On Detection and Structural Reconstruction of Small-World Random Networks

T. Tony Cai, Tengyuan Liang, and Alexander Rakhlin

***Abstract*—In this paper, we study detection and fast reconstruction of the celebrated Watts-Strogatz (WS) small-world random graph model (Watts and Strogatz, 1998) which aims to describe real-world complex networks that exhibit both high clustering and short average length properties. The WS model with neighborhood size k and rewiring probability probability β can be viewed as a continuous interpolation between a deterministic ring lattice graph and the Erdős-Rényi random graph. We study the computational and statistical aspects of detection and recovery of the deterministic ring lattice structure (strong ties) in the presence of random connections (weak ties). The phase diagram in terms of (k, β) is shown to consist of several regions according to the difficulty of the problem. We propose distinct methods for these regions.**

***Index Terms*—Small world networks, random graphs, spectral analysis, detection, reconstruction, computational boundary.**

I. Introduction

The "small-world" phenomenon aims to describe real-world complex networks that exhibit both high clustering and short average length properties. While most of the pairs of nodes are not friends, any node can be reached from another in a small number of hops. The Watts-Strogatz (WS) model, introduced in (Newman and Watts, 1999; Watts and Strogatz, 1998), is a popular generative model for networks that exhibit the small-world phenomenon. The WS model interpolates between the two extremes—the regular lattice graph for high clustering on the one hand, and the random graph exhibiting the short chain property on the other. Considerable effort has been spent on studying the asymptotic statistical behavior (degree distribution, average path length, clustering coefficient, etc.) and the empirical performance of the WS model (Amaral et al., 2000; Barrat and Weigt, 2000; Latora and Marchiori, 2001; Van Der Hofstad, 2009; Watts, 1999). Successful applications of the WS model have been found in a range of disciplines, such as psychology (Milgram, 1967), epidemiology (Moore and Newman, 2000), medicine and health (Stam et al., 2007), to name a few. In one of the first algorithmic studies of small-world networks, Kleinberg (2000) investigated the theoretical difficulty of finding the shortest path between any two nodes when one is restricted to use local algorithms, and further extended the small-world notion to long range percolation on graphs (Benjamini and Berger, 2000; Coppersmith et al., 2002).

In the present paper, we study detection and reconstruction of small-world networks. Our focus is on both statistical and computational aspects of these problems. Given a network, the first challenge is to detect whether it enjoys the small-world property (i.e., high clustering and short average path), or whether the observation may simply be explained by the Erdős-Rényi random graph (the null hypothesis). The second question is concerned with the reconstruction of the neighborhood structure if the network does exhibit the small-world phenomenon. In the language of social network analysis, the detection problem corresponds to detecting the existence of strong ties (close friend connections) in the presence of weak ties (random connections). The more difficult reconstruction problem corresponds to distinguishing between strong and weak ties. Statistical and computational difficulties of both detection and reconstruction are due to the latent high-dimensional permutation matrix which blurs the natural ordering of the ring structure on the nodes.

Department of Statistics, The Wharton School, University of Pennsylvania.

Let us parametrize the WS model in the following way: the number of nodes is denoted by n, the neighborhood size by k, and the rewiring probability by β. Provided the adjacency matrix $A \in \mathbb{R}^{n \times n}$, we are interested in identifying the tuples (n, k, β) when detection and reconstruction of the small-world random graph is possible. Specifically, we focus on the following two questions.

Detection Given the adjacency matrix A up to a permutation, when (in terms of n, k, β) and how (in terms of procedures) can one statistically distinguish whether it is a small-world graph ($\beta < 1$), or a random graph with matching degree ($\beta = 1$). What can be said if we restrict our attention to computationally efficient procedures?

Reconstruction Once the presence of the neighborhood structure is confirmed, when (in terms of n, k, β) and how (in terms of procedures) can one estimate the deterministic neighborhood structure? If one only aims to estimate the structure asymptotically consistently, are there computationally efficient procedures, and what are their limitations?

We address the above questions by presenting a phase diagram in Figure 1. The phase diagram divides the parameter space into four disjoint regions according to the difficulty of the problem. We propose distinct methods for the regions where solutions are possible.

A. Why the Small-World Model?

Finding and analyzing the appropriate statistical models for real-world complex networks is one of the main themes in network science. Many real empirical networks—for example, internet architecture, social networks, and biochemical pathways—exhibit two features simultaneously: high clustering among individual nodes and short distance between any two nodes. Consider the local tree rooted at a person. The high clustering property suggests prevalent existence of triadic closure, which significantly reduces the number of reachable people within a certain depth (in contrast to the regular tree case where this number grows exponentially with the depth), contradicting the short average length property. In a pathbreaking paper, Watts and Strogatz (1998) provided a mathematical model that resolves the above seemingly contradictory notions. The solution is surprisingly simple — interpolating between structural ring lattice graph and a random graph. The ring lattice provides the strong ties (i.e., homophily, connection to people who are similar to us) and triadic closure, while the random graph generates the weak ties (connection to people who are otherwise far-away), preserving the local-regular-branching-tree-like structure that induces short paths between pairs.

We remark that one can find different notions of "small-worldness" in the existing literature. For instance, "small-world" refers to "short chain" in (Kleinberg, 2000; Milgram, 1967), while it refers to both "high clustering" and "short chain" in (Watts and Strogatz, 1998). We adopt the latter definition in the current study.

B. Rewiring Model

Let us now define the WS model. Consider a ring lattice with n nodes, where each node is connected with its k nearest neighbors ($k/2$ on the left and $k/2$ on the right, k even for convenience). The rewiring process consists of two steps. First, we erase each currently connected edge with probability β, independently. Next, we reconnect each edge pair with probability $\beta\frac{k}{n-1}$, allowing multiplicity.[1] The observed symmetric adjacency matrix $A \in \{0,1\}^{n\times n}$ has the following structure under some unobserved permutation matrix $P_\pi \in \{0,1\}^{n\times n}$. For $1 \le i < j \le n$, the probability that

$$[P_\pi A P_\pi^T]_{ij} = 1$$

is given by

(i) $1-\beta(1-\beta\frac{k}{n-1})$, if $0 < |i-j| \le \frac{k}{2} \mod n-1-\frac{k}{2}$
(ii) $\beta\frac{k}{n-1}$ otherwise,

and the entries are independent of each other. Equivalently, we have for $1 \le i < j \le n$

$$A_{ij} = \kappa\left([P_\pi B P_\pi^T]_{ij}\right), \tag{1}$$

where $\kappa(\cdot)$ is the entry-wise i.i.d. Markov channel,

$$\kappa(0) \sim \mathsf{Bernoulli}\left(\beta\frac{k}{n-1}\right),$$
$$\kappa(1) \sim \mathsf{Bernoulli}\left(1-\beta(1-\beta\frac{k}{n-1})\right),$$

[1]The original rewiring process in Watts and Strogatz (1998) does not allow multiplicity; however, for the simplicity of technical analysis, we focus on reconnection allowing multiplicity. These two rewiring processes are asymptotically equivalent.

and $B \in \{0,1\}^{n\times n}$ indicates the support of the structural ring lattice

$$B_{ij} = \begin{cases} 1, & \text{if } 0 < |i-j| \le \frac{k}{2} \mod n-1-\frac{k}{2} \\ 0, & \text{otherwise} \end{cases}. \tag{2}$$

We denote by $\mathsf{WS}(n,k,\beta)$ the distribution of the random graph generated from the rewiring model, and denote by $\mathsf{ER}(n,\frac{k}{n-1})$ the Erdős-Rényi random graph distribution (with matching average degree k). Remark that if $\beta = 1$, the small-world graph $\mathsf{WS}(n,k,\beta)$ reduces to $\mathsf{ER}(n,\frac{k}{n-1})$, with no neighborhood structure. In contrast, if $\beta = 0$, the small-world graph $\mathsf{WS}(n,k,\beta)$ corresponds to the deterministic ring lattice, without random connections. We focus on the dependence of the gap $1-\beta = o(1)$ on n and k, such that distinguishing between $\mathsf{WS}(n,k,\beta)$ and $\mathsf{ER}(n,\frac{k}{n-1})$ or reconstructing the ring lattice structure is statistically and computationally possible.

C. Summary of Results

The main theoretical and algorithmic results are summarized in this section. We first introduce several regions in terms of (n,k,β), according to the difficulty of the problem instance, and then we present the results using the phase diagram in Figure 1. Except for the 'impossible region', we will introduce algorithms with distinct computational properties. The 'impossible region' is defined through a lower bound, while the other regions are classified according to upper bounds on performance of respective procedures.

Impossible region: $1-\beta \prec \sqrt{\frac{\log n}{n}} \vee \frac{\log n}{k}$. In this region, no multiple testing procedure (regardless of computational budget) can succeed in distinguishing, with vanishing error, among the class of models that includes all of $\mathsf{WS}(n,k,\beta)$ and $\mathsf{ER}(n,\frac{k}{n-1})$.

Hard region: $\sqrt{\frac{\log n}{n}} \vee \frac{\log n}{k} \preceq 1-\beta \prec \sqrt{\frac{1}{k}} \vee \frac{\sqrt{\log n}}{k}$. It is possible to distringuish between $\mathsf{WS}(n,k,\beta)$ and $\mathsf{ER}(n,\frac{k}{n-1})$ statistically with vanishing error; however the evaluation of the test statistic (5) requires exponential time complexity, to the best of our knowledge.

Easy region: $\sqrt{\frac{1}{k}} \vee \frac{\sqrt{\log n}}{k} \preceq 1-\beta \preceq \sqrt{\sqrt{\frac{\log n}{n}} \vee \frac{\log n}{k}}$. There exists an efficient spectral test that can distinguish between the small-world random graph $\mathsf{WS}(n,k,\beta)$ and the Erdős-Rényi graph $\mathsf{ER}(n,\frac{k}{n-1})$ in time nearly linear in the matrix size.

