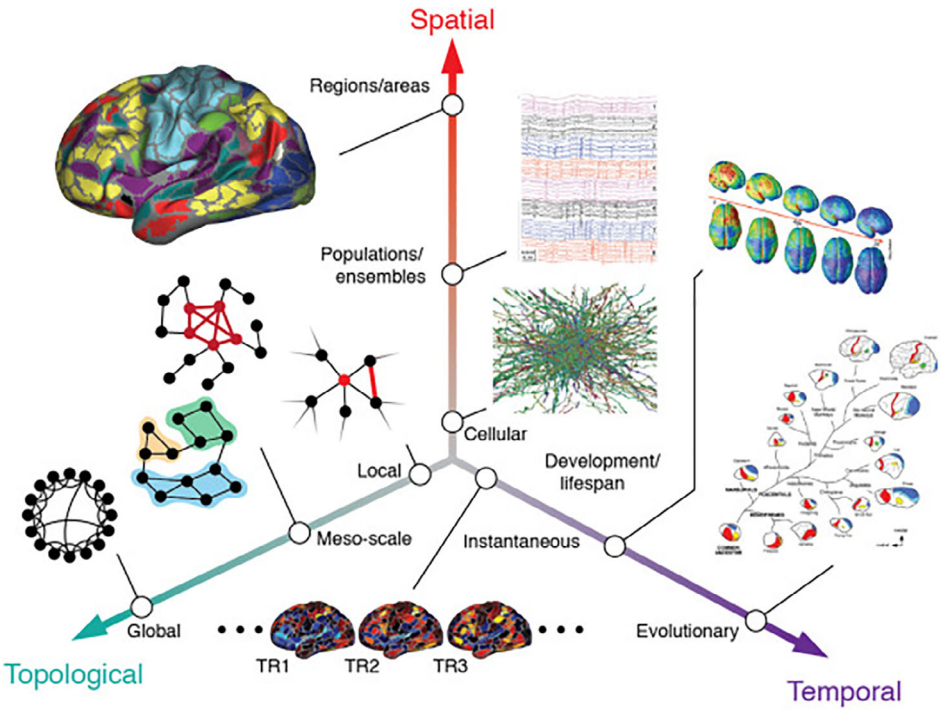



**FIGURE 10** Multiscale brain networks. The networks are organized across multiple spatio-temporal scales and also can be analyzed at topological (network) scales ranging from individual nodes to the network as a whole. Reprinted (adapted) figure with permission from Garcia et al. (2018)

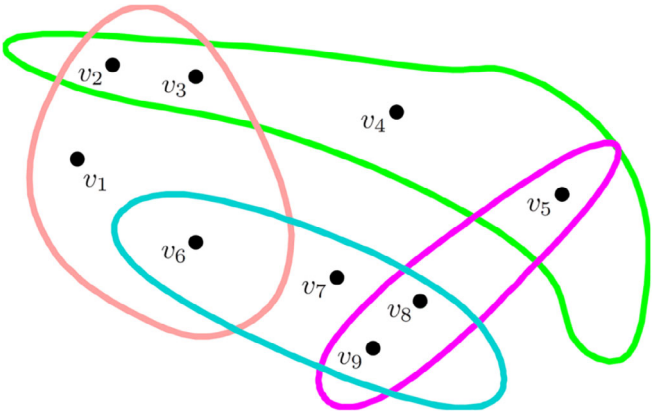



**FIGURE 11** A hypergraph with four hyperedges: $E = \{\{v_1, v_2, v_3, v_6\}, \{v_2, v_3, v_4, v_5\}, \{v_6, v_7, v_8, v_9\}, \{v_5, v_8, v_9\}\}$. Reprinted figure with permission from Gao et al. (2015)

Another actively developing research direction is the clustering of *hypergraph* data. A hypergraph is a generalization of a graph in which an edge can connect any number of nodes (called hyperedges), whereas, in a standard graph, an edge connects exactly two nodes (Berge, 1985; Bretto, 2013; Estrada & Rodríguez-Velázquez, 2006; Ouvrard, 2020; Voloshin, 2009). Hypergraphs are used in many areas, for example, in collaboration networks (Grossman & Ion, 1995), in chemical reactions (Temkin et al., 1996). Hypergraphs are also used in 3D-object retrieval and in image analysis (Gao et al., 2012; Wang et al., 2018; Zhu et al., 2015). Figure 11 shows an undirected hypergraph with nine nodes and four hyperedges.

Figure 12 represents two hypergraph examples: the left panel represents a music style hypergraph and the right panel represents a tag hypergraph (Wang et al., 2009). There are five common nodes in these two hypergraphs, which are "Angola Bond," "Who is he," "Dangerous," "Pleasure," and "Strip." The regions of different colors correspond to different hyperedges. Notice that the number of hyperedges correspond to music styles is four and the number of hyperedges correspond to music tags is three.

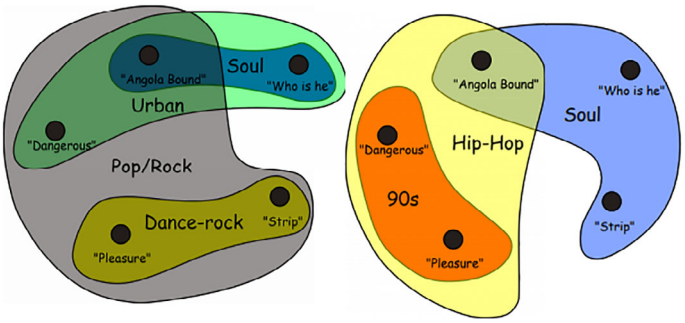



**FIGURE 12** Examples of the music style (left) and tag (right) hypergraphs. Reprinted figure with permission from Wang et al. (2009)

Clustering of hypergraphs can be approached using spectral methods (see, e.g., Zhou et al. (2006)). Many hypergraph clustering methods are based on the smallest possible hyperedge size. However, Purkait et al. (2017) introduce a hypergraph clustering algorithm based on large hyperedges. Most recently, Kumar et al. (2020) propose an iteratively reweighted modularity maximization algorithm for hypergraph clustering, while Takai et al. (2020) develop a clustering algorithm based on personalized PageRank on hypergraphs.

## 7 | OTHER POPULAR METHODS

### 7.1 | Modularity-based clustering

Modularity is one of the fundamental concepts in network community detection which can be viewed as a score that computes the difference between the observed structure of the network and the expected structure under some random (null) network model (Girvan & Newman, 2002; Hartwell et al., 1999; Ravasz et al., 2002). That is, in maximizing modularity, we compare the actual connections in a static or time-dependent network to the connections obtained from a random-graph ensemble that acts as a null model.

The most popular modularity optimization algorithm is proposed by Girvan and Newman (N–G model) (Newman, 2004, 2006). Let $c_i$ and $k_i$ denote the membership and degree of node $v_i$, respectively, $m$ is the number of edges in the network, then modularity $Q$ is defined as

$$Q = \frac{1}{2m}\sum_{ij}\left[A_{ij} - \frac{k_i k_j}{2m}\right]\delta(c_i, c_j),$$

where $\delta(c_i, c_j) = 1$ if $c_i = c_j$ and 0 otherwise, $A_{ij}$ is the number of edges between nodes $v_i$ and $v_j$, $k_i k_j/2m$ is the expected number of edges between nodes $v_i$ and $v_j$.

In the N–G null algorithm, the idea is to repeatedly merge two communities, leading to the largest increase in $Q$, This iterative procedure produces a dendrogram which shows the hierarchical decomposition of the network, and the stopping point is identified by modularity peak. The N–G null algorithm is computationally expensive and the computational complexity of the algorithm is $\mathcal{O}(m^2 n)$. Clauset et al. (2004) propose a greedy optimization of modularity to detect communities in large networks. Blondel et al. (2008) introduce a simple and efficient method for identifying communities in large networks. This method is commonly known as *Louvain method* for community detection, and the computational complexity of the method is $\mathcal{O}(n.\log^2 n)$. However, modularity optimization methods suffer a resolution limit and fail to detect small communities in larger networks (Fortunato & Barthélemy, 2007; Kumpula et al., 2007). To address the issue of the resolution limit, Arenas, Fernández, and Gomez (2008) propose a resistance parameter (self-loop) to every node. In turn, Reichardt and Bornholdt (2006) and Lancichinetti and Fortunato (2011) describe a parameter that can control the relative importance between within links of the communities and the null model. Recently, Hollocou et al. (2019) propose a relaxation of the modularity maximization problem which assigns each element of the dataset a probability to belong to a given cluster, whereas a solution of the standard modularity problem is a partition. Furthermore, Combe et al. (2020) develop modularity based clustering for complex networks with node attributes.

Finally, there are multiple results on modularity optimization using SDP and other relaxation methods in conjunction with SBM (see Section 3).

## 7.2 | Hierarchical clustering

Hierarchical clustering methods are based on building a hierarchy of clusters, which is often represented by a tree structure called a *dendrogram*. Hierarchical clustering can be divided into two main types: *agglomerative* and *divisive*. In agglomerative algorithms, clusters are iteratively merged if their similarity is sufficiently high. In divisive algorithms, the clusters are iteratively split by removing edges connecting nodes with low similarity.

Let $X_i$ and $X_j$ be any two subsets of $X$. The distance between $X_i$ and $X_j$, $\Delta(X_i, X_j)$, is called linkage distance (Nielsen, 2016). There are many linkage function, for example, *single linkage*, where $\Delta(X_i,X_j) = \min_{x_i \in X_i, x_j \in X_j} d(x_i,x_j)$, *complete linkage* (*or diameter*), where $\Delta(X_i,X_j) = \max_{x_i \in X_i, x_j \in X_j} d(x_i,x_j)$, and *average linkage*, where

$$\Delta(X_i,X_j) = \frac{1}{|X_i||X_j|} \sum_{x_i \in X_i} \sum_{x_j \in X_j} d(x_i,x_j),$$

where $d(x_i, x_j)$ is basic distance (e.g., Euclidean distance) between any two elements of $X$.

Due to its flexible data-driven and model-free framework allowing to easily visualize the obtained node partitions, hierarchical clustering has found its place in many applications of network sciences. For example, Liu, Yao, and Zhao (2020) use hierarchical clustering in fMRI data. Jasiński et al. (2020) apply hierarchical clustering for power quality measurements in a power grid network. Bandyopadhyay and Coyle (2003) design an energy efficient hierarchical clustering algorithm especially for wireless networks, while Brzoska et al. (2020) investigate partitions of the German cattle trade network using hierarchical SBM.

Furthermore, hierarchical clustering is often combined with other methods for complex network analysis. For instance, most recently Mémoli and Pinto (2020) describe hierarchical clustering method for directed network based on the concept of network motifs (for more details about motif based community detection see Section 6). Côme et al. (2020) introduce a hierarchical clustering which combine discrete latent variable models and classification likelihood. In turn, Lyzinski et al. (2017) propose a community detection method based on hierarchical SBMs, while Paez et al. (2019) introduce a hierarchical SBM for community detection in multiplex complex networks (for more details about community detection using stochastic block models see Section 3).

