Submitted to Bernoulli

Power Enhancement and Phase Transitions for Global Testing of the Mixed Membership Stochastic Block Model

LOUIS CAMMARATA[1] and ZHENG TRACY KE[2,*]

[1]*Statistics Department, Harvard University, Cambridge, MA 02138, USA.*
E-mail: louis_cammarata@g.harvard.edu
[2]*Statistics Department, Harvard University, Cambridge, MA 02138, USA.*
E-mail: [*]zke@fas.harvard.edu

The mixed-membership stochastic block model (MMSBM) is a common model for social networks. Given an n-node symmetric network generated from a K-community MMSBM, we would like to test $K = 1$ versus $K > 1$. We first study the degree-based χ^2 test and the orthodox Signed Quadrilateral (oSQ) test. These two statistics estimate an order-2 polynomial and an order-4 polynomial of a "signal" matrix, respectively. We derive the asymptotic null distribution and power for both tests. However, for each test, there exists a parameter regime where its power is unsatisfactory. It motivates us to propose a power enhancement (PE) test to combine the strengths of both tests. We show that the PE test has a tractable null distribution and improves the power of both tests. To assess the optimality of PE, we consider a randomized setting, where the n membership vectors are independently drawn from a distribution on the standard simplex. We show that the success of global testing is governed by a quantity $\beta_n(K, P, h)$, which depends on the community structure matrix P and the mean vector h of memberships. For each given (K, P, h), a test is called *optimal* if it distinguishes two hypotheses when $\beta_n(K, P, h) \to \infty$. A test is called *optimally adaptive* if it is optimal for all (K, P, h). We show that the PE test is optimally adaptive, while many existing tests are only optimal for some particular (K, P, h), hence, not optimally adaptive.

Keywords: Chi-square test; degree matching; mixed memberships; phase transition; signed cycles; stochastic block model.

1. Introduction

Statistical analysis of large social networks has received much recent attention. In this paper, we are interested in testing whether an undirected network has one community or multiple communities (a.k.a., global testing). This problem has several applications. One is measuring neighborhood diversity of individual nodes [9, 12]. For example, [12] constructed a co-authorship network of statisticians using papers in 36 journals from 1975 to 2015; they conducted global testing on the 'personalized network' of each node (author) and used the obtained *p*-value to measure the collaboration diversity of this author. Another application is recursive community detection, where a large network is partitioned into sub-networks in a hierarchical way and global testing is used repeatedly to decide whether each sub-network should be further partitioned [12]. Theoretically, studying global testing (especially the lower bound) also provides valuable insights for the statistical limits of related problems, such as community detection [23], mixed membership estimation [13], and estimation of the number of communities [16].

Consider an undirected network with n nodes. The adjacency matrix $A = (A_{ij})_{1 \le i,j \le n}$ is a symmetric matrix, where

$$A_{ij} = \begin{cases} 1, & \text{if } i \neq j \text{ and there is an edge between node } i \text{ and node } j, \\ 0, & \text{otherwise.} \end{cases} \tag{1.1}$$

Assuming that the network has K perceivable communities, we model A with the Mixed-Membership Stochastic Block Model (MMSBM) [1] as follows. The mixed-membership vector of node i is a weight vector $\pi_i \in \mathbb{R}^K$ such that $\pi_i(k) \geq 0$ and $\sum_{k=1}^K \pi_i(k) = 1$, with $\pi_i(k)$ denoting the 'weight' that node i puts on community k. For a symmetric nonnegative matrix $P \in \mathbb{R}^{K\times K}$ that models the community structure, we assume that the upper triangle of A contains independent Bernoulli variables, where

$$\mathbb{P}(A_{ij}=1) = \pi_i' P \pi_j, \qquad 1 \leq i < j \leq n. \tag{1.2}$$

The well-known Stochastic Block Model (SBM) and Erdös-Renyi (ER) model are special cases of MMSBM. When all π_i's are degenerate (meaning that one entry of π_i is 1 and all other entries are 0), MMSBM reduces to SBM; furthermore, if $K=1$, then SBM reduces to ER.