Reconstructable region: $\sqrt{\sqrt{\frac{\log n}{n}} \vee \frac{\log n}{k}} \prec 1-\beta \preceq 1$. In this region, not only is it possible to detect the existence of the lattice structure in a small-world graph, but it is also possible to consistently reconstruct the neighborhood structure via a novel computationally efficient correlation thresholding procedure.

The following phase diagram provides an intuitive illustration of the above theoretical results. If we parametrize $k \asymp n^x, 0 < x < 1$ and $1-\beta \asymp n^{-y}, 0 < y < 1$, each point $(x,y) \in [0,1]^2$ corresponds to a particular problem instance with parameter bundle $(n, k = n^x, \beta = 1-n^{-y})$. According to

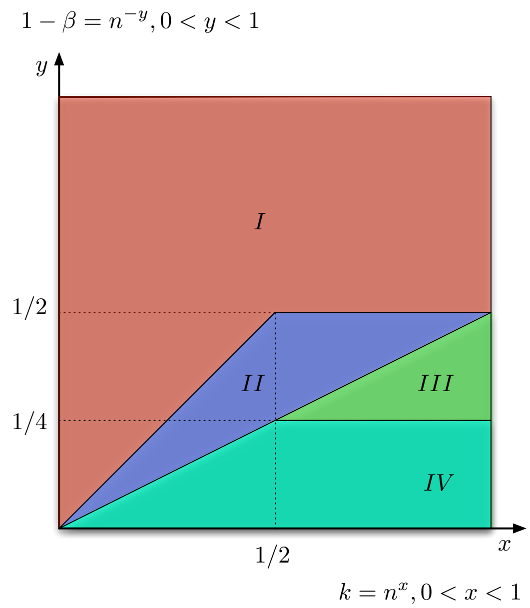

Fig. 1: Phase diagram for small-world network: impossible region (red region I), hard region (blue region II), easy region (green region III), and reconstructable region (cyan region IV).

the location of (x, y), the difficulty of the problem changes; for instance, the larger the x and the smaller the y is, the easier the problem becomes. The various regions are: impossible region (red region I), hard region (blue region II), easy region (green region III), reconstructable region (cyan region IV).

D. Notation

$A, B, Z \in \mathbb{R}^{n\times n}$ denote symmetric matrices: A is the adjacency matrix, B is the structural signal matrix as in Equation (2), and $Z = A - \mathbb{E}A$ is the noise matrix. We denote the matrix of all ones by J. Notations $\preceq, \succeq, \prec, \succ$ denote the asymptotic order: $a(n) \preceq b(n)$ if and only if $\limsup_{n\to\infty} \frac{a(n)}{b(n)} \le c$, with some constant $c > 0$, $a(n) \prec b(n)$ if and only if $\limsup_{n\to\infty} \frac{a(n)}{b(n)} = 0$. $C, C' > 0$ are universal constants that may change from line to line. For a symmetric matrix A, $\lambda_i(A)$, $1 \le i \le n$, denote the eigenvalues in a decreasing order. The inner-product $\langle A, B\rangle = \mathrm{tr}(A^T B)$ denotes both the Euclidian inner-product and matrix inner-product. For any integer n, $[n] := \{0, 1, \ldots, n-1\}$ denotes the index set. Denote the permutation in symmetric group $\pi \in S_n$ and its associated matrix form as P_π.

For a graph $G(V, E)$ generated from the Watts-Strogatz model $\mathsf{WS}(n, k, \beta)$ with associated permutation π, for each node $v_i \in V, 1 \le i \le |V|$, we denote

$$\mathcal{N}(v_i) := \left\{ v_j : 0 < |\pi^{-1}(i) - \pi^{-1}(j)| \le \frac{k}{2} \mod n - 1 - \frac{k}{2} \right\},$$

the ring neighborhood of v_i before permutation π is applied.

E. Organization of the Paper

The following sections are dedicated to the theoretical justification of the various regions in Section I-C. Specifically, Section II establishes the boundary for the impossible region I, where the detection problem is information-theoretically impossible. We contrast the hard region II with the regions III and IV in Section III; here, the difference arises in statistical and computational aspects of detecting the strong tie structure inside the random graph. Section IV studies a correlation thresholding algorithm that reconstructs the neighborhood structure consistently when the parameters lie within the reconstructable region IV. We also study a spectral ordering algorithm which succeeds in reconstruction in a part of region III. Whether the remaining part of region III admits a recovery procedure is an open problem. Additional further directions are listed in Section V.

II. The Impossible Region: Lower Bounds

We start with an information-theoretic result that describes the difficulty of distinguishing among a class of models. Theorem 1 below characterizes the impossible region, as in Section I-C, in the language of minimax multiple testing error. The proof is postponed to Section VI.

Theorem 1 (Impossible Region)**.** *Consider the following statistical models: $\mathcal{P}_0$ denotes the distribution of the Erdős-Rényi random graph $\mathsf{ER}(n, \frac{k}{n-1})$, and $\mathcal{P}_\pi, \pi \in S_{n-1}$ denote distributions of the Watts-Strogatz small-world graph $\mathsf{WS}(n, k, \beta)$ as in Equation (1) with different permutations π. Consider any selector $\phi : \{0,1\}^{n\times n} \to S_{n-1} \cup \{0\}$ that maps the adjacency matrix $A \in \{0,1\}^{n\times n}$ to a decision in $S_{n-1} \cup \{0\}$. Then for any fixed $0 < \alpha < 1/8$, the following lower bound on multiple testing error holds:*

$$\lim_{n\to\infty} \min_{\phi} \max \left\{ \mathcal{P}_0(\phi \ne 0), \frac{1}{(n-1)!} \sum_{\pi \in S_{n-1}} \mathcal{P}_\pi(\phi \ne \pi) \right\} \ge 1 - 2\alpha,$$

when the parameters satisfy

$$1 - \beta \le C_\alpha \cdot \sqrt{\frac{\log n}{n}} \quad \text{or} \quad 1 - \beta \le C'_\alpha \cdot \frac{\log n}{k} \cdot \frac{1}{\log \frac{n \log n}{k^2}},$$

with constants C_α, C'_α that only depend on α. In other words, if

$$1 - \beta \prec \sqrt{\frac{\log n}{n}} \vee \frac{\log n}{k},$$

no multiple testing procedure can succeed in distinguishing, with vanishing error, the class of models containing all of $\mathsf{WS}(n, k, \beta)$ and $\mathsf{ER}(n, \frac{k}{n-1})$.

The missing latent random variable, the permutation matrix P_π, is the object we are interested in recovering. A permutation matrix P_π induces a certain distribution on the adjacency matrix A. Thus the parameter space of interest, including models $\mathsf{WS}(n, k, \beta)$ and $\mathsf{ER}(n, \frac{k}{n-1})$, is of cardinality $(n-1)! + 1$. Based on the observed adjacency matrix, distinguishing among the models $\mathsf{WS}(n, k, \beta)$ and $\mathsf{ER}(n, \frac{k}{n-1})$ is equivalent to a multiple testing problem. The impossible region characterizes the information-theoretic difficulty of this reconstruction problem by establishing the condition that ensures non-vanishing minimax testing error as $n, k(n) \to \infty$.

The "high dimensional" nature of this problem is mainly driven by the unknown permutation matrix, and this latent structure introduces difficulty both statistically and computationally. Statistically, via Le Cam's method, one can build a distance metric on permutation matrices using the distance between the corresponding measures (measures on adjacency matrices induced by the permutation structure). In order to characterize the intrinsic difficulty of estimating the permutation structure, one needs to understand the richness of the set of permutation matrices within certain distance to one particular element, a combinatorial task. The combinatorial nature of the problem makes the "naive" approach computationally intensive.

III. Hard v.s. Easy Regions: Detection Statistics

This section studies the hard and easy regions in Section I-C. First, we propose a near optimal test, the **maximum likelihood test**, that detects the ring structure above the information boundary derived in Theorem 1. However, the evaluation of the maximum likelihood test requires $\mathcal{O}(n^n)$ time complexity. The maximum likelihood test succeeds outside of region I, and, in particular, succeeds (statistically) in the hard region II. We then propose another efficient test, the **spectral test**, that detects the ring structure in time $\mathcal{O}^*(n^2)$ via the power method. The method succeeds in regions III and IV.

Theorem 2 combines the results of Lemma 1 and Lemma 2 below.

Theorem 2 (Detection: Easy and Hard Boundaries)**.** *Consider the following statistical models: $\mathcal{P}_0$ denotes the distribution of the Erdős-Rényi random graph $\mathsf{ER}(n, \frac{k}{n-1})$, and $\mathcal{P}_\pi, \pi \in S_{n-1}$ denote distributions of the Watts-Strogatz small-world graph $\mathsf{WS}(n,k,\beta)$. Consider any selector $\phi : \{0,1\}^{n\times n} \to \{0,1\}$ that maps an adjacency matrix to a binary decision.*

We say that minimax detection for the small-world random model is possible when

$$\lim_{n\to\infty} \min_{\phi} \max \left\{ \mathcal{P}_0(\phi \neq 0), \frac{1}{(n-1)!} \sum_{\pi \in S_{n-1}} \mathcal{P}_\pi(\phi \neq 1) \right\} = 0. \tag{3}$$

If the parameter (n,k,β) satisfies

$$\text{hard boundary}: \quad 1-\beta \succeq \sqrt{\frac{\log n}{n}} \vee \frac{\log n}{k},$$

minimax detection is possible, and an exponential time **maximum likelihood test** *(5) ensures (3). If, in addition, the parameter (n,k,β) satisfies*

$$\text{easy boundary}: \quad 1-\beta \succeq \sqrt{\frac{1}{k}} \vee \frac{\sqrt{\log n}}{k},$$

then a near-linear time **spectral test** *(7) ensures (3).*

Proof of Theorem 2 consists of two parts, which will be addressed in the following two sections, respectively.