## 7.3 | Walk-based algorithms

The idea in a walk-based clustering algorithm is that nodes within the same community should be reached through relatively shorter path than the average path in the network. Rosvall and Bergstrom (2008) employ maps to describe the dynamics across the edges and nodes in networks that represent the local interactions among the subunits of a system (this method is commonly known as *Infomap*). *Infomap* uses a random walk as a proxy for the information flow and Huffman code (Huffman, 1952) to assign short codewords to common objects and long codewords to rare objects. Such an approach is applicable to weighted directed and undirected networks.

Pons and Latapy (2005) propose a clustering algorithm, called *Walktrap*, where the similarity between nodes is computed based on random walks. At each step, a walker is on a node and moves to a node with a random walk process (Markov chain), chosen randomly and uniformly among its neighbors. At each step, the transition probability from node $i$ to node $j$ is $P_{ij} = A_{ij}/d_i$. The probability of going from $i$ to $j$ through a random walk of length $t$ is defined as $P^t_{ij}$. That is, $P^t_{ij}$ is the probability to go from $i$ to $j$ in $t$ steps. The distance metric $r$ to measure node similarities is defined as

$$r_{ij} = \sqrt{\sum_{k=1}^{n} \frac{\left(P^t_{ik} - P^t_{jk}\right)^2}{d(k)}}.$$

$$r_{C_1C_1} = \sqrt{\sum_{k=1}^{n} \frac{\left(P^t_{C_1k} - P^t_{C_2k}\right)^2}{d(k)}}.$$

The *Walktrap* algorithm uses hierarchical clustering algorithm to find community structures based on distance metrics $r_{ij}$ and $r_{C_1C_1}$. Wei et al. (2009) propose a random walk based clustering algorithm (*DOCS*) to detect, especially, overlapping network communities.

Recently, Okuda et al. (2019) consider a restrained random-walk similarity measure to detect communities in a graph, where the starting nodes of random walks are grouped in the same community if the walkers pass similar sets of nodes. In turn, Hua et al. (2020) introduce a random walk gap mechanism for a signed graph. This method constructs two random walks, a random walk graph on positive edges, and a random walk graph on negative edge in the signed networks. Clustering is then conducted using the gap between these two random walk graphs. More generally, clustering based on similarity of various random walks, including Levy Flights (Michelitsch et al., 2019; Riascos & Mateos, 2012) constitutes one of the most promising and actively developing research directions.

## 7.4 | Spectral clustering

Due to its computational performance and tractability, spectral clustering is one of the most popular methods for network community detection. In spectral clustering the idea is to embed a graph $G$ into some Euclidean space of lower dimension, for example, $\mathbb{R}^k$, $k$ with $k \ll n$, where $n$ is the number of nodes (Chung & Lu, 1997). That is, as a result of such embedding we no longer deal with a graph structure but with its representation in a linear space. In particular, let $k$ be a given number of clusters, and let $\mathbf{x}_{\cdot j}$, $j = 1, ..., k$, be orthogonal eigenvectors of the *Laplacian* $L$, corresponding to the $k$ largest eigenvalues. Form an $n \times k$-matrix $X = [\mathbf{x}_{\cdot 1}, ..., \mathbf{x}_{\cdot K}]$, where each row of $X$, $\mathbf{x}_i \equiv \mathbf{x}_{i\cdot}$, provides a representation of a node in $\mathcal{V}$ in $\mathbb{R}^k$. We now cluster the resulting $n$ sample points in $\mathbb{R}^k$ using any appropriate classifier, such as, for instance, $k$-means, $k$-medians, or $k$-medoids (von Luxburg, 2007). If the number of clusters $k$ is unknown, which is typically the case in most applications, an optimal $k$ can be identified from an eigengap analysis, for example, a scree plot of leading eigenvalues. Alternatively, crossvalidation arguments as well as hierarchical divisive or agglomerative clustering can be employed to identify optimal $k$ (Chen & Lei, 2018; Garcia, 2016; Li, Levina, & Zhu, 2020).

A $k$-means algorithm is one of the primary tools within the spectral clustering framework. Given data points $\{\mathbf{x}_i\}_1^n, \mathbf{x}_i \in \mathbb{R}^p$, group the data $\{\mathbf{x}_i\}_1^n$ into $k$ clusters $\mathbf{C} = \{C_1, ..., C_k\}$ in such a way that the within-cluster sum of squares is minimized, that is

$$\underset{\mathbf{C}}{\operatorname{argmin}} \sum_{k=1}^{k} \sum_{\mathbf{x} \in C_k} \|\mathbf{x} - \boldsymbol{\mu}_k\|^2,$$

where $\boldsymbol{\mu}_k$ is the $k$th group mean and $||\mathbf{x} - \boldsymbol{\mu}_k||^2$ is the squared Euclidean distance between $\mathbf{x}$ and $\boldsymbol{\mu}_k$. Consistency and other theoretical properties of spectral clustering under SBM have been studied, for example, by Rohe et al. (2011), Lei and Rinaldo (2015), Zhang and Zhou (2016), and Bhattacharyya and Bickel (2016), while theoretical analysis under degree-corrected SBM includes, for instance, Lyzinski et al. (2014), Gulikers et al. (2017), and Gao et al. (2018) (see Zhao, 2017; Zhou & Amini, 2019 and references therein for a more detailed discussion of theoretical results of network spectral clustering). In turn, Leskovec et al. (2008) study the empirical performance of spectral clustering for community detection as a function of network order, that is, number of nodes.

However, many real world networks exhibit a phenomenon of an unbalanced structure with sparse connectivity but low diameters. That is, the majority of nodes in such sparse networks have only a few edges while some small number of nodes are highly linked. Such phenomena raise an issue of anomalies or outliers in network community detection, resulting in decreased network detection performance.

The idea of regularization in this context is to diminish the impact of low degree nodes, by viewing them as outliers and shrinking them toward the center of the spectrum. As a result, regularization leads to a higher concentration of the Laplacian. There are a number of regularization procedures ranging from brute-force trimming of outliers to sophisticated methods that are closely connected to regularization of covariance matrices (Le et al., 2017). One of the most

popular approaches, by analogy with a ridge regularization of covariance matrices, is to select some positive parameter $\tau$ and add $\tau/n$ to all entries of the adjacency matrix $A$ (Amini et al., 2013), that is

$$A_\tau = A + \frac{\tau}{n} J,$$

where $J = \mathbb{1}\mathbb{1}^T$, $\mathbb{1}$ is $n \times n$-matrix with all elements 1. The resulting regularized Laplacian then takes a form

$$L_\tau = D_\tau^{-1/2} A_\tau D_\tau^{-1/2},$$

where $D_{ii,\tau} = \sum_{j=1}^{n} A_{ij} + \tau$. The optimal regularizer $\tau$ can be then selected by minimizing the Davis–Kahan bound, that is the bound on the distance between the sample and population Laplacians (Dall'Amico et al., 2020; Joseph & Yu, 2016; Le et al., 2017; Nadler et al., 2009). Cao and Chen (2011) study the consistency of the regularized spectral clustering algorithm, while Cao and Chen (2012) and Yang et al. (2019) evaluate error analysis on graph Laplacian regularized estimator.

However, there are a number of challenges associated with such an approach. First, selecting an optimal regularizer $\tau$ tends to be computationally expensive. Second, the impact of small and weak network communities on regularized spectral graph methods for sparse networks may be nontrivial (Cai & Li, 2015) and this regularizer makes this approach much more difficult to apply sparse eigen-methods. Finally, the conventional $k$-means method is known to be sensitive to various types of outliers and anomalies, both in conjunction with network studies and multivariate analysis (Baruri et al., 2019; García-Escudero & Gordaliza, 1999; von Luxburg, 2007).

To address the sensitivity of spectral clustering to outliers in complex networks, there has been developed and analyzed multiple approaches, including, for example, $k$-medians (Lei & Rinaldo, 2015), various local algorithms Chen et al. (2016), and penalized $k$-means (Banerjee et al., 2019). Although these algorithms are more robust compared to $k$-means, similarly to $k$-means most of these methods do not account for the underlying geometry of the network data.

To address these challenges, Tian and Gel (2017) introduce the concept of nonparametric data depth (Mosler, 2013; Vardi & Zhang, 2000; Zuo & Serfling, 2000) into network analysis, particularly, within the spectral clustering framework and propose a $k$-depths algorithm for identifying network community structures. The key idea behind $k$-depths is to employ a data depth function, simultaneously accounting for the intrinsic network geometry, anomalies, and outliers, as the primary community cohesion metric rather than the Euclidean distance among node pairs as in $k$-means and $k$-medians. Figure 13 shows a schematic representation of the gains delivered by the $k$-depths approach in the presence of outliers. Dey et al. (2017) discuss the utility of $k$-depths based on various data depth functions, in application to optimal islanding (clustering) of power grid networks and show that spectral clustering with data depths tends to outperform spectral clustering with $k$-means in terms of $k$-way expansion. Figure 14 shows the partitioning of the Italian grid into five zones, based on $k$-means and spectral clustering with $k$-depths (with $L_1$ and Mahalanobis depth functions). In turn, Tian and Gel (2019) propose spectral clustering of complex networks based on the notion of data depth in

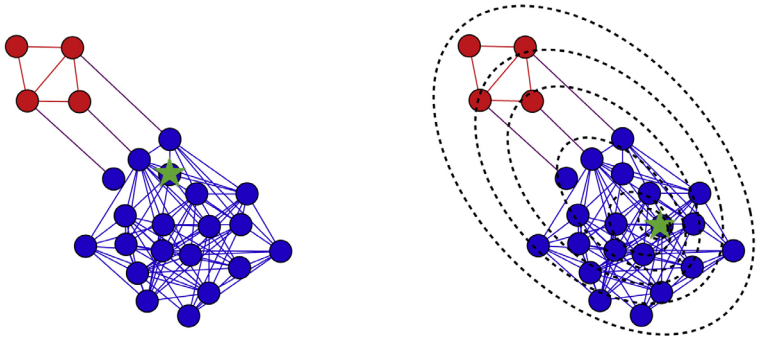


**FIGURE 13** Clustering with the presence of outliers (in red). Green stars represent community centers yielded by $k$-means (left panel) and depth-based clustering (right panel). Reprinted figure with permission from (Tian & Gel, 2019)