The global testing problem is formulated as testing between the two hypotheses:

$$H_0^{(n)} : K = 1 \qquad \text{versus} \qquad H_1^{(n)} : K > 1. \tag{1.3}$$

In MMSBM, write $\Omega = \Pi P \Pi'$, with $\Pi = [\pi_1, \pi_2, \ldots, \pi_n]' \in \mathbb{R}^{n\times K}$. We call Ω the Bernoulli probability matrix. It follows that, for some $\alpha_n > 0$,

$$\mathbb{E}A = \Omega - \mathrm{diag}(\Omega), \qquad \text{with} \quad \Omega = \begin{cases} \alpha_n \mathbf{1}_n \mathbf{1}_n', & \text{under } H_0^{(n)}, \\ \Pi P \Pi', & \text{under } H_1^{(n)}. \end{cases} \tag{1.4}$$

The signals to separate two hypotheses are captured by the following $n \times n$ matrix:

$$\widetilde{\Omega} = \Omega - \tilde{\alpha}_n \mathbf{1}_n \mathbf{1}_n', \qquad \text{where} \quad \tilde{\alpha}_n = n^{-2}(\mathbf{1}_n' \Omega \mathbf{1}_n). \tag{1.5}$$

The null hypothesis holds if and only if $\widetilde{\Omega}$ is a zero matrix. We will investigate testing ideas that aim to estimate a polynomial of (the entries of) $\widetilde{\Omega}$:

- We consider the *degree-based* χ^2 statistic, which targets on estimating $\mathbf{1}_n' \widetilde{\Omega}^2 \mathbf{1}_n$.
- For each $m \geq 3$, we consider the order-m *orthodox Signed Polygon* statistic, which targets on estimating $\mathrm{tr}(\widetilde{\Omega}^m)$.

Here, $\mathbf{1}_n' \widetilde{\Omega}^2 \mathbf{1}_n$ is a *second order polynomial* of $\widetilde{\Omega}$, and $\mathrm{tr}(\widetilde{\Omega}^m)$ is an *m-th order polynomial* of $\widetilde{\Omega}$, for $m \geq 3$. There is no natural testing idea based on estimating a first order polynomial of $\widetilde{\Omega}$. For example, $\mathbf{1}_n' \widetilde{\Omega} \mathbf{1}_n$ is always equal to zero and hence useless for global testing; $\mathrm{tr}(\widetilde{\Omega})$ is hard to estimate, partially because the diagonals of A are always zero (i.e., self-edges are not allowed).

We study the asymptotic performances of the above tests. We also derive information theoretic lower bounds for this global testing problem. By comparing the upper/lower bounds, we discover: (i) None of the above tests can attain the lower bound across all parameter regimes; (ii) In some parameter regimes, the degree-based χ^2 test (abbreviated as the χ^2 test) is optimal; and in the remaining parameter regimes, the order-4 orthodox Signed Polygon test (also called the orthodox Signed Quadrilateral test, abbreviated as the oSQ test) is optimal. This motivates us to design a new test statistic to combine the strengths of the χ^2 test and the oSQ test.

We propose the *Power Enhancement* (PE) test. It is inspired by a key result about the joint distribution of the χ^2 test statistic and the oSQ test statistic: Under the null hypothesis, they jointly converge to a bivariate normal distribution with a covariance matrix I_2; especially, the two test statistics are asymptotically uncorrelated. The PE test statistic is defined as the sum of squares of these two test statistics, and it converges to a $\chi_2^2(0)$ distribution under the null hypothesis. Therefore, we can conveniently control the level of the PE test.