A. Maximum Likelihood Test

Consider the test statistic T_1 as the objective value of the following optimization

$$T_1(A) := \max_{P_\pi} \langle P_\pi B P_\pi^T, A \rangle, \tag{4}$$

where $P_\pi \in \{0,1\}^{n\times n}$ is taken over all permutation matrices and A is the observed adjacency matrix. The **maximum likelihood test** $\phi_1 : A \to \{0,1\}$ based on T_1 by

$$\phi_1(A) = \begin{cases} 1 & \text{if } T_1(A) \geq \frac{k}{n-1}nk + 2\sqrt{\frac{k}{n-1}nk \cdot \log n!} + \frac{2}{3}\cdot \log n! \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

The threshold is chosen as the rate $k^2 + \mathcal{O}\left(\sqrt{k^2 n \log \frac{n}{e}} \vee n \log \frac{n}{e}\right)$: if the objective value is of a greater order, then we believe the graph is generated from the small-world rewiring process with strong ties; otherwise we cannot reject the null, the random graph model with only weak ties.

Lemma 1 (Guarantee for Maximum Likelihood Test)**.** *The maximum likelihood test ϕ_1 in Equation (5) succeeds in detecting the small-world random structure when*

$$1-\beta \succeq \sqrt{\frac{\log n}{n}} \vee \frac{\log n}{k},$$

in the sense that

$$\lim_{n,k(n)\to\infty} \max \left\{ P_0(\phi_1 \neq 0),\ \frac{1}{(n-1)!} \sum_{i=1}^{(n-1)!} P_i(\phi_1 \neq 1) \right\} = 0.$$

Remark 1. Lemma 1 can be viewed as the condition on the signal and noise separation. By solving the combinatorial optimization problem, the test statistic aggregates the signal that separates from the noise the most. An interesting open problem is: if we solve a relaxed version of the combinatorial optimization problem (4) in polynomial time, how much stronger the condition on $1-\beta$ needs to be to ensure power.

B. Spectral Test

For the spectral test, we calculate the second largest eigenvalue of the adjacency matrix A as the test statistic

$$T_2(A) := \lambda_2(A). \tag{6}$$

The **spectral test** $\phi_2 : A \to \{0,1\}$ is

$$\phi_2(A) = \begin{cases} 1 & \text{if } T_2(A) \succeq \sqrt{k} \vee \sqrt{\log n} \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

Namely, if $\lambda_2(A)$ passes the threshold, we classify the graph as a small-world graph. Evaluation of (7) requires near-linear time $\mathcal{O}^*(n^2)$ in the size of the matrix.

Lemma 2 (Guarantee for Spectral Test)**.** *The second eigenvalue test ϕ_2 in Equation (7) satisfies*

$$\lim_{n,k(n)\to\infty} \max \left\{ P_0(\phi_2 \neq 0),\ \frac{1}{(n-1)!} \sum_{i=1}^{(n-1)!} P_i(\phi_2 \neq 1) \right\} = 0$$

whenever

$$1-\beta \succeq \sqrt{\frac{1}{k}} \vee \frac{\sqrt{\log n}}{k}.$$

The main idea behind Lemma 2 is as follows. Let us look at the expectation of the adjacency matrix,

$$\mathbb{E}A = (1-\beta)(1-\beta\frac{k}{n-1}) \cdot P_\pi^T B P_\pi + \beta\frac{k}{n-1}\cdot(J-I),$$

where J is the matrix of all ones. The main structure matrix $P_\pi^T B P_\pi$ is a permuted version of the *circulant matrix* (see e.g. (Gray, 2006)). The spectrum of the circulant matrix B is highly structured, and is of distinct nature in comparison to the noise matrix $A - \mathbb{E}A$.

IV. Reconstructable Region: Fast Structural Reconstruction

In this section, we discuss reconstruction of the ring structure in the Watts-Strogatz model. We show that in the reconstructable region (region IV in Figure 1), a **correlation thresholding procedure** succeeds in reconstructing the ring neighborhood structure. As a by-product, once the neighborhood structure is known, one can distinguish between weak ties (random edges) and strong ties (neighborhood edges) for each node. A natural question is whether there is another algorithm that can work in a region (beyond region IV) where correlation thresholding fails. We show that in a certain regime with large k, a **spectral ordering procedure** outperforms the correlation thresholding procedure and succeeds in parts of regions III and IV (as depicted in Figure 2 below).

A. Correlation Thresholding

Consider the following **correlation thresholding procedure** for neighborhood reconstruction.

Algorithm 1: Correlation Thresholding for Neighborhood Reconstruction

---

Data: An adjacency matrix $A \in \mathbb{R}^{n\times n}$ for the graph $G(V, E)$.
Result: For each node $v_i, 1 \le i \le n$, an estimated set for neighborhood $\hat{\mathcal{N}}(v_i)$.
1. For each v_i, calculate correlations $\langle A_i, A_j\rangle$, $j \neq i$;
2. Sort $\{\langle A_i, A_j\rangle, j \in [n]\backslash\{i\}\}$ in a decreasing order, select the largest k ones to form the estimated set $\hat{\mathcal{N}}(v_i)$;
Output: $\hat{\mathcal{N}}(v_i)$, for all $i \in [n]$

The following lemma proves consistency of the above Algorithm 1. Note the computational complexity is $\mathcal{O}(n \cdot \min\{\log n, k\})$ for each node using quick-sort, with a total runtime $\mathcal{O}^*(n^2)$.

Lemma 3 (Consistency of Correlation Thresholding)**.** *Consider the Watts-Strogatz random graph* $\mathsf{WS}(n, k, \beta)$*. Under the reconstructable regime IV (in Figure 1), that is,*

$$1 - \beta \succ \sqrt{\frac{\log n}{k}} \vee \left(\frac{\log n}{n}\right)^{1/4}, \tag{8}$$

correlation thresholding provides a consistent estimate of the neighborhood set $\mathcal{N}(v_i)$ *w.h.p in the sense that*

$$\lim_{n,k(n)\to\infty} \max_{i\in[n]} \frac{|\hat{\mathcal{N}}(v_i) \triangle \mathcal{N}(v_i)|}{|\mathcal{N}(v_i)|} = 0,$$

where $\triangle$ *denotes the symmetric set difference.*

The condition under which consistency of correlation thresholding is ensured corresponds to the reconstructable region in Figure 1. One may ask if there is another algorithm that can provide a consistent estimate of the neighborhood set beyond region IV. The answer is yes, and we will show in the following section that under the regime when k is large (for instance, $k \succeq n^{\frac{15}{16}}$), indeed it is possible to slightly improve on Algorithm 1.

B. Spectral Ordering

Consider the following **spectral ordering procedure**, which approximately reconstructs the ring lattice structure when k is large, i.e., $k \succ n^{\frac{7}{8}}$.

Algorithm 2: Spectral Reconstruction of Ring Structure

Data: An adjacency matrix $A \in \mathbb{R}^{n\times n}$ for the graph $G(V, E)$.
Result: A ring embedding of the nodes V.
1. Calculate top 3 eigenvectors in the SVD $A = U\Sigma U^T$. Denote second and third eigenvectors as $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^n$, respectively;
2. For each node i and vector $A_{\cdot i} \in \mathbb{R}^n$, calculate the associated angle θ_i for the vector $(u^T A_{\cdot i}, v^T A_{\cdot i})$;
Output: the sorted sequence $\{\theta_i\}_{i=1}^n$ and the corresponding ring embedding of the nodes. For each node v_i, $\hat{\mathcal{N}}(v_i)$ are the closest k nodes in the ring embedding.

The following Lemma 4 shows that when k is large, Algorithm 2 also provides consistent reconstruction of the ring lattice. Its computational complexity is $\mathcal{O}^*(n^2)$.

Lemma 4 (Guarantee for Spectral Ordering)**.** *Consider the Watts-Strogatz graph* $\mathsf{WS}(n, k, \beta)$*. Assume k is large enough in the following sense:*

$$1 > \overline{\lim_{n,k(n)\to\infty}} \frac{\log k}{\log n} \ge \underline{\lim_{n,k(n)\to\infty}} \frac{\log k}{\log n} > \frac{7}{8}.$$

Under the regime

$$1 - \beta \succ \frac{n^{3.5}}{k^4}, \tag{9}$$

the spectral ordering provides consistent estimate of the neighborhood set $\mathcal{N}(v_i)$ *w.h.p. in the sense that*

$$\lim_{n,k(n)\to\infty} \max_{i\in[n]} \frac{|\hat{\mathcal{N}}(v_i) \triangle \mathcal{N}(v_i)|}{|\mathcal{N}(v_i)|} = 0,$$

where $\triangle$ *denotes the symmetric set difference.*

In Lemma 4, we can only prove consistency of spectral ordering under the technical condition that k is large. We do not believe this is due to an artifact of the proof. Even though the structural matrix (the signal) has large eigenvalues, the eigen-gap is not large enough. The spectral ordering succeeds when the spectral gap stands out over the noise level, which implies that k needs to be large enough.