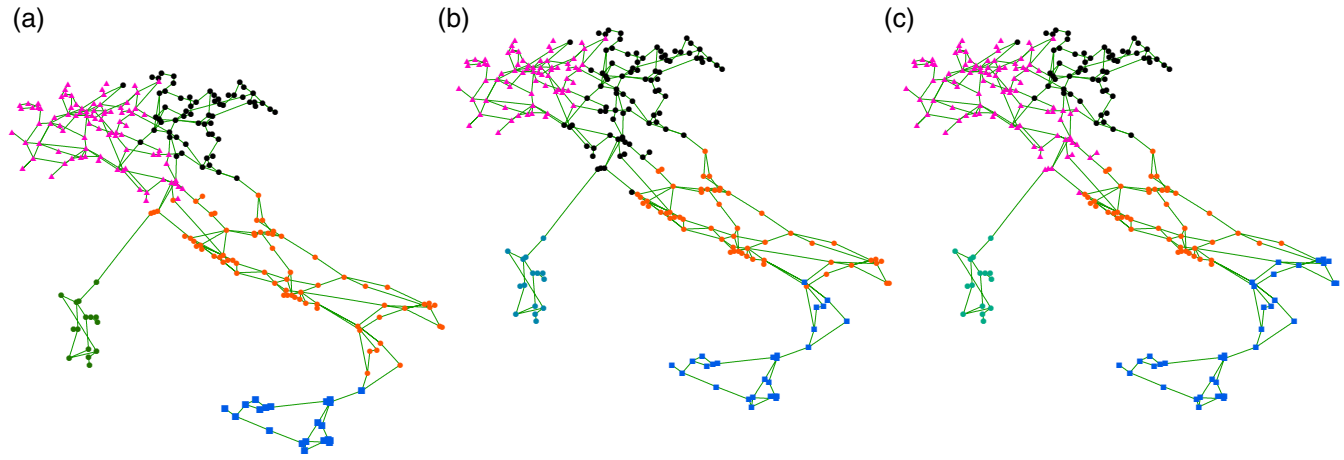



**FIGURE 14** The five zones of the Italian power grid network. (a) Spectral clustering with $k$-means, (b) spectral clustering with $k$-depth with $L_1$, (c) spectral clustering with $k$-depth with MhD. The *k-way expansion*, a performance metric in spectral clustering (a lower $k$-way expansion indicates a better clustering algorithm), of these three algorithms are 1.95%, 1.65%, and 1.75%, respectively. Reprinted figure with permission from Dey et al. (2017)

supervised settings, using tools of the Depth versus Depth (DD(G)) classifiers (Cuesta-Albertos et al., 2017; Li et al., 2012). Finally, Zhang et al. (2021) integrate the depth concept into the analysis of communities of complex networks with multiple attributes. The results of Tian and Gel (2019) and Zhang et al. (2021) indicate that data depth-based approaches for network community detection may yield noticeable gains in performance accuracy comparing to regularized spectral clustering, especially for sparse networks and in the presence of outliers.

Some most recent research directions on spectral clustering for complex networks include, *multiway spectral clustering* (Damle et al., 2018; Lee et al., 2014; Riolo & Newman, 2014), approximation algorithms for Cheeger ratio (Louis et al., 2011; Trevisan, 2013), recovery of overlapping communities in large-scale networks (Van Lierde et al., 2020; Zhang et al., 2020), as well as in multilayer and dynamic networks (see Sections 6 and 8, respectively).

## 8 | COMMUNITY DETECTION IN DYNAMIC NETWORKS

Most network community detection approaches focus on static networks, that is, a network representation at a certain fixed time point. However, such a scenario is often too simplistic to model the ever changing nature in the real world. In fact, temporal factors play an important role in characterizing network structure, that is, nodes/entities and edges/interactions may appear, disappear, merge, or continue as time progresses. The dynamic network with temporal dimension aggregates features from both evolving nature and features from entities and interactions and hence conveys critical information about the underlying community structure of the network. The goal of this section is to introduce a list of methods which are dedicated to tracking network topology and uncover community structures.

In particular, we introduce two different definitions of dynamic networks, depending on temporal granularity.

- *Temporal network*: A temporal network is defined as $G = (V, E, T)$, where $V$ is a set of triplets with the form $(v, t_s, t_e)$, $v$ is a node in the network, $t_s, t_e \in T$ are birth and death timestamps of node $v$, and $t_s < t_e$. $E$ is a set of quadruplets $(u, v, t_s, t_e)$, $u, v$ are nodes in the network, $t_s, t_e \in T$ are birth and death timestamps of the corresponding edge, and $t_s < t_e$. If the original data include accurate temporal information, like friend network, e-mail network or phone call network with exact timestamps, the temporal network will provide a fine representation of the data and a precise mathematical object for the following analysis.
- *Snapshot network*: A snapshot network is defined as an ordered set $< G_1, G_2, ..., G_t >$, where $G_i = (V_i, E_i)$ is the network at time $i$ or the aggregation of observed interactions during the certain time period.
  Snapshot network contains aggregated data, the analysis complexity depends on the level of aggregation and network size/order (number of nodes/edges), since dynamic community detection needs to run a clustering algorithm on each snapshot. The computational cost at one snapshot does not depend on the change between the current and past snapshots.

More definitions of the dynamic network could be found in Holme and Saramäki (2012), Viard et al. (2016), Latapy et al. (2018), and Rossetti and Cazabet (2018).

## 8.1 | Dynamic community and community operations

The aim of community detection in dynamic networks is to identify and track the evolution of community structures. The first step is to define what a dynamic community is. Given a time-evolving network, a dynamic community can be defined as a collection of nodes and corresponding time durations, that is, $(v_1, P_1)$, $(v_2, P_2)$, ..., $(v_n, P_n)$, where $P_n = ((t_{s0}, t_{e0}), (t_{s1}, t_{e1}), ..., (t_{sN}, t_{eN}))$, $t_{si} < t_{ei}$, $i = 1, 2, ..., N$.

Community tracking then focuses on the evolution of entities and their interaction, that is, appearance and vanishing of nodes and edges. Palla et al. (2007) and Cazabet and Amblard (2014) outline the following eight operations:

- *Birth*: the first appearance of a new community;
- *Death*: the vanishing of an existing community;
- *Growth*: the size of a community increases;
- *Contraction*: the size of a community decreases;
- *Merge*: two or more communities merge into a single one;
- *Split*: a community splits into two or more communities;
- *Continue*: the size and order of a community remain unchanged;
- *Resurgence*: a community appears again after vanishing for a period.

Following Rossetti and Cazabet (2018), current algorithms for dynamic network community detection can be largely classified into the following five categories:

*Instant-optimal community detection*: This class of approaches focuses on the current state of the network at time $t$ and identifies communities existing at $t$. Typically, two steps are included in this process. First, static communities are detected at each snapshot of the time evolving graph. Second, communities detected at the current snapshot are aligned with communities found in snapshots at earlier time points. These approaches allow for the direct incorporation of methods for static community detection into the analysis of time-dependent networks. Furthermore, parallelization of community detection at each timestamp can lead to substantial gains in computational efficiency. However, such approaches suffer from instability since the same algorithm may identify very different communities in the adjacent time snapshots of a time-dependent network, as a result, leading to challenges in distinguishing between the changes due to community evolvement, anomalies, and intrinsic algorithm uncertainties.

With the idea of instant-optimal community detection, Bourqui et al. (2009) focus on dynamic connections and hidden hierarchical structure, while Bóta et al. (2011) propose an extended approach that can efficiently handle community detection and community events in large networks. In turn, Bródka et al. (2013) develop group evolution discovery with the change indicator called *inclusion measure*. Martinet et al. (2020) propose a robust dynamic community detection algorithm for human brain functional networks based on the percolation method.

*Temporal trade-off community detection*: This class of methods focuses not only on the current network structure but also on the topology of the past network evolution. That is, these approaches utilize communities found in previous steps to identify communities in the current step. The following iterative steps are used in the category. First, initial communities are found based on the initial state of the network. Second, as time goes by, extract communities at the current state using both current network data and communities found in previous steps. This class of methods embeds previous information into the community detection, as a result, they are able to mitigate the instability concerns in instant-optimal approaches. However, parallelization of community detection at each state is challenging since the outputs at each step are related. Moreover, the community may drift due to the involvement of other states. Among contributions that use temporal trade-off methods for community detection are, for instance, Görke et al. (2010) who integrate modularity and derive a tight bound on the number of operations and Bansal et al. (2011) who discuss a fast community detection algorithm based on historical community information. In turn, a game-theoretic approach where community structure evolves until the game reaches the Nash equilibrium is proposed by Alvari et al. (2014). Zakrzewska and Bader (2015) consider a dynamic seed set expansion, with updating communities as the graph evolves. Finally, Messaoudi and Kamel (2019) propose a multi-objective algorithm for community detection on dynamic

networks, which simultaneously optimizes the modularity density and the normalized mutual information of the solutions as objective functions.

*Cross-time community detection*: This class of approaches focuses on identifying communities with the consideration of the whole network evolution, that is, a single classification/clustering for all time steps. First, a single network is created by aggregating all evolving information. Second, a static community detection method is applied to the created network. This process will filter out local outliers, anomalies, and slow evolvement through multiple aggregations, as a result diminishing the effect of algorithm instability and community drifts. However, this category of methods is inapplicable to address community detection in real time since the community identification is based on a single aggregated network. Examples of cross-time community detection include Duan et al. (2009) who study weighted and directed networks, Aynaud and Guillaume (2011) who focus on a unique decomposition that is connected with every time step, Ghasemian et al. (2016) who explore theoretical properties such as accuracy and limits of detectability, and Liu et al. (2018) who propose a global spectral clustering by eigenvector smoothing.

*Stochastic block model based community detection*: This method is proposed as a temporal SBM in which interactions between nodes evolve and have certain time periods (Corneli et al., 2016; Karaaslanli & Aviyente, 2020; Pensky & Zhang, 2019). Time intervals are divided into subintervals of fixed identical duration. The idea is to classify subintervals into homogeneous interaction patterns. Integrated complete-data likelihood is used for optimization, and the greedy search approach is employed to identify the number of nodes in a cluster, time intervals, and node memberships. The model proposed by Matias and Miele (2017) combines static parts in SBM and Markov chains to track the evolution of communities, allowing group membership and connectivity parameters to vary over time. Here model parameters and the resulting partitioning of nodes are performed using the variational expectation maximization procedure, and the numerical results indicate performance gains both in terms of accuracy and interpretability.

While originally methods for community detection in dynamic networks (arguably) have largely focused on social network analysis (Dakiche et al., 2019; Leitch et al., 2019; Nguyen et al., 2014; Oliveira & Gama, 2012; Palla et al., 2009; Rossetti & Cazabet, 2018), the discovery of community structure in time evolving graphs continues to gain popularity in many other disciplines, from finance to transportation systems to such less explored areas for network analysis as medicine and epidemiology. For example, Wong et al. (2018) investigate underlying mechanisms in disease progression by studying a dynamic graph which consists of genes and their interactions. In particular, Wong et al. (2018) evaluate how density and pattern of gene interactions evolve as disease progress with a goal to enhance our understanding of which subgroups of genes tend to be associated with early/late development of the disease. If these genes are identified, they may be considered as potential targets of treatment at different disease stages. To glean deeper insights into these hidden interactions, Wong et al. (2018) propose a novel algorithm called SDREGION which identifies subgroups whose density increase/decrease continuously over time. The objective function is then used to detect specific regions which represent medical mechanisms of targeted pathways, integrins, and focal adhesions. Wong et al. (2018) also provide an example of the application in the tumor progression in non-small cell lung cancer, and the results indicate that SDREGION may assist in discovering intrinsic mechanisms behind disease dynamics. These findings open a door to new research directions at the interface of community detection, epidemiology, and data science for network-based omics applications (Holzinger et al., 2019; Leitch et al., 2019).