To assess the power and optimality of the PE test, we adopt the phase transition framework in Jin et al. [15]. For arbitrary parameters (K, P) and distribution F on the probability simplex of $\mathbb{R}^K$, writing $h = \mathbb{E}_{\pi\sim F}[\pi]$, we consider the following pair of hypotheses:

$$\Omega = \begin{cases} \alpha_0 \mathbf{1}_n \mathbf{1}_n', \text{ where } \alpha_0 = h'Ph, & \text{under } H_0^{(n)}, \\ \Pi P \Pi', \quad \text{where } \pi_i \overset{iid}{\sim} F, & \text{under } H_1^{(n)}. \end{cases} \tag{1.6}$$

As $n \to \infty$, we fix K and allow (P, h) to depend on n (i.e., we consider a sequence of (P_n, h_n) indexed by n). Our lower bound result tells when the chi-square distance between two hypotheses converges to 0 for *every* (K, P_n, h_n). In particular, we identify a quantity (as before, $\alpha_0 = h_n' P_n h_n$)

$$\beta_n(K, P_n, h_n) \equiv \max\{n^{3/2}\alpha_0^{-1}\|P_n h_n - \alpha_0 \mathbf{1}_K\|^2,\ n^2\alpha_0^{-2}\|P_n - \alpha_0 \mathbf{1}_K \mathbf{1}_K'\|^4\}, \tag{1.7}$$

such that the chi-square distance between two hypotheses tends to 0 if $\beta_n(K, P_n, h_n) \to 0$. We call the parameter regimes where $\beta_n(K, P_n, h_n) \to 0$ the *Region of Impossibility*. In this region, the two hypotheses are asymptotically inseparable. We call the parameter regimes where $\beta_n(K, P_n, h_n) \to \infty$ the *Region of Possibility*. A test is called *optimally adaptive* if it is able to distinguish two hypotheses for *any* (K, P_n, h_n) in the Region of Possibility. We show that the PE test is optimally adaptive.

1.1. Related literature

The likelihood ratio test (LRT) was studied by Mossel et al. [20] and Banerjee and Ma [4] for a special case of $K = 2$, $P_n = (a_n - b_n)I_2 + b_n \mathbf{1}_2 \mathbf{1}_2'$ and $h_n = (1/2, 1/2)'$. The LRT may be generalized to other (K, P_n, h_n), but it requires prior knowledge of parameters in the alternative hypothesis. By the Neyman-Pearson lemma, the LRT has the highest power; however, it is not a polynomial-time test. The tests we study here, χ^2, oSQ and PE, need no prior knowledge of the parameters, and are polynomial-time tests.

The eigenvalue-based tests were also studied before. For example, Lei [18] used the maximum singular value of the centered and rescaled adjacency matrix as test statistic. However, the eigenvalue-based tests are not optimally adaptive: their SNRs are linked to the second term in (1.7); hence, in the Region of Possibility, for those (K, P_n, h_n) such that the first term in (1.7) $\to \infty$ but the second term $\to 0$, the eigenvalue-based tests are unable to separate two hypotheses.

Arias-Castro and Verzelen [2] considered the testing of a planted clique model v.s. the ER model. The planted clique model can be viewed as a special case of MMSBM with $K = 2$, $P_n = b_n \mathbf{1}_2 \mathbf{1}_2' + (a_n - b_n)e_1 e_1'$ and $h = (\epsilon_n, 1 - \epsilon_n)'$, where $\epsilon_n = o(1)$ and $a_n > b_n > 0$. They derived the optimal detection boundaries for many different cases of $(n, a_n, b_n, \epsilon_n)$ and proposed optimal tests. When b_n is unknown and $n\epsilon_n \gg n^{3/2}$, they used the χ^2 test (called the *degree variance test* in their paper). Our result about the SNR of the χ^2 test agrees with their result (see their Table 1, the bottom right cell) in this special case. However, there are major differences between two papers: First, they focused on a particular (K, P_n, h_n), for which the first term in (1.7) always dominates, so that the χ^2 test alone is enough to achieve optimality (provided that $n\epsilon_n \gg n^{3/2}$). In contrast, we seek to find an *optimally adaptive* test that works for a broad collection of (K, P_n, h_n), where the power enhancement idea is crucial. Second, we focus on $\epsilon_n \asymp 1$, but their main interest was in $\epsilon_n = o(1)$. In their setting, they could take advantage of sparsity by using the scan tests, which is unnecessary in our setting. Last, we provide the asymptotic null distribution for the χ^2 test, which was not given in [2].