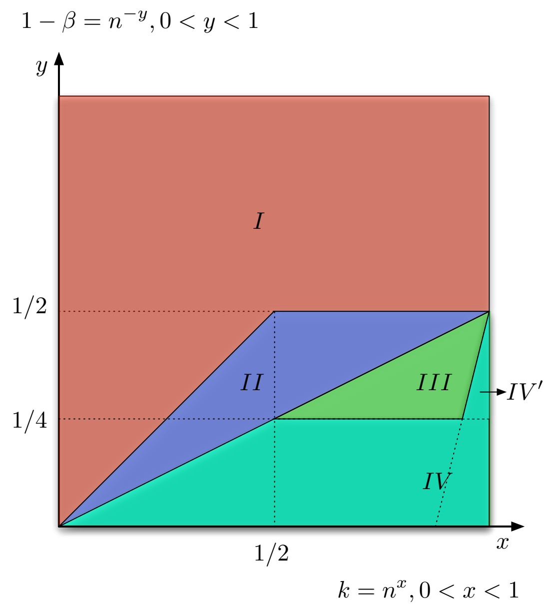

Fig. 2: Phase diagram for small-world networks: impossible region (red region I), hard region (blue region II), easy region (green region III), and reconstructable region (cyan region IV and IV'). Compared to Figure 1, the spectral ordering procedure extends the reconstructable region (IV) when $k \succ n^{\frac{15}{16}}$ (IV').

Let us compare the region described in Equation (9) with the reconstructable region in Equation (8). We observe that spectral ordering pushes slightly beyond the reconstructable region when $k \succ n^{\frac{15}{16}}$, as shown in Figure 2.

C. Numerical Study

To see how the ring embedding Algorithm 2 performs on real dataset, we implemented it in Python, on the co-appearance network of characters in the novel Les Misérable[2] (Knuth, 1993). Figure 3 summarizes the visualization (zoom in for better resolution). Each node represents one character, and the color and size illustrate its degree, with darker color and larger size meaning higher degree. The lines connecting nodes on the ring represent co-appearance relationship in the chapters of the book, with the line width summarizing the co-appearance intensity. As one can see in the embedding, the obvious triangle is among the three main characters – Valjean, Marius and Cosette. In the ring embedding, Valjean and Javert are next to each other, so does Marius and Eponine, as they have very strong ties (enemies and friends) in the plot. The algorithm embeds the main characters – Valjean, Marius, Fantine, Thenadler, etc – in a rather spread out fashion on the ring, with each main character communicating with many other minor characters as in the novel. The structure assures the "short chain" property – any two characters can reach each other through these few main characters as middle points. One can also see many triadic closures in the ring neighborhood around main character, supporting the local "high clustering" feature.

[2] The data is downloaded from Prof. Mark Newman's website http://www-personal.umich.edu/~mejn/netdata/.

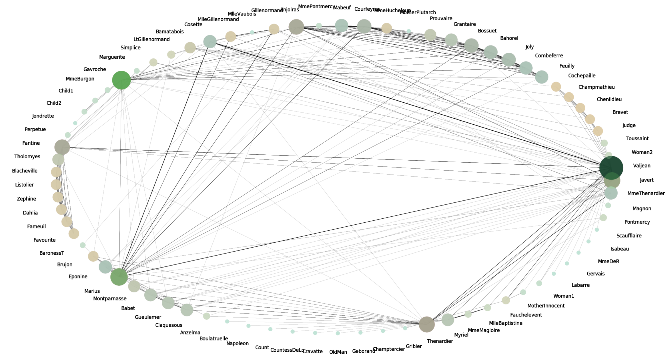

Fig. 3: Ring embedding of Les Misérable co-appearance network.

V. Discussion

a) Comparison to stochastic block models: Recently, stochastic block models (SBM) have attracted considerable amount of attention from researchers in various fields (Decelle et al., 2011; Massoulié, 2014; Mossel et al., 2013). Community detection in stochastic block models focuses on recovering the hidden community structure obscured by noise in the adjacency matrix and further concealed by the latent permutation on the nodes.

Detectability or weak recovery of the hidden community is one of the central question in studying SBM in the constant degree regime. Drawing insights from statistical physics, Decelle et al. (2011) conjectured a sharp transition threshold (also known as the Kesten-Stigum bound) for detection in the symmetric two-community case, above which recovering the community better than random guessing is possible, and below which – impossible. Massoulié (2014); Mossel et al. (2013) proved the conjecture independently, one using spectral analysis on the non-backtracking matrix (Hashimoto, 1989), the other through analyzing non-backtracking walks. Later, for partial recovery and strong recovery (reconstruction) of multiple communities beyond the symmetric case, Abbe and Sandon (2015) characterized the recovery threshold in terms of the Chernoff-Hellinger divergence.

The hidden community structure for classic SBM is illustrated in Figure 4 (left) as a block diagonal matrix. An interesting but theoretically more challenging extension to the classic SBM is the mixed-membership SBM, where each node may simultaneously belong to several communities. Consider an easy case of the model, where the mixed-membership occurs only within neighborhood communities, as shown in the middle image of Figure 4. The small-world network we are investigating in this paper can be seen as an extreme case (shown on the right-most figure) of this easy mixed-membership SBM, where each node falls in effectively k local clusters. In the small-world network model, identifying the structural links and random links becomes challenging since

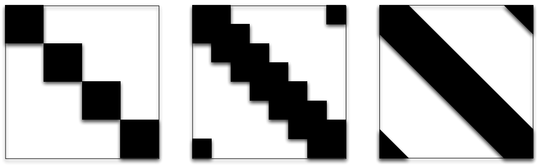

Fig. 4: The structural matrices for stochastic block model (left), mixed membership SBM (middle), and small-world model (right). The black location denotes the support of the structural matrix.

there are many local clusters, in constrast to the relatively small number of communities in SBM. The multitude of local clusters makes it difficult to analyze the effect of the hidden permutation on the structural matrix. We view the current paper as an initial attempt at tackling this problem.

b) Relations to graph matching: Small world reconstruction, community membership reconstruction, planted clique localization etc., can be cast as solving for the latent permutation matrix P_π with different structural matrix B, in $\arg\max_{P_\pi \in \Pi} \langle P_\pi A P_\pi^T, B\rangle$ as suggested in Eq. (4). This is also called graph matching (Lyzinski et al., 2014). As one aims to match the observed adjacency matrix A to the structural matrix B via latent permutation matrix P_π. As written in the above form, the reconstruction task is reduced to a quadratic assignment problem (QAP), which is known to be NP-hard (Burkard et al., 1998; Cela, 2013). Due to the NP-hard nature of QAP, various relaxations on the permutation matrix constraints have been proposed: for instance, orthogonal matrices, doubly stochastic matrices, and SDP relaxations (Chandrasekaran et al., 2012). Quantifying the loss due to a relaxation for each model is a challenging task.

c) Reconstructable region: We addressed the reconstruction problem via two distinct procedures: correlation thresholding and spectral ordering. Whether there exist other computationally efficient algorithms that can significantly improve upon the current reconstructable region is an open problem. Designing new algorithms requires a deeper insight into the structure of the small-world model, and will probably shed light on better algorithms for mixed membership models.

VI. Technical Proofs

Proof of Theorem 1. Denote the circulant matrix by B (it is B_π for any $\pi \in S_{n-1}$). The log-likelihood for WS model on symmetric matrix X (with diagonal elements being 0) is

$$\begin{aligned}\log \mathcal{L}_{n,k,\beta}(X|B) &= \log \frac{1-\beta(1-\beta\frac{k}{n-1})}{\beta(1-\beta\frac{k}{n-1})}\cdot\langle X, B\rangle \\ &+ \log\frac{\beta\frac{k}{n-1}}{1-\beta\frac{k}{n-1}}\cdot\langle X, J-I-B\rangle \\ &+ nk\log(\beta(1-\beta\frac{k}{n-1})) + n(n-1-k)\log(1-\beta\frac{k}{n-1})\end{aligned}$$

For the Erdős-Rényi model, the log likelihood is

$$\log \mathcal{L}_{n,k}(X) = \log\frac{\frac{k}{n-1}}{1-\frac{k}{n-1}}\cdot\langle X, J-I\rangle + n(n-1)\log(1-\frac{k}{n-1}).$$

The Kullback-Leibler divergence between these two models is

$$\begin{aligned}&\mathrm{KL}(P_B||P_0) = \mathbb{E}_{X\sim P_B}\log\frac{P_B(X)}{P_0(X)} \\ &= \mathbb{E}_{X\sim P_B}\Bigg\{-\left(\log\frac{\frac{k}{n-1}}{1-\frac{k}{n-1}} - \log\frac{\beta\frac{k}{n-1}}{1-\beta\frac{k}{n-1}}\right)\cdot\langle X, J-I\rangle \\ &\quad -n(n-1)\log(1-\frac{k}{n-1}) \\ &\quad +\left(\log\frac{1-\beta(1-\beta\frac{k}{n-1})}{\beta(1-\beta\frac{k}{n-1})} - \log\frac{\beta\frac{k}{n-1}}{1-\beta\frac{k}{n-1}}\right)\cdot\langle X, B\rangle \\ &\quad +nk\log(\beta(1-\beta\frac{k}{n-1})) + n(n-1-k)\log(1-\beta\frac{k}{n-1})\Bigg\}\end{aligned}$$