## 9 | FUTURE DIRECTIONS

Network clustering is one of the most actively developing research areas spanning a broad range of disciplines and applications, from statistics and machine learning to social sciences, neuroscience, and finance. While there exists a multitude of emerging research directions in network clustering, here we highlight two of them, namely, network clustering with parallel computing and clustering on blockchain networks.

In particular, given vast amounts of graph structured data especially in data streaming scenarios, the more conventional clustering algorithms described in this paper often become computationally expensive if not prohibitive for large real-world networks. To address this limitation, nowadays, there has been a surge of interest in *parallel graph clustering algorithms* which are based on the concept of parallel computing. For example, Pan et al. (2015) propose a parallel correlation clustering for large complex networks. Meyerhenke et al. (2017) introduce a scalable parallelization of the size-constrained label propagation algorithm for large complex networks clustering. Shun et al. (2016) introduce a parallel *local* graph clustering, which can find good clusters in a network with work proportional to the size of the cluster rather than that of the entire graph. A number of parallel clustering methods

have also been developed to reduce the communication bottlenecks in supercomputing workloads. For instance, Huo et al. (2019) propose a novel parallel clustering method for logs data to better serve network system and network security. Dafir et al. (2020) summarize recent developments in parallel clustering methods in both the horizontal and vertical scaling platforms.

In turn, the rapidly emerging blockchain technology poses a set of new challenging problems in network sciences and clustering, in particular. First, in addition to sheer amounts of information to be analyzed, blockchain transaction graphs are highly sparse and time-evolving, making application of many conventional clustering approaches infeasible (Akcora et al., 2019; Akcora, Gel, & Kantarcioglu, 2020; Liu, Jiang, et al., 2020). Second, one of the primary goals of cluster analysis on blockchain networks is to identify addresses which are managed by the same user, rather than to group multiple users based on their similarity (Victor, 2020). The results of such cluster analysis, for instance, are of high importance for law enforcement agencies in record linkage associated with money laundering detection on the blockchain. This direction of network clustering yet remains in its infancy and requires developing of a new set of methods and algorithms (Akcora, Purusotham, et al., 2020).

## 10 | CONCLUSION

Analysis of community structures in complex networks is one of the cornerstones of network sciences, spanning such disciplines as statistics, machine learning, social studies, computer, and information sciences. Applications of community detection are even more diverse, from biology and epidemiology to finance and marketing to cyber-security, power grids, and critical infrastructures.

Given such a diverse set of methods and their application areas, no survey paper on community detection can be exhaustive. The goal of our paper is to contribute to this very rich and constantly evolving body of knowledge by providing an overview of some recent emerging directions in community extraction approaches and by offering a broader picture of how these interdisciplinary tools and algorithms are interlinked methodologically.

### CONFLICT OF INTEREST

The authors have declared no conflicts of interest for this article.

### AUTHOR CONTRIBUTIONS

**Asim K. Dey:** Conceptualization; data curation; formal analysis; investigation; methodology; project administration; resources; validation; writing-original draft; writing-review & editing. **Yahui Tian:** Conceptualization; data curation; formal analysis; investigation; methodology; resources; validation; visualization; writing-original draft; writing-review & editing. **Yulia Gel:** Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; supervision; validation; visualization; writing-original draft; writing-review & editing.

### DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

### ORCID

*Yulia R. Gel* https://orcid.org/0000-0002-4500-6495

### ENDNOTE

[1] Throughout the paper we use the notions of graphs and networks interchangeably.

### RELATED WIREs ARTICLES

A survey on theoretical advances of community detection in networks
An overview of social network analysis
Community detection in social networks
Leader-based community detection algorithm for social networks

## REFERENCES

Abay, N. C., Akcora, C. G., Gel, Y. R., Kantarcioglu, M., Islambekov, U. D., Tian, Y., & Thuraisingham, B. (2019). Chainnet: Learning on blockchain graphs with topological features. In *2019 IEEE international conference on data mining (ICDM)* (pp. 946–951).

Abbe, E. (2018). Community detection and stochastic block models: Recent developments. *Journal of Machine Learning Research*, *18*, 1–86.

Afsarmanesh, N., & Magnani, M. (2016). Finding overlapping communities in multiplex networks. *CoRR*, abs/1602.03746.

Agarwal, A., Marwan, N., Rathinasamy, M., Merz, B., & Kurths, J. (2017). Multiscale complex network analysis: An approach to study spatio-temporal rainfall pattern in Germany. In *European Geophysical Union*.

Agasse-Duval, M., & Lawford, S. (2018). *Subgraphs and motifs in a dynamic airline network*. arXiv preprint arXiv:1807.02585.

Ahmed, N. K., Neville, J., Rossi, R. A., Duffield, N., & Willke, T. L. (2016). Graphlet decomposition: Framework, algorithms, and applications. *Knowledge and Information Systems (KAIS)*, *50*, 1–32.

Aicher, C., Jacobs, A. Z., & Clauset, A. (2014). Learning latent block structure in weighted networks. *Journal of Complex Networks*, *3*, 221–248.

Airoldi, E. M., Blei, D. M., Fienberg, S. E., & Xing, E. P. (2008). Mixed membership stochastic blockmodels. *Journal of Machine Learning Research*, *9*, 1981–2014.

Akcora, C. G., Dixon, M. F., Gel, Y. R., & Kantarcioglu, M. (2019). Blockchain data analytics. *Intelligent Informatics*, *20*(1), 1–7.

Akcora, C. G., Gel, Y. R., & Kantarcioglu, M. (2020). *Blockchain: A graph primer*. arXiv:1708.08749.

Akcora, C. G., Li, Y., Gel, Y. R., & Kantarcioglu, M. (2020). BitcoinHeist: Topological data analysis for ransomware detection on the bitcoin blockchain. In *International joint conference on artificial intelligence (IJCAI)*.

Akcora, C. G., Purusotham, S., Gel, Y. R., Krawiec-Thayer, M., & Kantarcioglu, M. (2020). *How to not get caught when you launder money on blockchain*? arXiv:2010.15082.

Al-sharoa, E., Al-khassaweneh, M., & Aviyente, S. (2019). Tensor based temporal and multilayer community detection for studying brain dynamics during resting state fmri. *IEEE Transactions on Biomedical Engineering*, *66*, 695–709.

Alvari, H., Hajibagheri, A., & Sukthankar, G. (2014). Community detection in dynamic social networks: A game-theoretic approach. In *Proceedings of the 2014 IEEE/ACM international conference on advances in social networks analysis and mining* (pp. 101–107). IEEE Press.

Amelio, A., Mangioni, G., & Tagarelli, A. (2020). Modularity in multilayer networks using redundancy-based resolution and projection-based inter-layer coupling. *IEEE Transactions on Network Science and Engineering*, *7*, 1198–1214.

Amini, A. A., Chen, A., Bickel, P. J., & Levina, E. (2013). Pseudo-likelihood methods for community detection in large sparse networks. *The Annals of Statistics*, *41*, 2097–2122.

Amini, A. A., & Levina, E. (2018). On semidefinite relaxations for the block model. *The Annals of Statistics*, *46*, 149–179.

Anandkumar, A., Ge, R., Hsu, D., Kakade, S. M., & Telgarsky, M. (2014). Tensor decompositions for learning latent variable models. *Journal of Machine Learning Research*, *15*, 2773–2832.

Arenas, A., Fernández, A., Fortunato, S., & Gomez, S. (2008). Motif-based communities in complex networks. *Journal of Physics A: Mathematical and Theoretical*, *41*, 224001.

Arenas, A., Fernández, A., & Gomez, S. (2008). Analysis of the structure of complex networks at different resolution levels. *New Journal of Physics*, *10*, 053039.

Arinik, N., Figueiredo, R., & Labatut, V. (2020). Multiple partitioning of multiplex signed networks: Application to European parliament votes. *Social Networks*, *60*, 83–102.

Ashourvan, A., Telesford, Q. K., Verstynen, T., Vettel, J. M., & Bassett, D. S. (2019). Multi-scale detection of hierarchical community architecture in structural and functional brain networks. *PLoS One*, *14*, 1–36.

Aynaud, T., & Guillaume, J.-L. (2011). Multi-step community detection and hierarchical time segmentation in evolving networks. In *Proceedings of the 5th SNA-KDD workshop*.

Bachmann, I., Bustos, J., & Bustos, B. (2020). A survey on frameworks used for robustness analysis on interdependent networks. *Complexity*, *2020*, 1–17.

Bandyopadhyay, S., & Coyle, E. (2003). An energy efficient hierarchical clustering algorithm for wireless sensor networks. In *IEEE INFOCOM 2003. Twenty-second annual joint conference of the IEEE computer and communications societies (IEEE Cat. No. 03CH37428)* (Vol. 3, pp. 1713–1723).

Banerjee, S., Akbani, R., & Baladandayuthapani, V. (2019). Spectral clustering via sparse graph structure learning with application to proteomic signaling networks in cancer. *Computational Statistics & Data Analysis*, *132*, 46–69 (Special issue on Biostatistics).

Bansal, S., Bhowmick, S., & Paymal, P. (2011). Fast community detection for dynamic complex networks. In *Complex networks* (pp. 196–207). Springer.

Barbillon, P., Donnet, S., Lazega, E., & Bar-Hen, A. (2017). Stochastic block models for multiplex networks: An application to a multilevel network of researchers. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *180*, 295–314.

Baruri, R., Ghosh, A., Chanda, S., Banerjee, R., Das, A., Mandal, A., & Halder, T. (2019). A comparative study on *k*-means clustering method and analysis. In *International conference on emerging technologies in computer engineering* (pp. 113–127).

Battiston, F., Nicosia, V., Chavez, M., & Latora, V. (2017). Multilayer motif analysis of brain networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *27*, 047404.

Bazzi, M., Porter, M. A., Williams, S., McDonald, M., Fenn, D. J., & Howison, S. D. (2016). Community detection in temporal multilayer networks, with an application to correlation networks. *Multiscale Modeling & Simulation*, *14*, 1–41.

Bedi, P., & Sharma, C. (2016). Community detection in social networks. *WIREs Data Mining & Knowledge Discovery*, *6*, 115–135.

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

Berge, C. (1985). *Graphs and hypergraphs*. Elsevier Science Ltd.

Berlingerio, M., Coscia, M., & Giannotti, F. (2011). Finding and characterizing communities in multidimensional networks. In *2011 international conference on advances in social networks analysis and mining* (pp. 490–494).

Berlingerio, M., Pinelli, F., & Calabrese, F. (2013). ABACUS: Frequent pattern mining-based community discovery in multidimensional networks. *Data Mining and Knowledge Discovery*, *27*, 294–320.

Betzel, R. F., & Bassett, D. S. (2017). Multi-scale brain networks. *NeuroImage*, *160*, 73–83.

Betzel, R. F., Fukushima, M., He, Y., Zuo, X.-N., & Sporns, O. (2016). Dynamic fluctuations coincide with periods of high and low modularity in resting-state functional brain networks. *NeuroImage*, *127*, 287–297.