The cycle count statistics were also studied in recent literature [3, 5, 7, 9, 14, 15, 19]. Our oSQ test is the same as the order-4 *signed-cycle* statistic introduced by Bubeck et al. [7] (also, see Banerjee [3]).

Under a 2-community SBM model (and the related contextual SBM model and Gaussian covariance model), Banerjee and Ma [5] and Lu and Sen [19] derived asymptotic distributions of order-m cycle count statistics for a general m. However, these works focused on the special case of $K = 2$, $P_n = (a_n - b_n)I_2 + b_n \mathbf{1}_2\mathbf{1}_2'$ and $h = (1/2, 1/2)'$, in which the oSQ test alone is enough to attain optimality. We seek to find an optimally adaptive test that works for rather arbitrary (K, P_n, h_n), where we do need to combine oSQ with the χ^2 test to achieve optimality. Moreover, these works only studied the asymptotic behavior of cycle counts, but we study the joint distribution of the cycle count and the χ^2 statistic (this is one of our key results that inspires the PE test). Last, none of these works revealed the phase transition in (1.7).

Jin et al. [15] studied the phase transition of global testing under the Degree-Corrected Mixed Membership (DCMM) model [13], a model more general than the MMSBM considered here. They proposed the Signed Quadrilateral (SQ) test and showed that it is optimally adaptive. Although our model is a special case of DCMM with no degree heterogeneity, the phase transition and the optimal test are different. Restricting from DCMM to MMSBM, the prior knowledge of 'no degree heterogeneity' brings additional signals for separating two hypotheses. For example, under DCMM, one can construct a pair of null and alternative hypotheses such that the expected degree of each node is perfectly matched under two hypotheses [15], and so the χ^2 statistic contains no signals and power enhancement is useless. But such 'degree-matched' hypothesis pairs do not exist when we restrict to MMSBM; for MMSBM, power enhancement is crucial for achieving optimality. [15] showed that the Region of Possibility and Region of Impossibility for global testing under DCMM are determined by $\gamma_n = |\lambda_2(\Omega)|/\sqrt{\lambda_1(\Omega)}$. This quantity restricted to MMSBM is different from β_n in (1.7). Moreover, the SQ test in [15] is also different from our oSQ test, so their results about the asymptotic behavior of the SQ test cannot imply our results about the oSQ test.

1.2. Content

We have made several contributions in this paper:

- We derive the phase transitions for global testing under MMSBM, where the Region of Impossibility and Region of Possibility are determined by the simple quantity $\beta_n(K, P_n, h_n)$.
- We study the (degree-based) χ^2 test and the oSQ test. For each test statistic, we derive its asymptotic distribution under the null hypothesis and SNR under the alternative hypothesis. We also derive the asymptotic *joint null* distribution of these two test statistics.
- We propose the Power Enhancement (PE) test to combine the strengths of the oSQ test and the χ^2 test while overcoming their respective limitations.
- We show that the PE test statistic has an asymptotic null distribution of $\chi_2^2(0)$. We also show that the PE test is optimally adaptive. In comparison, several popular tests are not optimally adaptive.

Below, in Section 2, we formally introduce the χ^2, oSQ and PE tests, and explain how PE combines strengths of the other two tests. In Section 3, we present the main theoretical results, including asymptotic properties of three test statistics, lower bounds and phase transitions. Section 4 is of independent interest, where we discuss the identifiability of parameters of MMSBM, especially, the identifiability of K. Section 5 contains simulation studies, and Section 6 concludes the paper.