which is equal to

$$\begin{aligned}& -\left(\log\frac{\frac{k}{n-1}}{1-\frac{k}{n-1}} - \log\frac{\beta\frac{k}{n-1}}{1-\beta\frac{k}{n-1}}\right)\cdot \\ &\qquad \left\langle (1-\beta)(1-\beta\frac{k}{n-1})B + \beta\frac{k}{n-1}(J-I), J-I\right\rangle \\ &+\left(\log\frac{1-\beta(1-\beta\frac{k}{n-1})}{\beta(1-\beta\frac{k}{n-1})} - \log\frac{\beta\frac{k}{n-1}}{1-\beta\frac{k}{n-1}}\right)\cdot \\ &\qquad \left\langle (1-\beta)(1-\beta\frac{k}{n-1})B + \beta\frac{k}{n-1}(J-I), B\right\rangle \\ &- n(n-1)\log(1-\frac{k}{n-1}) + nk\log(\beta(1-\beta\frac{k}{n-1})) \\ &+ n(n-1-k)\log(1-\beta\frac{k}{n-1}) \\ &= n(n-1)\log\frac{1-\beta\frac{k}{n-1}}{1-\frac{k}{n-1}} - nk\log\frac{1}{\beta} \qquad (10) \\ &- \left[\log\frac{1}{\beta} + \log\frac{1-\beta\frac{k}{n-1}}{1-\frac{k}{n-1}}\right] nk\left[1-(1-\beta)\beta\frac{k}{n-1}\right] \\ &+ \left[\log\frac{1}{\beta} + \log\frac{1-\beta(1-\beta\frac{k}{n-1})}{\beta\frac{k}{n-1}}\right] nk\left[1-\beta(1-\beta\frac{k}{n-1})\right] \\ &= -\log\frac{1}{\beta}\cdot nk\left[1+\beta-\beta\frac{k}{n-1}\right] \\ &+ \log\frac{1-\beta\frac{k}{n-1}}{1-\frac{k}{n-1}}\cdot n\left[(n-1-k)+(1-\beta)\beta\frac{k^2}{n-1}\right] \\ &+ \log\frac{1-\beta(1-\beta\frac{k}{n-1})}{\beta\frac{k}{n-1}}\cdot nk\left[1-\beta(1-\beta\frac{k}{n-1})\right]. \qquad (11)\end{aligned}$$

Via the inequality $\log(1+x) < x$ for all $x > -1$, we can further simplify the above expression as

$$\begin{aligned}&\mathrm{KL}(P_B||P_0) \\ &\le nk(1-\beta)\left[-\beta+\beta\frac{k}{n-1} + (1-\beta)\beta\frac{k^2}{n(n-1-k)}\right] \\ &\quad + \frac{(1-\beta)(1-\beta\frac{k}{n-1})}{\beta\frac{k}{n-1}} nk\left[(1-\beta)+\beta^2\frac{k}{n-1}\right] \\ &\le nk(1-\beta)\left[(1-\beta)\beta\frac{k}{n-1} + (1-\beta)\beta\frac{k^2}{n(n-1-k)}\right] \\ &\quad + \frac{(1-\beta)^2(1-\beta\frac{k}{n-1})}{\beta} n(n-1) \le C\cdot n^2(1-\beta)^2, \quad (12)\end{aligned}$$

where $0 < C < \frac{1}{2}\frac{k^2}{n(n-1)} + \frac{1}{\beta}$ is some universal constant (note we are interested in the case when β is close to 1).

When $k \preceq n^{1/2}$, the above bound can be further strengthened, in the following sense (recall equation (11)):

$$\begin{aligned}
&\mathrm{KL}(P_B||P_0)\\
&\leq nk(1-\beta)\left[-\beta+\beta\frac{k}{n-1}+(1-\beta)\beta\frac{k^2}{n(n-1-k)}\right]\\
&+\log\frac{1-\beta(1-\beta\frac{k}{n-1})}{\beta\frac{k}{n-1}}\cdot nk\left[1-\beta(1-\beta\frac{k}{n-1})\right]\\
&\leq\left\{\log\frac{1-\beta(1-\beta\frac{k}{n-1})}{\beta\frac{k}{n-1}}\cdot\frac{1-\beta(1-\beta\frac{k}{n-1})}{\beta\frac{k}{n-1}}\right\}\cdot k^2\beta\frac{n}{n-1}.
\end{aligned}$$

Denote $t:=\frac{1-\beta(1-\beta\frac{k}{n-1})}{\beta\frac{k}{n-1}}=\frac{1-\beta}{\beta}\frac{n-1}{k}+\beta$. Thus we have

$$\mathrm{KL}(P_B||P_0)\leq t\log t\cdot k^2\beta\frac{n}{n-1}. \tag{13}$$

Suppose for some constant $\alpha_*>0$, and $\alpha=\alpha_*\cdot\frac{1}{\beta}(1-\frac{1}{n})^2$, we have the following

$$t\leq\alpha\frac{n\log\frac{n}{e}}{k^2}\cdot\frac{1}{\log\alpha\frac{n\log\frac{n}{e}}{k^2}} \tag{14}$$

and
$$t\log t\leq\alpha\frac{n\log\frac{n}{e}}{k^2}\cdot\left(1-\frac{\log\log\alpha\frac{n\log\frac{n}{e}}{k^2}}{\log\alpha\frac{n\log n}{k^2}}\right)<\alpha\frac{n\log\frac{n}{e}}{k^2}. \tag{15}$$

Plugging in the expression for t into (14), if

$$\begin{aligned}
\frac{1-\beta}{\beta}&\leq\alpha(1+\frac{1}{n-1})\cdot\frac{\log\frac{n}{e}}{k}\cdot\frac{1}{\log\alpha\frac{n\log\frac{n}{e}}{k^2}}-\frac{k}{n-1} \qquad (16)\\
&\asymp\frac{\log n}{k}\frac{1}{\log\frac{n\log\frac{n}{e}}{k^2}}
\end{aligned}$$

we have

$$t\leq\alpha\frac{n\log\frac{n}{e}}{k^2}\cdot\frac{1}{\log\alpha\frac{n\log\frac{n}{e}}{k^2}}\quad\Rightarrow\quad t\log t<\alpha\frac{n\log\frac{n}{e}}{k^2}$$

which further implies, via Equation (12),

$$\begin{aligned}
\frac{1}{(n-1)!}\sum_{\pi\in S_{n-1}}\mathrm{KL}(P_{B_\pi}||P_0)&\leq t\log t\cdot k^2\beta\frac{n}{n-1}\\
&\leq\alpha_*\cdot\log(n-1)!.
\end{aligned}$$

Recalling the bound on KL-divergence, if

$$1-\beta\leq\sqrt{\frac{\alpha_*}{C}\cdot\frac{(n-1)\log\frac{n}{e}}{n^2}}\asymp\sqrt{\frac{\log n}{n}} \tag{17}$$

we have

$$\frac{1}{(n-1)!}\sum_{\pi\in S_{n-1}}\mathrm{KL}(P_{B_\pi}||P_0)\leq n^2(1-\beta)^2\leq\alpha_*\cdot\log(n-1)!.$$

We invoke the following Lemma on minimax error through Kullbak-Leibler divergence.

Lemma 5 (Tsybakov (2009), Proposition 2.3)**.** *Let P_0, P_1,..., P_M be probability measures on $(\mathcal{X},\mathcal{A})$ satisfying*

$$\frac{1}{M}\sum_{j=1}^{M}\mathrm{KL}(P_j||P_0)\leq\alpha\cdot\log M \tag{18}$$

with $0<\alpha<\frac{1}{8}$. Then for any $\psi:\mathcal{X}\to[M+1]$

$$\begin{aligned}
&\max\left\{P_0(\psi\neq0),\frac{1}{M}\sum_{j=1}^{M}P_j(\psi\neq j)\right\}\\
&\qquad\geq\frac{\sqrt{M}}{\sqrt{M}+1}\left(1-2\alpha-\sqrt{\frac{2\alpha}{\log M}}\right).
\end{aligned}$$

Hence, if either one of the conditions in Equations (16) and (17) holds, we have

$$\frac{1}{(n-1)!}\sum_{\pi\in S_{n-1}}\mathrm{KL}(P_{B_\pi}||P_0)\leq\alpha_*\cdot\log(n-1)!. \tag{19}$$

Putting everything together, Equation (19) holds whenever if

$$1-\beta\prec\sqrt{\frac{\log n}{n}}\vee\frac{\log n}{k}.$$

Applying Lemma 5, we complete the proof:

$$\begin{aligned}
&\lim_{n\to\infty}\ \min_{\phi}\ \max\left\{P_0(\phi\neq0),\ \frac{1}{(n-1)!}\sum_{i=1}^{(n-1)!}P_i(\phi\neq i)\right\}\\
&\geq\lim_{n\to\infty}\frac{\sqrt{(n-1)!}}{1+\sqrt{(n-1)!}}\left(1-2\alpha-\sqrt{\frac{2\alpha}{\log(n-1)!}}\right)=1-2\alpha.
\end{aligned}$$

□

Proof of Lemma 1. Let us state the well-known Bernstein's inequality (Boucheron et al. (2013), Theorem 2.10), which will be used in the proof of this lemma.