Betzel, R. F., Griffa, A., Avena-Koenigsberger, A., Goñi, J., Thiran, J.-P., Hagmann, P., & Sporns, O. (2013). Multi-scale community organization of the human structural connectome and its relationship with resting-state functional connectivity. *Network Science*, *1*, 353–373.

Bhattacharyya, S., & Bickel, P. J. (2016). Spectral clustering and block models: A review and a new algorithm. In A. Frigessi, P. Bühlmann, I. K. Glad, M. Langaas, S. Richardson, & M. Vannucci (Eds.), *Statistical analysis for high-dimensional data* (pp. 67–90). Springer International Publishing.

Bickel, P. J., & Sarkar, P. (2016). Hypothesis testing for automated community detection in networks. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *78*, 253–273.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*, P10008.

Bóta, A., Krész, M., & Pluhár, A. (2011). Dynamic communities and their detection. *Acta Cybernetica*, *20*, 35–52.

Bourqui, R., Gilbert, F., Simonetto, P., Zaidi, F., Sharan, U., & Jourdan, F. (2009). Detecting structural changes and command hierarchies in dynamic social networks. In *2009 International conference on advances in social network analysis and mining* (pp. 83–88). IEEE.

Bretto, A. (2013). *Hypergraph theory: An introduction*. Springer Publishing Company, Incorporated.

Bródka, P., Saganowski, S., & Kazienko, P. (2013). GED: The method for group evolution discovery in social networks. *Social Network Analysis and Mining*, *3*, 1–14.

Bron, C., & Kerbosch, J. (1973). Algorithm 457: Finding all cliques of an undirected graph. *Communications of the ACM*, *16*, 575–577.

Brumback, T., Castro, N., Jacobus, J., & Tapert, S. (2016). Effects of marijuana use on brain structure and function: Neuroimaging findings from a neurodevelopmental perspective. *International Review of Neurobiology*, *129*, 33–65.

Brzoska, L., Fischer, M., & Lentz, H. H. K. (2020). Hierarchical structures in livestock trade networks—A stochastic block model of the German cattle trade network. *Frontiers in Veterinary Science*, *7*, 281.

Thang Bui, Chaudhuri, S., Leighton, T., & Sipser, M. (1984). Graph bisection algorithms with good average case behavior. In *25th annual symposium on foundations of computer science* (pp. 181–192).

Burgess, M., Adar, E., & Cafarella, M. (2016). Link-prediction enhanced consensus clustering for complex networks. *PLoS One*, *11*, 1–23.

Cai, T. T., & Li, X. (2015). Robust and computationally feasible community detection in the presence of arbitrary outlier nodes. *The Annals of Statistics*, *43*, 1027–1059.

Cao, Y., & Chen, D. (2012). Generalization errors of Laplacian regularized least squares regression. *Science China Mathematics*, *55*, 1859–1868.

Cao, Y., & Chen, D.-R. (2011). Consistency of regularized spectral clustering. *Applied and Computational Harmonic Analysis*, *30*, 319–336.

Cazabet, R., & Amblard, F. (2014). Dynamic community detection. In *Encyclopedia of social network analysis and mining* (pp. 404–414). Springer.

Celisse, A., Daudin, J.-J., & Pierre, L. (2012). Consistency of maximum-likelihood and variational estimators in the stochastic block model. *Electronic Journal of Statistics*, *6*, 1847–1899.

Chen, K., & Lei, J. (2018). Network cross-validation for determining the number of communities in network data. *Journal of the American Statistical Association*, *113*, 241–251.

Chen, Y., Kamath, G., Suh, C., & Tse, D. (2016). Community recovery in graphs with locality. In *International conference on machine learning* (pp. 689–698).

Chen, Y., Li, X., & Xu, J. (2018). Convexified modularity maximization for degree-corrected stochastic block models. *The Annals of Statistics*, *46*, 1573–1602.

Chien, I., Lin, C.-Y., and Wang, I.-H. (2018). Community detection in hypergraphs: Optimal statistical limit and efficient algorithms. In *International conference on artificial intelligence and statistics* (pp. 871–879).

Chodrow, P., & Mellor, A. (2020). Annotated hypergraphs: Models and applications. *Applied Network Science*, *5*, 9.

Chung, F., & Lu, L. (1997). *Spectral graph theory*. American Mathematical Society.

Clauset, A. (2013). *Lecture notes in network analysis and modeling*. Santa Fe Institute.

Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, *70*, 066111.

Combe, D., Largeron, C., Jeudy, B., Fogelman-Soulié, F., & Wang, J. (2020). *Attributed networks partitioning based on modularity optimization*, chap. 8 (pp. 169–185). John Wiley & Sons, Ltd.

Côme, E., Latouche, P., Jouvin, N., & Bouveyron, C. (2020). *Hierarchical clustering with discrete latent variable models and the integrated classification likelihood*. arXiv preprint arXiv:2002.11577.

Contisciani, M., Power, E., & De Bacco, C. (2020). Community detection with node attributes in multilayer networks. *Scientific Reports*, *10*, 15736.

Corneli, M., Latouche, P., & Rossi, F. (2016). Exact ICL maximization in a non-stationary temporal extension of the stochastic block model for dynamic networks. *Neurocomputing*, *192*, 81–91.

Cuesta-Albertos, J. A., Febrero-Bande, M., & de la Fuente, M. O. (2017). The $DD^g$-classifier in the functional setting. *Test*, *26*, 119–142.

Dafir, Z., Lamari, Y., & Slaoui, S. C. (2020). A survey on parallel clustering algorithms for big data. *Artificial Intelligence Review*, *54*, 2411–2443.

Dakiche, N., Tayeb, F. B.-S., Slimani, Y., & Benatchba, K. (2019). Tracking community evolution in social networks: A survey. *Information Processing & Management*, *56*, 1084–1102.

Dall'Amico, L., Couillet, R., & Tremblay, N. (2020). Optimal Laplacian regularization for sparse spectral community detection. In *ICASSP 2020 – 2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 3237–3241).

Damle, A., Minden, V., & Ying, L. (2018). Simple, direct and efficient multi-way spectral clustering. *Information and Inference: A Journal of the IMA*, *8*, 181–203.

De Domenico, M., Granell, C., Porter, M., & Arenas, A. (2016). The physics of spreading processes in multilayer networks. *Nature Physics*, *12*, 901–906.

De Domenico, M., Lancichinetti, A., Arenas, A., & Rosvall, M. (2015). Identifying modular flows on multilayer networks reveals highly overlapping organization in interconnected systems. *Physical Review X*, *5*, 011027.

DeFord, D. R., & Pauls, S. D. (2019). Spectral clustering methods for multiplex networks. *Physica A: Statistical Mechanics and its Applications*, *533*, 121949.

Dey, A., Gel, Y. R., & Poor, H. V. (2017). Intentional islanding of power grids with data depth. *In 2017 IEEE 7th international workshop on computational advances in multi-sensor adaptive processing (CAMSAP)* (pp. 1–5).

Dey, A. K., Akcora, C., Gel, Y., & Kantarcioglu, M. (2020). On the role of local blockchain network features in cryptocurrency price formation: Local network features in cryptocurrency price analysis. *Canadian Journal of Statistics*, *48*, 561–581.

Dey, A. K., Gel, Y. R., & Poor, H. V. (2019). What network motifs tell us about resilience and reliability of complex networks. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 19368–19373.

Duan, D., Li, Y., Jin, Y., & Lu, Z. (2009). Community mining on dynamic weighted directed graphs. In *Proceedings of the 1st ACM international workshop on complex networks meet information & knowledge management* (pp. 11–18). ACM.

Estrada, E., & Rodríguez-Velázquez, J. A. (2006). Subgraph centrality and clustering in complex hyper-networks. *Physica A: Statistical Mechanics and its Applications*, *364*, 581–594.

Faskowitz, J., Yan, X., Zuo, X.-N., & Sporns, O. (2018). Weighted stochastic block models of the human connectome across the life span. *Scientific Reports*, *8*, 12997.

Filbey, F. M., Aslan, S., Calhoun, V. D., Spence, J. S., Damaraju, E., Caprihan, A., & Segall, J. (2014). Long-term effects of marijuana use on the brain. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 16913–16918.

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, *486*, 75–174.

Fortunato, S., & Barthélemy, M. (2007). Resolution limit in community detection. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 36–41.

Fortunato, S., & Hric, D. (2016). Community detection in networks: A user guide. *Physics Reports*, *659*, 1–44.

Fügenschuh, M., Gera, R., & Tagarelli, A. (2021). Topological analysis of synthetic models for air transportation multilayer networks. In R. M. Benito, C. Cherifi, H. Cherifi, E. Moro, L. M. Rocha, & M. Sales-Pardo (Eds.), *Complex networks & their applications IX* (pp. 206–217). Springer International Publishing.

Funke, T., & Becker, T. (2019). Stochastic block models: A comparison of variants and inference methods. *PLoS One*, *14*, 1–40.

Gao, C., Ma, Z., Zhang, A. Y., & Zhou, H. H. (2018). Community detection in degree-corrected block models. *The Annals of Statistics*, *46*, 2153–2185.

Gao, J., Zhao, Q., Ren, W., Swami, A., Ramanathan, R., & Bar-Noy, A. (2015). Dynamic shortest path algorithms for hypergraphs. *IEEE/ACM Transactions on Networking*, *23*, 1805–1817.

Gao, Y., Wang, M., Tao, D., Ji, R., & Dai, Q. (2012). 3-D object retrieval and recognition with hypergraph analysis. *IEEE Transactions on Image Processing*, *21*, 4290–4303.

Gao, Z.-K., Yang, Y.-X., Zhai, L.-S., Weidong, D., Yu, J.-L., & Jin, N. (2016). Multivariate multiscale complex network analysis of vertical upward oil-water two-phase flow in a small diameter pipe. *Scientific Reports*, *6*, 20052.

Garcia, C. (2016). Boclust: Bootstrap clustering stability algorithm for community detection. *PLoS One*, *11*, 1–15.

Garcia, J. O., Ashourvan, A., Muldoon, S., Vettel, J. M., & Bassett, D. S. (2018). Applications of community detection techniques to brain graphs: Algorithmic considerations and implications for neural function. *Proceedings of the IEEE*, *106*, 846–867.

García-Escudero, L. A., & Gordaliza, A. (1999). Robustness properties of *k* means and trimmed *k* means. *Journal of the American Statistical Association*, *94*, 956–969.

Ghasemian, A., Hosseinmardi, H., & Clauset, A. (2019). Evaluating overfit and underfit in models of network community structure. *IEEE Transactions on Knowledge and Data Engineering.*, *32*, 1.

Ghasemian, A., Zhang, P., Clauset, A., Moore, C., & Peel, L. (2016). Detectability thresholds and optimal algorithms for community structure in dynamic networks. *Physical Review X*, *6*, 031005.

Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. *PNAS*, *99*, 7821–7826.