2. Three test statistics

Recall that A is the adjacency matrix. Let $d = A\mathbf{1}_n$ be the vector of degrees, and $\bar{d} = \frac{1}{n}\sum_{i=1}^n d_i$ be the average degree. Under the null hypothesis, $\Omega = \alpha_n \mathbf{1}_n \mathbf{1}_n'$, and a good estimate of α_n is

$$\hat{\alpha}_n = [n(n-1)]^{-1}\mathbf{1}_n' A \mathbf{1}_n = (n-1)^{-1}\bar{d}. \tag{2.1}$$

Under the null hypothesis, $(d_i - \bar{d})/\sqrt{(n-1)\hat{\alpha}_n(1-\hat{\alpha}_n)} \approx \mathcal{N}(0,1)$, for each $1 \leq i \leq n$. Aggregating these terms for all i gives rise to the *degree-based* χ^2 *test statistic* (also known as the *degree of variance statistic* [2]; throughout this paper, we call it the χ^2 test for short):

$$X_n = [(n-1)\hat{\alpha}_n(1-\hat{\alpha}_n)]^{-1}\sum_{i=1}^n (d_i - \bar{d})^2.^{1} \tag{2.2}$$

The χ^2 test looks for evidence against the null hypothesis through degree heterogeneity. Since all nodes have the same expected degree under the null hypothesis, a significant degree heterogeneity provides a strong evidence against the null. By some simple calculations, we find that

$$[(n-1)\hat{\alpha}_n(1-\hat{\alpha}_n)](X_n - n) = \sum_{i_1,i_2,i_3 \text{ (distinct)}} (A_{i_1 i_2} - \hat{\alpha}_n)(A_{i_2 i_3} - \hat{\alpha}_n). \tag{2.3}$$

Recall that we have introduced $\widetilde{\Omega}$ in (1.5). The matrix $A - \hat{\alpha}_n \mathbf{1}_n \mathbf{1}_n'$ is a stochastic proxy of $\widetilde{\Omega}$. Therefore, the right hand side above is approximately $\sum_{i_1,i_2,i_3 \text{ (distinct)}} \widetilde{\Omega}_{i_1 i_2}\widetilde{\Omega}_{i_2 i_3} \approx \mathbf{1}_n'\widetilde{\Omega}^2\mathbf{1}_n$. This suggests that $[(n-1)\hat{\alpha}_n(1-\hat{\alpha}_n)](X_n - n)$ is an estimate of $\mathbf{1}_n'\widetilde{\Omega}^2\mathbf{1}_n$, under the alternative hypothesis.

The *orthodox Signed Polygon* is a family of statistics that extends the *Signed Triangle* statistic [7], where for $m = 3, 4, \ldots$, the m-th order statistic in the family is defined as

$$U_n^{(m)} = \sum_{i_1,i_2,\ldots i_m \text{ (distinct)}} (A_{i_1 i_2} - \hat{\alpha}_n)(A_{i_2 i_3} - \hat{\alpha}_n)\ldots(A_{i_m i_1} - \hat{\alpha}_n). \tag{2.4}$$

Centering each A_{ij} by $\hat{\alpha}_n$ is reasonable, because $\mathbb{E}[A_{ij}] = \alpha_n$ under the null hypothesis. It was noted in [15] that the Signed Polygon statistic may experience signal cancellation if m is odd, but it can avoid signal cancellation if m is even. For this reason, it is preferred to only consider the even order statistics in the family. In this paper, we focus our discussion on the orthodox Signed Quadrilateral (oSQ), which corresponds to the smallest even order m (i.e. $m = 4$):

$$Q_n = \sum_{i_1,i_2,i_3,i_4 \text{ (distinct)}} (A_{i_1 i_2} - \hat{\alpha}_n)(A_{i_2 i_3} - \hat{\alpha}_n)(A_{i_3 i_4} - \hat{\alpha}_n)(A_{i_4 i_1} - \hat{\alpha}_n). \tag{2.5}$$

Again, since $A - \hat{\alpha}_n \mathbf{1}_n \mathbf{1}_n'$ is a stochastic proxy of $\widetilde{\Omega}$, the right hand side above is approximately equal to $\sum_{i_1,i_2,i_3,i_4 \text{ (distinct)}} \widetilde{\Omega}_{i_1 i_2}\widetilde{\Omega}_{i_2 i_3}\widetilde{\Omega}_{i_3 i_4}\widetilde{\Omega}_{i_4 i_1} \approx \mathrm{tr}(\widetilde{\Omega}^4)$. In other words, Q_n is an estimate of $\mathrm{tr}(\widetilde{\Omega}^4)$, under the alternative hypothesis.