Lemma 6 (Bernstein's inequality)**.** *Let $X_1,\ldots,X_n$ be independent bounded real-valued random variables. Assume that there exist positive numbers v and c such that*

$$\begin{aligned}
&\sum_{i=1}^{n}\mathbb{E}[X_i^2]\leq v,\\
&X_i\leq3c,\forall1\leq i\leq n\quad a.s.
\end{aligned}$$

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

$$\mathbb{P}\left(\sum_{i=1}^{n}(X_i-\mathbb{E}X_i)\geq\sqrt{2vt}+ct\right)\leq e^{-t}. \tag{20}$$

First, let us consider the case when the adjacency matrix A is generated from the Erdős-Rényi random graph $\mathsf{ER}(n,\frac{k}{n-1})$. For any P_π with $\pi\in S_{n-1}$, we know $\langle P_\pi BP_\pi^T,A\rangle$ has the same distribution as $\langle B,A\rangle$. Thus, in view of Lemma 6,

$$\begin{aligned}
\langle P_\pi BP_\pi^T,A\rangle&\overset{\text{in law}}{=\!=}\langle B,A\rangle=2\sum_{i>j}A_{ij}B_{ij}\\
&=2\sum_{i>j}\mathbb{E}[A_{ij}]B_{ij}+2\sum_{i>j}(A_{ij}-\mathbb{E}[A_{ij}])B_{ij}\\
&\leq\frac{k}{n-1}nk+2\sqrt{\frac{k}{n-1}nkt}+\frac{2}{3}t
\end{aligned}$$

with probability at least $1-\exp(-t)$. Indeed, there are $nk/2$ non-zero $B_{i,j},i>j$, and it is clear that $A_{ij}\sim\mathsf{Bernoulli}(\frac{k}{n-1})$

and $2\sum_{i>j}\mathbb{E}[A_{ij}]B_{ij} = nk\frac{k}{n-1}$, implying the choice of $c = \frac{1}{3}$ and

$$v = \sum_{i<j}\mathbb{E}[(A_{ij}B_{ij})^2] = \sum_{i<j}\mathbb{E}[A_{ij}^2]B_{ij} = \frac{nk}{2}\frac{k}{n-1}$$

in Lemma 6. Via the union bound, taking $t = \log n!$, we have

$$\begin{aligned}&\max_{P_\pi}\ \langle P_\pi B P_\pi^T, A\rangle\\ &\leq \frac{k}{n-1}nk + 2\sqrt{\frac{k}{n-1}nk\cdot\log n!} + \frac{2}{3}\cdot\log n!\end{aligned}$$

with probability at least $1-(n-1)!\exp(-\log n!) = 1-\frac{1}{n}$.

Alternatively, suppose A is from the small-world rewiring model $\mathsf{WS}(n,k,\beta)$, with permutation being identity. With probability at least $1-\exp(-\log n) = 1-\frac{1}{n}$,

$$\begin{aligned}\max_{P_\pi}\ \langle P_\pi B P_\pi^T, A\rangle &\geq \langle B, A\rangle\\ &= \langle B, \mathbb{E}[A]\rangle + \langle B, A-\mathbb{E}[A]\rangle\\ &\geq (1-\beta+\beta^2\frac{k}{n-1})nk - \sqrt{nk\cdot\log n}\end{aligned}$$

where the last step follows via Hoeffding's inequality: it is clear that for (i,j) with $B_{ij}\neq 0$,

$$\mathbb{E}[A_{ij}] = 1-\beta+\beta^2\frac{k}{n-1},$$

and $0\leq A_{ij}\leq 1$ almost surely.

Therefore if there exist a threshold $T>0$ such that

$$(1-\beta+\beta^2\frac{k}{n-1})nk - \sqrt{nk\cdot\log n} > T$$

$$\text{and } T > \frac{k}{n-1}nk + 2\sqrt{\frac{k}{n-1}nk\cdot\log n!} + \frac{2}{3}\cdot\log n! \quad (21)$$

we have that

$$\begin{aligned}&\lim_{n,k(n)\to\infty}\ \max\ \left\{P_0(\phi_1\neq 0),\ \frac{1}{(n-1)!}\sum_{i=1}^{(n-1)!}P_i(\phi_1\neq 1)\right\}\\ &\leq \lim_{n,k(n)\to\infty}\ \frac{1}{n} = 0.\end{aligned}$$

The detailed calculation of Equation (21) yields that the test succeeds with high probability whenever

$$1-\beta \succeq \sqrt{\frac{\log n}{n}} \vee \frac{\log n}{k}.$$

□

Proof of Lemma 2. Under the model $\mathsf{WS}(n,k,\beta)$ with permutation P_π,

$$A = (1-\beta)(1-\beta\frac{k}{n-1})\cdot P_\pi^T B P_\pi + \beta\frac{k}{n-1}\cdot(J-I) + Z$$

where $J = 11^T\in\mathbb{R}^{n\times n}$, B is the ring structured signal matrix defined in Equation (2), and Z is a zero-mean noise random matrix.

We first study the random fluctuation part, $Z = A-\mathbb{E}A$. Let us bound the expectation $\mathbb{E}\|A-\mathbb{E}A\|$ as the first step, for any adjacency matrix $A\in\mathbb{R}^{n\times n}$ using the symmetrization trick. Denote $A'\sim A$ as the independent copy of A sharing the same distribution. Take $E,G\in\mathbb{R}^{n\times n}$ as random symmetric Rademacher and Gaussian matrices with entries E_{ij}, G_{ij} being, respectively, independent Rademacher and Gaussian. Denoting matrix Hadamard product as $A\circ B$, we have

$$\begin{aligned}&\mathbb{E}\|A-\mathbb{E}A\| = \mathbb{E}\sup_{\|v\|_{\ell_2}=1}\langle(A-\mathbb{E}A)v,v\rangle\\ &= \mathbb{E}\sup_{\|v\|_{\ell_2}=1}\langle(A-\mathbb{E}_{A'}A')v,v\rangle \leq \mathbb{E}_A\mathbb{E}_{A'}\sup_{\|v\|_{\ell_2}=1}\langle(A-A')v,v\rangle\\ &= \mathbb{E}_E\mathbb{E}_A\mathbb{E}_{A'}\sup_{\|v\|_{\ell_2}=1}\langle[E\circ(A-A')]v,v\rangle\\ &\leq \mathbb{E}_A\mathbb{E}_E\sup_{\|v\|_{\ell_2}=1}\langle[E\circ A]v,v\rangle + \mathbb{E}_{A'}\mathbb{E}_E\sup_{\|v\|_{\ell_2}=1}\langle[-E\circ A']v,v\rangle\\ &= 2\mathbb{E}_A\mathbb{E}_E\sup_{\|v\|_{\ell_2}=1}\langle[E\circ A]v,v\rangle\\ &\leq \frac{2}{\sqrt{2/\pi}}\cdot\mathbb{E}_A\mathbb{E}_E\sup_{\|v\|_{\ell_2}=1}\langle[\mathbb{E}_G[|G|]\circ E\circ A]v,v\rangle\\ &\leq \sqrt{\frac{\pi}{2}}\cdot\mathbb{E}_A\mathbb{E}_E\mathbb{E}_G\sup_{\|v\|_{\ell_2}=1}\langle[|G|\circ E\circ A]v,v\rangle\\ &= \sqrt{\frac{\pi}{2}}\cdot\mathbb{E}_A\mathbb{E}_G\sup_{\|v\|_{\ell_2}=1}\langle[G\circ A]v,v\rangle\\ &= \sqrt{\frac{\pi}{2}}\cdot\mathbb{E}_A\left(\mathbb{E}_G\|G\circ A\|\right).\end{aligned}$$

Recall the following Lemma from (Bandeira and van Handel, 2014).

Lemma 7 (Bandeira and van Handel (2014), Theorem 1.1)**.** *Let X be the $n\times n$ symmetric random matrix with $X = G\circ A$, where $G_{ij}, i<j$ are i.i.d. $N(0,1)$ and A_{ij} are given scalars. Then*

$$\mathbb{E}_G\|X\| \precsim \max_i\sqrt{\sum_j A_{ij}^2} + \max_{ij}|A_{ij}|\cdot\sqrt{\log n}.$$

Thus via Jensen's inequality and the above Lemma, we upper bound

$$\begin{aligned}&\mathbb{E}\|A-\mathbb{E}A\| \leq \sqrt{\frac{\pi}{2}}\cdot\mathbb{E}_A\left(\mathbb{E}_G\|G\circ A\|\right)\\ &\precsim \mathbb{E}_A\left[\max_i\sqrt{\sum_j A_{ij}^2} + \max_{ij}|A_{ij}|\cdot\sqrt{\log n}\right]\\ &\leq \sqrt{\mathbb{E}_A\max_i\sum_j A_{ij}^2} + \sqrt{\log n}\\ &\leq \sqrt{k + C_1 2\sqrt{k\log n} + C_2\log n} + \sqrt{\log n} \asymp \sqrt{k}\vee\sqrt{\log n},\end{aligned}$$

where the last step uses Bernstein inequality Lemma 6. Moving from expectation $\mathbb{E}\|A-\mathbb{E}A\|$ to concentration on $\|A-\mathbb{E}A\|$ is through Talagrand's concentration inequality (see, Talagrand (1996) and Tao (2012) Theorem 2.1.13), since $\|\cdot\|$ is $1-$Lipschitz convex function in our case (and the entries are bounded), thus with probability at least $1-\frac{1}{n}$,

$$\|A-\mathbb{E}A\| \leq \mathbb{E}\|A-\mathbb{E}A\| + C\cdot\sqrt{\log n} \asymp \sqrt{k}\vee\sqrt{\log n}.$$

Now let us study the structural signal part. Matrix B is of the form circulant matrix, the associated polynomial is

$$f(x) = (x + x^{n-k/2})\cdot\frac{x^{k/2}-1}{x-1}.$$

The eigenvectors can be analytically calculated: collect for all $j = 0, 1, ..., n/2$

$$(\cos 0, \cos\frac{2\pi j}{n}, \cos\frac{2\pi 2j}{n}, \ldots, \cos\frac{2\pi nj}{n})$$

and

$$(\sin 0, \cos\frac{2\pi j}{n}, \sin\frac{2\pi 2j}{n}, \ldots, \sin\frac{2\pi nj}{n})$$

and the corresponding eigenvalues are

$$\lambda_j = f(w_j) = 2\sum_{i=1}^{k/2}\cos\left(i\frac{2\pi j}{n}\right).$$

Let us first assume $\frac{k}{n} \le \frac{1}{2}$, thus the second largest eigenvalue

$$\lambda_2 = 2\sum_{i=1}^{k/2}\cos\left(i\frac{2\pi}{n}\right) = \frac{2\sin\frac{k\pi}{2n}}{\sin\frac{\pi}{n}}\cos\frac{(k+2)\pi}{2n} \asymp k.$$