Goldsmith, D., Grauer, K., & Shmalo, Y. (2020). Analyzing hack subnetworks in the bitcoin transaction graph. *Applied Network Science*, *5*, 1–20.

Görke, R., Maillard, P., Staudt, C., & Wagner, D. (2010). Modularity-driven clustering of dynamic graphs. In *International symposium on experimental algorithms* (pp. 436–448). Springer.

Grossman, J., & Ion, P. (1995). On a portion of the well-known collaboration graph. *Congressus Numerantium*, *108*, 129–131.

Guédon, O., & Vershynin, R. (2016). Community detection in sparse networks via Grothendieck's inequality. *Probability Theory and Related Fields*, *165*, 1025–1049.

Gulikers, L., Lelarge, M., & Massoulié, L. (2017). A spectral method for community detection in moderately sparse degree-corrected stochastic block models. *Advances in Applied Probability*, *49*, 686–721.

Hao, Y., Wang, Q., Duan, Z., & Chen, G. (2019). Controllability of Kronecker product networks. *Automatica*, *110*, 108597.

Hartwell, L., Hopfield, J., Leibler, S., & Murray, A. (1999). From molecular to modular cell biology. *Nature*, *402*, C47–C52.

He, Z., Liang, H., Chen, Z., Zhao, C., & Liu, Y. (2020). Computing exact *p*-values for community detection. *Data Mining and Knowledge Discovery*, *34*, 833–869.

Helal, N. A., Ismail, R. M., Badr, N. L., & Mostafa, M. G. M. (2017). Leader-based community detection algorithm for social networks. *WIREs Data Mining and Knowledge Discovery*, *7*, e1213.

Hidalgo, S. J. T., & Ma, S. (2018). Clustering multilayer omics data using MuNCut. *BMC Genomics*, *19*, 1–13.

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

Hollocou, A., Bonald, T., & Lelarge, M. (2019). Modularity-based sparse soft graph clustering. In *AISTATS 2019 – 22nd international conference on artificial intelligence and statistics*. Naha, Okinawa, Japan.

Holme, P., & Saramäki, J. (2012). Temporal networks. *Physics Reports*, *519*, 97–125.

Holzinger, A., Haibe-Kains, B., & Jurisica, I. (2019). Why imaging data alone is not enough: AI-based integration of imaging, omics, and clinical data. *European Journal of Nuclear Medicine and Molecular Imaging*, *46*, 2722–2730.

Hua, J., Yu, J., & Yang, M.-S. (2020). Fast clustering for signed graphs based on random walk gap. *Social Networks*, *60*, 113–128.

Huang, L., Wang, C.-D., & Chao, H.-Y. (2019). Higher-order multi-layer community detection. In *AAAI*.

Huffman, D. A. (1952). A method for the construction of minimum-redundancy codes. *Proceedings of the Institute of Radio Engineers*, *40*, 1098–1101.

Huo, J., Weng, J., & Qu, H. (2019). A parallel clustering algorithm for logs data based on Hadoop platform. In *Proceedings of the 3rd international conference on high performance compilation, computing and communications* (pp. 90–94).

Hurd, Y. L., Manzoni, O. J., Pletnikov, M. V., Lee, F. S., Bhattacharyya, S., & Melis, M. (2019). Cannabis and the developing brain: Insights into its long-lasting effects. *Journal of Neuroscience*, *39*, 8250–8258.

Jasiński, M., Sikorski, T., Leonowicz, Z., Borkowski, K., & Jasińska, E. (2020). The application of hierarchical clustering to power quality measurements in an electrical power network with distributed generation. *Energies*, *13*, 2407.

Joseph, A., & Yu, B. (2016). Impact of regularization on spectral clustering. *Annals of Statistics*, *44*, 1765–1791.

Jourdan, M., Blandin, S., Wynter, L., & Deshpande, P. (2018). Characterizing entities in the bitcoin blockchain. In *2018 IEEE international conference on data mining workshops (ICDMW)* (pp. 55–62). IEEE.

Karaaslanli, A., & Aviyente, S. (2020). Constrained spectral clustering for dynamic community detection. In *ICASSP 2020 – 2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 8474–8478).

Karrer, B., & Newman, M. E. J. (2011). Stochastic blockmodels and community structure in networks. *Physical Review E*, *83*, 016107.

Kernighan, B. W., & Lin, S. (1970). An efficient heuristic procedure for partitioning graphs. *The Bell System Technical Journal*, *49*, 291–307.

Kim, M., & Leskovec, J. (2010). Multiplicative attribute graph model of real-world networks. In R. Kumar & D. Sivakumar (Eds.), *Algorithms and models for the web-graph* (pp. 62–73). Springer Berlin Heidelberg.

Kim, M., & Leskovec, J. (2011). Modeling social networks with node attributes using the multiplicative attribute graph model. In *Proceedings of the twenty-seventh conference on uncertainty in artificial intelligence, UAI'11* (pp. 400–409). Arlington, VA: AUAI Press.

Kim, M., & Leskovec, J. (2012). Multiplicative attribute graph model of real-world networks. *Internet Mathematics*, *8*, 113–160.

Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., & Porter, M. A. (2014). Multilayer networks. *Journal of Complex Networks*, *2*, 203–271.

Kumar, A., Rai, P., & Daumé III, H. (2010). Co-regularized spectral clustering with multiple kernels. In *NIPS 2010 workshop: New directions in multiple kernel learning*.

Kumar, T., Vaidyanathan, S., Ananthapadmanabhan, H., Parthasarathy, S., & Ravindran, B. (2020). Hypergraph clustering by iteratively reweighted modularity maximization. *Applied Network Science*, *5*, 52.

Kumpula, J., Saramäki, J., Kaski, K., & Kertész, J. (2007). Limited resolution in complex network community detection with Potts model approach. *European Physical Journal B. Condensed Matter and Complex Systems*, *56*, 41–45.

Kuncheva, Z., & Montana, G. (2015). Community detection in multiplex networks using locally adaptive random walks. In *2015 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)* (pp. 1308–1315).

Lancichinetti, A., & Fortunato, S. (2009). Community detection algorithms: A comparative analysis. *Physical Review E*, *80*, 056117.

Lancichinetti, A., & Fortunato, S. (2011). Limits of modularity maximization in community detection. *Physical Review E*, *84*, 066122.

Lancichinetti, A., Radicchi, F., Ramasco, J. J., & Fortunato, S. (2011). Finding statistically significant communities in networks. *PLoS One*, *6*, 1–18.

Latapy, M., Viard, T., & Magnien, C. (2018). Stream graphs and link streams for the modeling of interactions over time. *Social Network Analysis and Mining*, *8*, 1–29.

Latouche, P., Birmelé, E., & Ambroise, C. (2011). Overlapping stochastic block models with application to the FRENCH political blogosphere. *The Annals of Applied Statistics*, *5*, 309–336.

Le, C. M., Levina, E., & Vershynin, R. (2017). Concentration and regularization of random graphs. *Random Structures & Algorithms*, *51*, 538–561.

Lee, C., & Wilkinson, D. J. (2019). A review of stochastic block models and extensions for graph clustering. *Applied Network Science*, *4*, 1–50.

Lee, J. R., Gharan, S. O., & Trevisan, L. (2014). Multiway spectral partitioning and higher-order Cheeger inequalities. *Journal of the ACM*, *61*, 1–30.

Lei, J., & Rinaldo, A. (2015). Consistency of spectral clustering in stochastic block models. *The Annals of Statistics*, *43*, 215–237.

Leitch, J., Alexander, K., & Sengupta, S. (2019). Toward epidemic thresholds on temporal networks: A review and open questions. *Applied Network Science*, *4*, 105.

Leskovec, J. (2009). Networks, communities and Kronecker products. In *Proceedings of the 1st ACM international workshop on complex networks meet information & knowledge management, CNIKM '09* (pp. 1–2). New York, NY: Association for Computing Machinery.

Leskovec, J., Chakrabarti, D., Kleinberg, J., & Faloutsos, C. (2005). Realistic, mathematically tractable graph generation and evolution, using kronecker multiplication. In A. M. Jorge, L. Torgo, P. Brazdil, R. Camacho, & J. Gama (Eds.), *Knowledge discovery in databases: PKDD 2005* (pp. 133–145). Springer Berlin Heidelberg.

Leskovec, J., Chakrabarti, D., Kleinberg, J., Faloutsos, C., & Ghahramani, Z. (2010). Kronecker graphs: An approach to modeling networks. *Journal of Machine Learning Research*, *11*, 985–1042.

Leskovec, J., & Faloutsos, C. (2007). Scalable modeling of real graphs using Kronecker multiplication. In *Proceedings of the 24th international conference on machine learning, ICML '07* (pp. 497–504). New York, NY: Association for Computing Machinery.

Leskovec, J., Lang, K. J., Dasgupta, A., & Mahoney, M. W. (2008). Statistical properties of community structure in large social and information networks. In *Proceedings of the 17th international conference on World Wide Web, WWW '08* (pp. 695–704). New York, NY: Association for Computing Machinery.

Li, C.-B., Yang, H., & Komatsuzaki, T. (2008). Multiscale complex network of protein conformational fluctuations in single-molecule time series. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 536–541.

Li, J., Cuesta-Albertos, J. A., & Liu, R. Y. (2012). DD-classifier: Nonparametric classification procedure based on DD-plot. *Journal of the American Statistical Association*, *107*, 737–753.

Li, P., Dau, H., Puleo, G., & Milenkovic, O. (2017). Motif clustering and overlapping clustering for social network analysis. In *IEEE INFOCOM 2017 – IEEE conference on computer communications* (pp. 1–9).

Li, P., & Milenkovic, O. (2017). Inhomogeneous hypergraph clustering with applications. In *Advances in neural information processing systems*.

Li, P.-Z., Huang, L., Wang, C.-D., & Lai, J.-H. (2019). EdMot: An edge enhancement approach for motif-aware community detection. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining, KDD '19* (pp. 479–487). New York, NY, USA: Association for Computing Machinery.

Li, P.-Z., Huang, L., Wang, C.-D., Lai, J.-H., & Huang, D. (2020). Community detection by motif-aware label propagation. *ACM Transactions on Knowledge Discovery from Data*, *14*, 1–19.

Li, Q., Dong, S., & Mostafavi, A. (2019). Modeling of inter-organizational coordination dynamics in resilience planning of infrastructure systems: A multilayer network simulation framework. *PLoS One*, *14*, 1–21.

Li, T., Levina, E., & Zhu, J. (2020). Network cross-validation by edge sampling. *Biometrika*, *107*, 257–276.

Li, W., Ahn, S., & Welling, M. (2016). Scalable MCMC for mixed membership stochastic blockmodels. In A. Gretton & C. C. Robert (Eds.), *Proceedings of the 19th international conference on artificial intelligence and statistics, vol. 51 of proceedings of machine learning research* (pp. 723–731). PMLR.

Liu, F., Choi, D., Xie, L., & Roeder, K. (2018). Global spectral clustering in dynamic networks. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 927–932.

Liu, N., Yao, L., & Zhao, X. (2020). Evaluating the amygdala network induced by neurofeedback training for emotion regulation using hierarchical clustering. *Brain Research*, *1740*, 146853.