These statistics are reminiscent of the classical moment statistics, as X_n and Q_n estimate an order-2 polynomial and an order-4 polynomial of (the entries of) $\widetilde{\Omega}$, respectively. We now recall

[1] In the rare event that $\hat{\alpha}_n = 0$ or $\hat{\alpha}_n = 1$, we replace $\hat{\alpha}_n$ by $\frac{2}{n(n-1)}$ and $\frac{n(n-1)-2}{n(n-1)}$, respectively, to make the χ^2 test statistic well defined. Similar operations apply to the oSQ and PE tests.

some conventional insights of classical moment statistics. Suppose that we observe independent data $X_i \sim \mathcal{N}(\mu_i, \sigma^2)$, for $1 \le i \le n$, and we would like to test the global null hypothesis

$$H_0: \; \mu_1 = \mu_2 = \cdots = \mu_n = 0.$$

Consider two moment statistics $S_1 = \frac{1}{n}\sum_{i=1}^n X_i$ and $S_2 = \frac{1}{n}\sum_{i=1}^n X_i^2$. It is well-known that, on the one hand, if the μ_i's have the same sign, the lower-order moment statistic S_1 has a better detection boundary than the higher-order moment statistic S_2; on the other hand, if the μ_i's have different signs, S_1 faces 'signal cancellation', but S_2 has no such issue. Hence, if one only cares about the worst-case performance, using S_2 is enough. However, going beyond the 'worst case', there are many cases where the power of S_2 is inferior to S_1, so using S_1 to enhance power will be useful.

In the network global testing we consider here, X_n is analogous to S_1, and Q_n is analogous to S_2. In Section 2.1, we will study the signal-to-noise ratios (SNRs) of X_n and Q_n. We have the following observations: (i) The SNR of X_n can be zero even when $\widetilde{\Omega}$ is a nonzero matrix, so the χ^2 test faces potential signal cancellation. (ii) The SNR of Q_n is always nonzero as long as $\widetilde{\Omega}$ is a nonzero matrix, but there exist cases where the SNR of Q_n is strictly smaller than the SNR of X_n. Therefore, if we want a test that performs uniformly well in all cases, we should combine these two statistics to simultaneously avoid signal cancellation and enhance power.

In order for the combined test statistic to have a tractable null distribution, we must derive the asymptotic joint distribution of X_n and Q_n. We show in Theorem 3.1 that with mild regularity conditions,

$$\begin{bmatrix} (2n)^{-1/2}(X_n - n) \\ \big(2\sqrt{2}n^2\hat{\alpha}_n^2\big)^{-1} Q_n \end{bmatrix} \qquad \xrightarrow[n\to\infty]{\mathcal{L}} \qquad \mathcal{N}(0, I_2), \qquad \text{under } H_0^{(n)}.$$

Hence, a convenient way to combine both test statistics is to construct

$$S_n = (2n)^{-1}(X_n - n) + (8n^4\hat{\alpha}_n^4)^{-1} Q_n^2, \tag{2.6}$$

which asymptotically follows the $\chi_2^2(0)$ distribution under the null hypothesis. We call S_n the *Power Enhancement (PE)* statistic. We will show that the PE test combines the strengths of the χ^2 test and the oSQ test while overcoming their respective limitations.