Now if there exist a threshold $T > 0$ such that w.h.p., the second eigenvalue of the adjacency matrix generated from WS model A_{WS} separates from that of the adjacency matrix generated from ER model A_{ER} in the following sense

$$\lambda_2(A_{\mathsf{WS}}) > T > \lambda_2(A_{\mathsf{ER}}),$$

we have

$$\lim_{n,k(n)\to\infty} \max\left\{P_0(\phi_2 \neq 0),\ \frac{1}{(n-1)!}\sum_{i=1}^{(n-1)!} P_i(\phi_2 \neq 1)\right\} = 0.$$

Using Weyl's interlacing inequality, we have

$$\begin{aligned}\lambda_2(A_{\mathsf{WS}}) &\ge \lambda_2(\mathbb{E}[A_{\mathsf{WS}}]) - \|Z\| \\ &\ge (1-\beta)(1-\beta\frac{k}{n-1})\lambda_2 - \sqrt{k} \vee \sqrt{\log n},\end{aligned}$$

while

$$\lambda_2(A_{\mathsf{ER}}) \le \sqrt{k} \vee \sqrt{\log n}.$$

Therefore, we have the condition for which the second eigenvalue test succeeds:

$$(1-\beta)(1-\beta\frac{k}{n-1})\lambda_2 > \sqrt{k} \vee \sqrt{\log n}$$

which means

$$(1-\beta)(1-\beta\frac{k}{n-1}) > \frac{\sqrt{k} \vee \sqrt{\log n}}{\frac{2\sin\frac{k\pi}{2n}}{\sin\frac{\pi}{n}}\cos\frac{(k+2)\pi}{2n}} \asymp \sqrt{\frac{1}{k}} \vee \frac{\sqrt{\log n}}{k}.$$

□

Proof of Lemma 3. Take any two rows $A_{i\cdot}, A_{j\cdot}$ of the adjacency matrix. Define the distance $x = |\pi^{-1}(i) - \pi^{-1}(j)|_{\text{ring}}$. Equivalently, the Hamming distance of the corresponding signal vectors satisfies $\mathrm{H}(B_{i\cdot}, B_{j\cdot}) = 2x$. Therefore the union of signal nodes for i, j-th row is of cardinality $|S_i \cup S_j| = k+x$, common signal nodes are of cardinality $|S_i \cap S_j| = k - x$, unique signal nodes are of cardinality $|S_i \triangle S_j| = 2x$, and $|S_i^c \cap S_j^c| = n - k - x - 2$. Each signal coordinate is 1 with probability $p = 1 - \beta(1 - \beta\frac{k}{n-1})$, while non-signal coordinate is 1 with probability $q = \beta\frac{k}{n-1}$, and we have

$$\langle A_{i\cdot}, A_{j\cdot}\rangle = \sum_{l\in S_i\cap S_j} A_{il}A_{jl} + \sum_{l\in S_i\triangle S_j} A_{il}A_{jl} + \sum_{l\in S_i^c\cap S_j^c} A_{il}A_{jl}.$$

Observe as long as $l \neq i, j$, A_{il} and A_{jl} are independent, and $\{A_{il}A_{jl}, l \in [n]\backslash\{i,j\}\}$ are independent of each other.

Let us bound each term via Bernstein's inequality Lemma 6,

$$\begin{aligned}\sum_{l\in S_i\cap S_j} A_{il}A_{jl} &\in p^2|S_i \cap S_j| \pm \left(\sqrt{2p^2|S_i\cap S_j|t} + \frac{1}{3}t\right)\\ \sum_{l\in S_i\triangle S_j} A_{il}A_{jl} &\in pq|S_i \triangle S_j| \pm \left(\sqrt{2pq|S_i\triangle S_j|t} + \frac{1}{3}t\right)\\ \sum_{l\in S_i^c\cap S_j^c} A_{il}A_{jl} &\in q^2|S_i^c \cap S_j^c| \pm \left(\sqrt{2q^2|S_i^c\cap S_j^c|t} + \frac{1}{3}t\right)\end{aligned}$$

with probability at least $1 - 6\exp(-t)$. We take $t = (2+\epsilon)\log n$ for any $\epsilon > 0$, such that with probability at least $1 - Cn^{-\epsilon}$, the above bound holds for all pairs (i, j).

Thus for all $|\pi^{-1}(i) - \pi^{-1}(j)|_{\text{ring}} > k$ pairs,

$$\begin{aligned}\langle A_{i\cdot}, A_{j\cdot}\rangle &\le 2kpq + (n-2k-2)q^2 \\ &\quad + \left(\sqrt{4kpqt} + \sqrt{2(n-2k-2)q^2t} + t\right),\end{aligned}$$

for $|\pi^{-1}(i) - \pi^{-1}(j)|_{\text{ring}} \le x$ pairs

$$\begin{aligned}&\langle A_{i\cdot}, A_{j\cdot}\rangle \ge (k-x)p^2 + 2xpq + (n-k-x-2)q^2 \\ &- \left(\sqrt{2(k-x)p^2t} + \sqrt{4xpqt} + \sqrt{2(n-k-x-2)q^2t} + t\right).\end{aligned}$$

Thus, with $t = (2+\epsilon)\log n$, $p = 1 - \beta(1-\beta\frac{k}{n-1})$ and $q = \beta\frac{k}{n-1}$, if $x < x_0$ with

$$\frac{x_0}{k} := 1 - C_1\sqrt{\frac{\log n}{k}}\frac{1}{1-\beta} - C_2\sqrt{\frac{\log n}{n}}\frac{1}{(1-\beta)^2},$$

we have

$$\begin{aligned}(k-x)(p-q)^2 &\ge 2t + (2\sqrt{2}+1)\left(\sqrt{kp^2} + \sqrt{nq^2}\right)\sqrt{2t}\\ &\ge 2t + \left(\sqrt{2kpq} + \sqrt{(n-2k-2)q^2} + \sqrt{(k-x)p^2}\right.\\ &\quad \left.+\sqrt{2xpq} + \sqrt{(n-k-x-2)q^2}\right)\sqrt{2t},\end{aligned}$$

which further implies,

$$\min_{j:|\pi^{-1}(i)-\pi^{-1}(j)|_{\text{ring}}\le x_0} \langle A_{i\cdot}, A_{j\cdot}\rangle \ge \max_{j\notin\mathcal{N}(v_i)} \langle A_{i\cdot}, A_{j\cdot}\rangle, \forall i$$

$$\begin{aligned}\max_{i\in[n]} \frac{|\hat{\mathcal{N}}(v_i)\triangle\mathcal{N}(v_i)|}{|\mathcal{N}(v_i)|} &\le \frac{k-x_0}{k}\\ &= C_1\sqrt{\frac{\log n}{k}}\frac{1}{1-\beta} + C_2\sqrt{\frac{\log n}{n}}\frac{1}{(1-\beta)^2}.\end{aligned}$$

Therefore we can reconstruct the neighborhood consistently, under the condition

$$1-\beta \succ \sqrt{\frac{\log n}{k}} \vee \left(\frac{\log n}{n}\right)^{1/4}.$$

□

Proof of Lemma 4. Since eigen structure is not affected by permutation, we will work under the case when the true permutation is identity. We work under a mild technical assumption that we have two independent observation of the adjacency matrix, one used for calculating the eigen-vector, the other used for projection to reduce dependency. Note this technical assumption only affect the signal $(1-\beta)$ to noise (k/n) ratio by a constant factor. Recall that $A = M + Z$, where $M = (1-\beta)(1-\beta\frac{k}{n-1}) \cdot B + \beta\frac{k}{n-1} \cdot (J - I)$ is the signal matrix. Denote the eigenvectors of M to be $U \in \mathbb{R}^{n\times n}$, and eigenvectors of A to be $\hat{U} \in \mathbb{R}^{n\times n}$. Denote the projection matrix corresponding the subspace of the second and third eigenvector $U_{\cdot 2}, U_{\cdot 3}$ to be H. Similarly $\hat{H}$ denotes the projection matrix to the 2-dim space spanned by $\hat{U}_{\cdot 2}, \hat{U}_{\cdot 3}$.

Classic Davis-Kahan perturbation bound informs us that two dimensional subspace $\hat{H}$ and H are close in spectral norm

$$\|\hat{H} - H\| \le \frac{\|Z\|}{\Delta\lambda - \|Z\|},$$

where the spectral gap $\Delta\lambda$ of M is

$$\begin{aligned}
&\Delta\lambda := (1-\beta)(1-\beta\frac{k}{n-1}) \cdot (\lambda_2 - \lambda_3) \\
&= (1-\beta)(1-\beta\frac{k}{n-1}) \cdot \left[2\sum_{i=1}^{k/2} \cos\left(i\frac{2\pi}{n}\right) - 2\sum_{i=1}^{k/2} \cos\left(i\frac{2\pi \cdot 2}{n}\right)\right] \\
&= (1-\beta)(1-\beta\frac{k}{n-1}) \left[\frac{2\sin\frac{k\pi}{2n}}{\sin\frac{\pi}{n}} \cos\frac{(k+2)\pi}{2n} - \frac{2\sin\frac{k\pi}{n}}{\sin\frac{2\pi}{n}} \cos\frac{(k+2)\pi}{n}\right] \\
&\asymp (1-\beta)(1-\beta\frac{k}{n-1})\frac{k^3}{n^2}.
\end{aligned}$$