Liu, X. F., Jiang, X.-J., Liu, S.-H., & Tse, C. K. (2020). *Knowledge discovery in cryptocurrency transactions: A survey*. arXiv:2010.01031.

Louis, A., Raghavendra, P., Tetali, P., & Vempala, S. (2011). Algorithmic extensions of Cheeger's inequality to higher eigenvalues and partitions. In L. A. Goldberg, K. Jansen, R. Ravi, & J. D. P. Rolim (Eds.), *Approximation, randomization, and combinatorial optimization. Algorithms and techniques* (pp. 315–326). Springer Berlin Heidelberg.

Lyzinski, V., Sussman, D. L., Tang, M., Athreya, A., & Priebe, C. E. (2014). Perfect clustering for stochastic blockmodel graphs via adjacency spectral embedding. *Electronic Journal of Statistics*, *8*, 2905–2922.

Lyzinski, V., Tang, M., Athreya, A., Park, Y., & Priebe, C. E. (2017). Community detection and classification in hierarchical stochastic blockmodels. *IEEE Transactions on Network Science and Engineering*, *4*, 13–26.

Ma, T., Wang, J., & Zhang, M. (2019). The restricted edge-connectivity of Kronecker product graphs. *Parallel Processing Letters*, *29*, 1950012.

Mahapatra, S., & Chandola, V. (2015). Modeling graphs using a mixture of Kronecker models. In *2015 IEEE international conference on big data (big data)* (pp. 727–736).

Makarychev, K., Makarychev, Y., & Vijayaraghavan, A. (2016). Learning communities in the presence of errors. In *Conference on learning theory* (pp. 1258–1291).

Malliaros, F. D., & Vazirgiannis, M. (2013). Clustering and community detection in directed networks: A survey. *Physics Reports*, *533*, 95–142.

Märtens, M., Meier, J., Hillebrand, A., Tewarie, P., & Mieghem, P. V. (2017). Brain network clustering with information flow motifs. *Applied Network Science*, *2*, 25.

Martinet, L.-E., Kramer, M., Viles, W., Perkins, L., Spencer, E., Chu, C., Cash, S., & Kolaczyk, E. (2020). Robust dynamic community detection with applications to human brain functional networks. *Nature Communications*, *11*, 2785.

Matias, C., & Miele, V. (2017). Statistical clustering of temporal networks through a dynamic stochastic block model. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *79*, 1119–1141.

Matuszewski, P., & Szabó, G. (2019). Are echo chambers based on partisanship? Twitter and political polarity in Poland and Hungary. *Social Media+ Society*, *5*, 2056305119837671.

Meier, J., Märtens, M., Hillebrand, A., Tewarie, P., & Van Mieghem, P. (2017). Motif-based analysis of effective connectivity in brain networks. In H. Cherifi, S. Gaito, W. Quattrociocchi, & A. Sala (Eds.), *Complex networks & their applications V* (pp. 685–696). Springer International Publishing.

Mémoli, F., & Pinto, G. V. F. (2020). *Motivic clustering schemes for directed graphs*. arXiv preprint arXiv:2001.00278.

Menck, P. J., Heitzig, J., Kurths, J., & Schellnhuber, H. J. (2014). How dead ends undermine power grid stability. *Nature Communications*, *5*, 3969.

Mercado, P., Tudisco, F., and Hein, M. (2019). Spectral clustering of signed graphs via matrix power means. In *Proceedings of the 36th international conference on machine learning*.

Messaoudi, I., & Kamel, N. (2019). A multi-objective bat algorithm for community detection on dynamic social networks. *Applied Intelligence*, *49*, 2119–2136.

Messina, F., Giombini, E., Agrati, C., Vairo, F., Bartoli, T. A., Al Moghazi, S., Piacentini, M., Locatelli, F., Kobinger, G., Maeurer, M., Zumla, A., Capobianchi, M. R., Lauria, F. N., Ippolito, G., & COVID 19 INMI Network Medicine for IDs Study Group. (2020). Covid-19: Viral–host interactome analyzed by network based-approach model to study pathogenesis of SARS-CoV-2 infection. *Journal of Translational Medicine*, *18*, 1–10.

Meyerhenke, H., Sanders, P., & Schulz, C. (2017). Parallel graph partitioning for complex networks. *IEEE Transactions on Parallel and Distributed Systems*, *28*, 2625–2638.

Michelitsch, T., Riascos, A. P., Collet, B., Nicolleau, F., & Nowakowski, A. (2019). *Fractional dynamics on networks and lattices*. Wiley Online Library.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. (2002). Network motifs: Simple building blocks of complex networks. *Science*, *298*, 824–827.

Moitra, A., Perry, W., & Wein, A. S. (2016). How robust are reconstruction thresholds for community detection? In *Proceedings of the forty-eighth annual ACM symposium on theory of computing* (pp. 828–841).

Moreno, S., Kirshner, S., Neville, J., & Vishwanathan, S. V. N. (2010). Tied Kronecker product graph models to capture variance in network populations. In *2010 48th annual Allerton conference on communication, control, and computing (Allerton)* (pp. 1137–1144).

Moreno, S., Robles, P., & Neville, J. (2013). Block Kronecker product graph model.

Mosler, K. (2013). Depth statistics. In C. Becker, R. Fried, & S. Kuhnt (Eds.), *Robustness and complex data structures* (pp. 17–34). Springer.

Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., & Onnela, J.-P. (2010). Community structure in time-dependent, multiscale, and multiplex networks. *Science*, *328*, 876–878.

Nadler, B., Srebro, N., & Zhou, X. (2009). Semi-supervised learning with the graph Laplacian: The limit of infinite unlabelled data. In *Neural information processing systems (NIPS)*.

Newman, M. E. J. (2004). Fast algorithm for detecting community structure in networks. *Physics Review*, *69*, 066133.

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

Nguyen, N. P., Dinh, T. N., Shen, Y., & Thai, M. T. (2014). Dynamic social community detection and its applications. *PLoS One*, *9*, e91431.

Nielsen, F. (2016). *Hierarchical clustering* (pp. 195–211). Springer International Publishing.

Nowicki, K., & Snijders, T. A. B. (2001). Estimation and prediction for stochastic blockstructures. *Journal of the American Statistical Association*, *96*, 1077–1087.

Okuda, M., Satoh, S., Sato, Y., & Kidawara, Y. (2019). Community detection using restrained random-walk similarity. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*, 89–103.

Oliveira, M., & Gama, J. (2012). An overview of social network analysis. *WIREs Data Mining and Knowledge Discovery*, *2*, 99–115.

Ouvrard, X. (2020). *Hypergraphs: An introduction and review*. arXiv preprint:2002.05014.

Paez, M. S., Amini, A. A., & Lin, L. (2019). *Hierarchical stochastic block model for community detection in multiplex networks*. arXiv preprint arXiv:1904.05330.

Palla, G., Barabási, A.-L., & Vicsek, T. (2007). Quantifying social group evolution. *Nature*, *446*, 664–667.

Palla, G., Pollner, P., Barabási, A.-L., & Vicsek, T. (2009). Social group dynamics in networks. In *Adaptive networks* (pp. 11–38). Springer.

Pamfil, A. R., Howison, S. D., Lambiotte, R., & Porter, M. A. (2019). Relating modularity maximization and stochastic block models in multilayer networks. *SIAM Journal on Mathematics of Data Science*, *1*, 667–698.

Pan, X., Papailiopoulos, D., Oymak, S., Recht, B., Ramchandran, K., & Jordan, M. I. (2015). Parallel correlation clustering on big graphs. In *Proceedings of the 28th international conference on neural information processing systems – Volume 1, NIPS'15* (pp. 82–90). Cambridge, MA: MIT Press.

Paul, S., & Chen, Y. (2016a). Consistent community detection in multi-relational data through restricted multi-layer stochastic blockmodel. *Electronic Journal of Statistics*, *10*, 3807–3870.

Paul, S., & Chen, Y. (2016b). *Null models and modularity based community detection in multi-layer networks*. arXiv preprint arXiv:1608.00623.

Paul, S., & Chen, Y. (2020). Spectral and matrix factorization methods for consistent community detection in multi-layer networks. *The Annals of Statistics*, *48*, 230–250.

Pavlović, D. M., Vértes, P. E., Bullmore, E. T., Schafer, W. R., & Nichols, T. E. (2014). Stochastic blockmodeling of the modules and core of the *Caenorhabditis elegans* connectome. *PLoS One*, *9*, e97584.

Peel, L., Delvenne, J.-C., & Lambiotte, R. (2018). Multiscale mixing patterns in networks. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 4057–4062.

Peixoto, T. P. (2015). Model selection and hypothesis testing for large-scale network models with overlapping groups. *Physical Review X*, *5*, 011033.

Peixoto, T. P. (2018). Nonparametric weighted stochastic block models. *Physical Review E*, *97*, 012306.

Pensky, M., & Zhang, T. (2019). Spectral clustering in the dynamic stochastic block model. *Electronic Journal of Statistics*, *13*, 678–709.

Pfeiffer, J. J., Moreno, S., La Fond, T., Neville, J., & Gallagher, B. (2014). Attributed graph models: Modeling network structure with correlated attributes. In *Proceedings of the 23rd international conference on World Wide Web, WWW '14* (pp. 831–842). New York, NY: Association for Computing Machinery.

Pilosof, S., Porter, M. A., Pascual, M., & Kéfi, S. (2017). The multilayer nature of ecological networks. *Nature Ecology & Evolution*, *1*, 0101.

Pons, P., & Latapy, M. (2005). Computing communities in large networks using random walks. In *International symposium on computer and information sciences* (pp. 284–293). Springer.

Pržulj, N. (2007). Biological network comparison using graphlet degree distribution. *Bioinformatics*, *23*, e177–e183.

Purkait, P., Chin, T., Sadri, A., & Suter, D. (2017). Clustering with hypergraphs: The case for large hyperedges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*, 1697–1711.

Qian, X., Chen, Y., & Minca, A. (2019). *Clustering degree-corrected stochastic block model with outliers*. arXiv preprint arXiv:1906.03305.

Qu, S., & Makowski, A. M. (2019). *Node isolation in large homogeneous binary multiplicative attribute graph models*.

Radcliffe, M., & Young, S. J. (2014). The spectra of multiplicative attribute graphs. *Linear Algebra and its Applications*, *462*, 39–58.

Ranshous, S., Joslyn, C. A., Kreyling, S., Nowak, K., Samatova, N. F., West, C. L., & Winters, S. (2017). Exchange pattern mining in the bitcoin transaction directed hypergraph. In *International conference on financial cryptography and data security* (pp. 248–263). Springer.

Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., & Barabási, A.-L. (2002). Hierarchical organization of modularity in metabolic networks. *Science*, *297*, 1551–1555.

Reichardt, J., & Bornholdt, S. (2006). Statistical mechanics of community detection. *Physical Review E*, *74*, 016110.

Riascos, A. P., & Mateos, J. L. (2012). Long-range navigation on complex networks using lévy random walks. *Physical Review E*, *86*, 056110.