From the proof of Lemma 2, we know with high probability

$$\|Z\| \preceq \sqrt{k} \vee \sqrt{\log n}.$$

Note we have for the true signal matrix M and true projection H

$$\begin{aligned}
HM_{\cdot i} &= \langle U_{\cdot 2}, M_{\cdot i}\rangle \cdot U_{\cdot 2} + \langle U_{\cdot 3}, M_{\cdot i}\rangle \cdot U_{\cdot 3}, \\
&= \frac{(1-\beta)\lambda_2}{\sqrt{n}} \cos\frac{(i-1)2\pi}{n} \cdot U_{\cdot 2} + \frac{(1-\beta)\lambda_2}{\sqrt{n}} \sin\frac{(i-1)2\pi}{n} \cdot U_{\cdot 3};
\end{aligned} \tag{22}$$

however, one only observes the noisy version $\hat{H}A_{\cdot i} \in \mathbb{R}^n$ (of the signal $HM_{\cdot i} \in \mathbb{R}^n$), which satisfies the equality

$$\hat{H}A_{\cdot i} = HM_{\cdot i} + (\hat{H} - H)M_{\cdot i} + \hat{H}Z_{\cdot i}.$$

Hence we have, uniformly for all i,

$$\begin{aligned}
\|\hat{H}A_{\cdot i} - HM_{\cdot i}\| &\le \|(\hat{H} - H)M_{\cdot i}\| + \|\hat{H}Z_{\cdot i}\| \\
&\le \|\hat{H} - H\|\|M_{\cdot i}\| + \|\hat{H}Z_{\cdot i}\| \\
&\le \frac{\sqrt{k} \vee \sqrt{\log n}}{\Delta\lambda - \sqrt{k} \vee \sqrt{\log n}} \cdot \sqrt{k}(1-\beta) + C\sqrt{\log n}
\end{aligned}$$

with probability $1 - n^{-c}$, with some constants $c, C > 0$. Here the last line follows from Davis-Kahan bound on $\|\hat{H} - H\|$ and Azuma-Hoeffding's inequality for $\langle \hat{U}_{\cdot 2}, Z_{\cdot i}\rangle$ and $\langle \hat{U}_{\cdot 3}, Z_{\cdot i}\rangle$ condition on $\hat{H}$. Denote this stochastic error as

$$\begin{aligned}
\delta &:= \frac{\sqrt{k} \vee \sqrt{\log n}}{\Delta\lambda - \sqrt{k} \vee \sqrt{\log n}} \cdot \sqrt{k}(1-\beta) + C\sqrt{\log n}, \\
&\asymp \frac{k}{\frac{k^3}{n^2}} = \frac{n^2}{k^2}.
\end{aligned}$$

The second line follows under the condition $1 - \beta \succsim \frac{n^2}{k^{2.5}}$, which is ensured under Eq. (23).

For any i, j with $|j - i|_{\text{ring}} = m$, Eq. (22) together with simple geometry implies

$$\begin{aligned}
&\|HM_{\cdot i} - HM_{\cdot j}\| \\
&= \frac{(1-\beta)\lambda_2}{\sqrt{n}} \cdot \left[\left(\cos\frac{(i-1)2\pi}{n} - \cos\frac{(j-1)2\pi}{n}\right)^2\right. \\
&\quad \left. + \left(\sin\frac{(i-1)2\pi}{n} - \sin\frac{(j-1)2\pi}{n}\right)^2\right]^{1/2} \\
&= \frac{(1-\beta)\lambda_2}{\sqrt{n}} \cdot 2\sin\frac{m\pi}{n}.
\end{aligned}$$

Therefore, fix any i, for $j \notin \mathcal{N}(v_i)$ not in i's neighborhood, using triangle inequality we have

$$\begin{aligned}
&\min_{j\notin\mathcal{N}(v_i)} \|\hat{H}A_{\cdot i} - \hat{H}A_{\cdot j}\| \ge \min_{j\notin\mathcal{N}(v_i)} \|HM_{\cdot i} - HM_{\cdot j}\| - 2\delta \\
&\ge \frac{(1-\beta)\lambda_2}{\sqrt{n}} \cdot 2\sin\frac{k\pi}{n} - 2\delta \\
&= \left(\frac{(1-\beta)\lambda_2}{\sqrt{n}} \cdot 2\sin\frac{k\pi}{n} - 4\delta\right) + 2\delta \\
&\ge \max_{|j-i|_{\text{ring}} < m} \|\hat{H}A_{\cdot i} - \hat{H}A_{\cdot j}\|
\end{aligned}$$

with

$$m = \frac{n}{\pi} \arcsin\left(\sin\frac{k\pi}{n} - 2\delta\frac{\sqrt{n}}{\lambda_2}\frac{1}{1-\beta}\right).$$

Therefore the following bound on symmetric set difference holds

$$\begin{aligned}
\max_{i\in[n]} \frac{|\hat{\mathcal{N}}(v_i) \triangle \mathcal{N}(v_i)|}{|\mathcal{N}(v_i)|} &\le 1 - \frac{\arcsin\left(\sin\frac{k\pi}{n} - 2\delta\frac{\sqrt{n}}{\lambda_2}\frac{1}{1-\beta}\right)}{\frac{k\pi}{n}} \\
&\le C' \cdot \frac{\frac{n^2}{k^2}\frac{\sqrt{n}}{k}\frac{1}{1-\beta}}{\frac{k}{n}} \asymp \frac{n^{3.5}}{k^4}\frac{1}{1-\beta}.
\end{aligned}$$

In summary under the condition

$$1 - \beta \succ \frac{n^{3.5}}{k^4}, \tag{23}$$

one can recover the neighborhood consistently w.h.p. in the sense

$$\lim_{n,k(n)\to\infty} \max_{i\in[n]} \frac{|\hat{\mathcal{N}}(v_i) \triangle \mathcal{N}(v_i)|}{|\mathcal{N}(v_i)|} = 0.$$

□

ACKNOWLEDGEMENT

The authors thank Elchanan Mossel for many helpful discussions, and thank anonymous referees for valuable suggestions on improving the paper.

References

Emmanuel Abbe and Colin Sandon. Community detection in general stochastic block models: fundamental limits and efficient recovery algorithms. *arXiv preprint arXiv:1503.00609*, 2015.

Luıs A Nunes Amaral, Antonio Scala, Marc Barthelemy, and H Eugene Stanley. Classes of small-world networks. *Proceedings of the national academy of sciences*, 97(21): 11149–11152, 2000.

Afonso S Bandeira and Ramon van Handel. Sharp nonasymptotic bounds on the norm of random matrices with independent entries. *arXiv preprint arXiv:1408.6185*, 2014.

Alain Barrat and Martin Weigt. On the properties of small-world network models. *The European Physical Journal B-Condensed Matter and Complex Systems*, 13(3):547–560, 2000.

Itai Benjamini and Noam Berger. The diameter of long-range percolation clusters on finite cycles. *arXiv preprint math/0012070*, 2000.

Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence*. OUP Oxford, 2013.

Rainer E Burkard, Eranda Çela, Panos M Pardalos, and Leonidas S Pitsoulis. *The quadratic assignment problem*. Springer, 1998.

Eranda Cela. *The quadratic assignment problem: theory and algorithms*, volume 1. Springer Science & Business Media, 2013.

Venkat Chandrasekaran, Pablo A Parrilo, and Alan S Willsky. Convex graph invariants. *SIAM Review*, 54(3):513–541, 2012.

Don Coppersmith, David Gamarnik, and Maxim Sviridenko. The diameter of a long range percolation graph. In *Proceedings of the thirteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 329–337. Society for Industrial and Applied Mathematics, 2002.

Aurelien Decelle, Florent Krzakala, Cristopher Moore, and Lenka Zdeborová. Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. *Physical Review E*, 84(6):066106, 2011.

Robert M Gray. *Toeplitz and circulant matrices: A review*. now publishers inc, 2006.

Ki-ichiro Hashimoto. Zeta functions of finite graphs and representations of p-adic groups. *Automorphic forms and geometry of arithmetic varieties.*, pages 211–280, 1989.

Jon Kleinberg. The small-world phenomenon: An algorithmic perspective. In *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, pages 163–170. ACM, 2000.

Donald Ervin Knuth. *The Stanford GraphBase: a platform for combinatorial computing*, volume 37. Addison-Wesley Reading, 1993.

Vito Latora and Massimo Marchiori. Efficient behavior of small-world networks. *Physical review letters*, 87(19): 198701, 2001.

Vince Lyzinski, Donniell E Fishkind, and Carey E Priebe. Seeded graph matching for correlated erdös-rényi graphs. *Journal of Machine Learning Research*, 15(1):3513–3540, 2014.

Laurent Massoulié. Community detection thresholds and the weak ramanujan property. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 694–703. ACM, 2014.

Stanley Milgram. The small world problem. *Psychology today*, 2(1):60–67, 1967.

Cristopher Moore and Mark EJ Newman. Epidemics and percolation in small-world networks. *Physical Review E*, 61(5):5678, 2000.

Elchanan Mossel, Joe Neeman, and Allan Sly. A proof of the block model threshold conjecture. *arXiv preprint arXiv:1311.4115*, 2013.

Mark EJ Newman and Duncan J Watts. Scaling and percolation in the small-world network model. *Physical Review E*, 60(6):7332, 1999.

CJ Stam, BF Jones, G Nolte, M Breakspear, and Ph Scheltens. Small-world networks and functional connectivity in alzheimer's disease. *Cerebral cortex*, 17(1):92–99, 2007.

Michel Talagrand. A new look at independence. *The Annals of probability*, pages 1–34, 1996.

Terence Tao. *Topics in random matrix theory*, volume 132. American Mathematical Soc., 2012.

Alexandre B Tsybakov. *Introduction to nonparametric estimation*, volume 11. Springer Series in Statistics, 2009.

Remco Van Der Hofstad. Random graphs and complex networks. *Available on http://www. win. tue. nl/rhofstad/NotesRGCN. pdf*, page 11, 2009.

Duncan J Watts. *Small worlds: the dynamics of networks between order and randomness*. Princeton university press, 1999.

Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world'networks. *nature*, 393(6684):440–442, 1998.