Ribeiro, P., Paredes, P., Silva, M. E., Aparicio, D., & Silva, F. (2019). *A survey on subgraph counting: Concepts, algorithms and applications to network motifs and graphlets*. arXiv preprint arXiv:1910.13011.

Riolo, M. A., & Newman, M. E. J. (2014). First-principles multiway spectral partitioning of graphs. *Journal of Complex Networks*, *2*, 121–140.

Rocklin, M., & Pinar, A. (2013). On clustering on graphs with multiple edge types. *Internet Mathematics*, *9*, 82–112.

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

Rossetti, G., & Cazabet, R. (2018). Community discovery in dynamic networks: A survey. *ACM Computing Surveys (CSUR)*, *51*, 1–37.

Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 1118–1123.

Roy, S., Ghosh, P., Barua, D., & Das, S. K. (2020). Motifs enable communication efficiency and fault-tolerance in transcriptional networks. *Scientific Reports*, *10*, 1–15.

Saldaña, D. F., Yu, Y., & Feng, Y. (2017). How many communities are there? *Journal of Computational and Graphical Statistics*, *26*, 171–181.

Sanna Passino, F., & Heard, N. (2020). Bayesian estimation of the latent dimension and communities in stochastic blockmodels. *Statistics and Computing*, *30*, 1291–1307.

Schultz, P., Heitzig, J., & Kurths, J. (2014). Detours around basin stability in power networks. *New Journal of Physics*, *16*, 125001.

Seshadhri, C., Pinar, A., & Kolda, T. G. (2013). An in-depth analysis of stochastic Kronecker graphs. *Journal of the ACM*, *60*, 1–32.

Shen-Orr, S. S., Milo, R., Mangan, S., & Alon, U. (2002). Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature Genetics*, *31*, 64–68.

Shun, J., Roosta-Khorasani, F., Fountoulakis, K., & Mahoney, M. W. (2016). Parallel local graph clustering. *Proceedings of the VLDB Endowment*, *9*, 1041–1052.

Stone, L., Simberloff, D., & Artzy-Randrup, Y. (2019). Network motifs and their origins. *PLoS Computational Biology*, *15*, e1006749.

Tagarelli, A., Amelio, A., & Gullo, F. (2017). Ensemble-based community detection in multilayer networks. *Data Mining and Knowledge Discovery*, *31*, 1506–1543.

Takai, Y., Miyauchi, A., Ikeda, M., & Yoshida, Y. (2020). *Hypergraph clustering based on PageRank* (pp. 1970–1978). Association for Computing Machinery.

Takes, F. W., Kosters, W. A., Witte, B., & Heemskerk, E. M. (2018). Multiplex network motifs as building blocks of corporate networks. *Applied Network Science*, *3*, 1–22.

Tang, L., Jing, K., He, J., & Stanley, H. (2015). Complex interdependent supply chain networks: Cascading failure and robustness. *Physica A: Statistical Mechanics and its Applications*, *443*, 58–69.

Tarjan, R. (1971). Depth-first search and linear graph algorithms. In *12th annual symposium on switching and automata theory (SWAT 1971)* (pp. 114–121).

Temkin, O., Zeigarnik, A., & Bonchev, D. (1996). *Chemical reaction networks: A graph-theoretical approach*. Taylor & Francis.

Tian, Y., & Gel, Y. (2019). Fusing data depth with complex networks: Community detection with prior information. *Computational Statistics & Data Analysis*, *139*, 99–116.

Tian, Y., & Gel, Y. R. (2017). Fast community detection in complex networks with a k-depths classifier. In S. E. Ahmed (Ed.), *Big and complex data analysis: Methodologies and applications* (pp. 139–157). Springer.

Tien, J. H., Eisenberg, M. C., Cherng, S. T., & Porter, M. A. (2020). Online reactions to the 2017 'unite the right' rally in Charlottesville: Measuring polarization in twitter networks using media followership. *Applied Network Science*, *5*, 1–27.

Trevisan, L. (2013). *Is Cheeger-type approximation possible for nonuniform sparsest cut*? arXiv preprint arXiv:1303.2730.

Tsourakakis, C. E., Pachocki, J., & Mitzenmacher, M. (2017). Scalable motif-aware graph clustering. In *Proceedings of the 26th international conference on World Wide Web, WWW '17* (pp. 1451–1460). Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee.

Vajdi, A., Juher, D., Saldaña, J., & Scoglio, C. (2020). A multilayer temporal network model for STD spreading accounting for permanent and casual partners. *Scientific Reports*, *10*, 3846.

Vallès-Català, T., Massucci, F. A., Guimerà, R., & Sales-Pardo, M. (2016). Multilayer stochastic block models reveal the multilayer structure of complex networks. *Physical Review X*, *6*, 011036.

Van Lierde, H., Chow, T. W. S., & Chen, G. (2020). Scalable spectral clustering for overlapping community detection in large-scale networks. *IEEE Transactions on Knowledge and Data Engineering*, *32*, 754–767.

Vardi, Y., & Zhang, C.-H. (2000). The multivariate L1-median and associated data depth. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 1423–1426.

Viard, T., Latapy, M., & Magnien, C. (2016). Computing maximal cliques in link streams. *Theoretical Computer Science*, *609*, 245–252.

Victor, F. (2020). Address clustering heuristics for Ethereum. In J. Bonneau & N. Heninger (Eds.), *Financial cryptography and data security* (pp. 617–633). Springer International Publishing.

Voloshin, V. (2009). *Introduction to graph and hypergraph theory*. Matematicas (E-libro). Nova Science Publishers.

von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics & Computing*, *17*, 395–416.

Vu, D. Q., Hunter, D. R., & Schweinberger, M. (2013). Model-based clustering of large networks. *The Annals of Applied Statistics*, *7*, 1010–1039.

Wang, F., Wang, X., Shao, B., Li, T., & Ogihara, M. (2009). Tag integrated multi-label music style classification with hypergraph. In *ISMIR*.

Wang, Y., Zhu, L., Qian, X., & Han, J. (2018). Joint hypergraph learning for tag-based image retrieval. *IEEE Transactions on Image Processing*, *27*, 4437–4451.

Wang, Y. X. R., & Bickel, P. J. (2017). Likelihood-based model selection for stochastic block models. *The Annals of Statistics*, *45*, 500–528.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*, 440–442.

Weaver, I. S., Williams, H., Cioroianu, I., Williams, M., Coan, T., & Banducci, S. (2018). Dynamic social media affiliations among UKpoliticians. *Social Networks*, *54*, 132–144.

Wei, F., Qian, W., Wang, C., & Zhou, A. (2009). Detecting overlapping community structures in networks. *World Wide Web*, *12*, 235–261.

Weichsel, P. (1962). The Kronecker product of graphs. *Proceedings of the American Mathematical Society*, *13*, 47–52.

Weng, L., Menczer, F., & Ahn, Y.-Y. (2013). Virality prediction and community structure in social networks. *Scientific Reports*, *3*, 2522.

Wider, N., Garas, A., Scholtes, I., & Schweitzer, F. (2016). An ensemble perspective on multi-layer. *Networks*, *1*, 37–59.

Wong, S. W., Pastrello, C., Kotlyar, M., Faloutsos, C., & Jurisica, I. (2018). SDREGION: Fast spotting of changing communities in biological networks. In *Proceedings of the 24th ACM SIGKDD international conference on Knowledge discovery & data mining* (pp. 867–875). ACM.

Yang, K., Dong, X., & Toni, L. (2019). *Error analysis on graph Laplacian regularized estimator*. arXiv preprint arXiv:1902.03720.

Yang, Y., Xu, K.-J., & Hong, C. (2021). Network dynamics on the Chinese air transportation multilayer network. *International Journal of Modern Physics C*, *32*, 2150070.

Yenerdag, E. (2016). *Contagion analysis in European financial markets through the lens of weighted stochastic block model: Systematically important communities of financial institutions* (Ph.D. thesis).

Yildirimoglu, M., & Kim, J. (2017). Identification of communities in urban mobility networks using multi-layer graphs of network traffic. *Transportation Research Procedia*, *27*, 1034–1041.

Yuvaraj, M., Dey, A., Lyubchich, V., Gel, Y., & Poor, H. (2021). Topological clustering of multilayer networks. *Proceedings of the National Academy of Sciences of the United States of America* (in press). *118*, 1–9.

Zachary, W. W. (1977). An information flow model for conflict and fission in small groups. *Journal of Anthropological Research*, *33*, 452–473.

Zakrzewska, A., & Bader, D. A. (2015). A dynamic algorithm for local community detection in graphs. In *Proceedings of the 2015 IEEE/ACM international conference on advances in social networks analysis and mining 2015* (pp. 559–564). ACM.

Zhang, A. Y., & Zhou, H. H. (2016). Minimax rates of community detection in stochastic block models. *The Annals of Statistics*, *44*, 2252–2280.

Zhang, X., Tian, Y., Guan, G., & Gel, Y. (2021). Depth-based classification for relational data with multiple attributes. *Journal of Multivariate Analysis*, *184*, 104732.

Zhang, Y., Levina, E., & Zhu, J. (2020). Detecting overlapping communities in networks using spectral methods. *SIAM Journal on Mathematics of Data Science*, *2*, 265–283.

Zhao, Y. (2017). A survey on theoretical advances of community detection in networks. *WIREs Computational Statistics*, *9*, e1403.

Zhao, Y., Levina, E., & Zhu, J. (2012). Consistency of community detection in networks under degree-corrected stochastic block models. *The Annals of Statistics*, *40*, 2266–2292.

Zhou, D., Huang, J., & Schölkopf, B. (2006). Learning with hypergraphs: Clustering, classification, and embedding. In *Proceedings of the 19th international conference on neural information processing systems, NIPS'06* (pp. 1601–1608). Cambridge, MA: MIT Press.

Zhou, Z., & Amini, A. A. (2019). Analysis of spectral clustering algorithms for community detection: The general bipartite setting. *Journal of Machine Learning Research*, *20*(47), 1–47.

Zhu, L., Shen, J., Jin, H., Zheng, R., & Xie, L. (2015). Content-based visual landmark search via multimodal hypergraph learning. *IEEE Transactions on Cybernetics*, *45*, 2756–2769.

Zitnik, M., Sosič, R., Feldman, M. W., & Leskovec, J. (2019). Evolution of resilience in protein interactomes across the tree of life. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 4426–4433.

Znidi, F., Davarikia, H., Iqbal, K., & Barati, M. (2019). Multi-layer spectral clustering approach to intentional islanding in bulk power systems. *Journal of Modern Power Systems and Clean Energy*, *7*, 1044–1055.

Zuo, Y., & Serfling, R. (2000). General notions of statistical depth function. *The Annals of Statistics*, *28*, 461–482.

**How to cite this article:** Dey, A. K., Tian, Y., & Gel, Y. R. (2021). Community detection in complex networks: From statistical foundations to data science applications. *Wiley Interdisciplinary Reviews: Computational Statistics*, e1566. https://doi.org/10.1002/wics.1566