2.1. Comparison of the signal-to-noise ratios

Let $\mathbb{E}_0$ and Var_0 denote the expectation and variance operators under the null hypothesis. Similarly, we have the notations $\mathbb{E}_1$ and Var_1 for the alternative hypothesis. The signal-to-noise ratio (SNR) of a test statistic T_n is defined as

$$\text{SNR}(T_n) \equiv \frac{|\mathbb{E}_1[T_n] - \mathbb{E}_0[T_n]|}{\sqrt{\max\{\text{Var}_0(T_n), \text{Var}_1(T_n)\}}}.$$

Under the alternative hypothesis, given (Π, P), let $h = n^{-1}\sum_{i=1}^n \pi_i$ and $\alpha_0 = h'Ph$. In Theorem 3.2, we show that the SNR of the χ^2 statistic X_n is captured by

$$\delta_n(K, P, h) = n^{3/2}\alpha_0^{-1}\|Ph - \alpha_0 \mathbf{1}_K\|^2. \tag{2.7}$$

Note that for any node i, the difference of expected degree under two hypothesis is $\mathbb{E}_1[d_i] - \mathbb{E}_0[d_i] = \sum_{j\neq i}(\pi_i' P \pi_j - \alpha_0) \approx n\pi_i'(Ph - \alpha_0\mathbf{1}_K)$. Hence, the χ^2 test finds evidence to reject the null hypothesis from node degrees. The χ^2 test can successfully separate two hypotheses as long as $\delta_n \to \infty$. The oSQ test looks for evidence against the null hypothesis from $\mathrm{tr}(\widetilde{\Omega}^4)$. In Theorem 3.3, we show that the SNR of the oSQ statistic Q_n is captured by

$$\tau_n(K, P, h) = n^2 \alpha_0^{-2} \|P - \alpha_0 \mathbf{1}_K \mathbf{1}_K'\|^4. \tag{2.8}$$

It can successfully separate two hypotheses as long as $\tau_n \to \infty$. The SNR of the PE statistic is captured by the quantity

$$\sqrt{\delta_n^2 + \tau_n^2} \quad \asymp \quad \beta_n(K, P, h) \equiv \max\{\delta_n, \tau_n\}. \tag{2.9}$$

The PE test can successfully separate two hypotheses as long as $\beta_n \to \infty$.

The SNR of PE improves those of χ^2 and oSQ. Such improvement can be significant. Consider the case of $Ph \propto \mathbf{1}_K$ (e.g., when P has equal diagonals and equal off-diagonals, and the communities have equal size), $\delta_n = 0$; also, it can be shown that τ_n is always nonzero. In this case, the χ^2 test faces signal cancellation and loses power, but the PE test still has power. At the same time, when the community signals are very weak (e.g., P is only a tiny perturbation of $\alpha_0 \mathbf{1}_K \mathbf{1}_K'$), there exist cases where $\tau_n \to 0$ but $\delta_n \to \infty$. Then, the oSQ test is unsatisfactory, but the PE test is still satisfactory. It is illuminating to understand these cases from examples.

Example 1 (*Two-community model*). Fix $K = 2$. For $a_n, b_n, d_n > 0$ and $\epsilon_n \in (0, 1)$,

$$P = \begin{bmatrix} a_n & b_n \\ b_n & d_n \end{bmatrix}, \qquad \text{and} \qquad h = \begin{bmatrix} \epsilon_n \\ 1 - \epsilon_n \end{bmatrix}.$$

Write $\bar{a}_n = (a_n + d_n)/2$. Suppose that $|a_n - b_n| = O(a_n + b_n)$. By calculations in Appendix A [8], we obtain the order of δ_n and τ_n in several cases ("S" stands for "symmetric", and "AS" stands for "asymmetric"):

Case	Symmetry in h	Symmetry in P	SNR of X_n	SNR of Q_n
S	$\epsilon_n = 1/2$	$d_n = a_n$	0	$n^2 \big[\frac{(a_n - b_n)^2}{a_n + b_n}\big]^2$
AS1	$\epsilon_n \neq 1/2$	$d_n = a_n$	$(1 - 2\epsilon_n)^2 n^{3/2} \frac{(a_n - b_n)^2}{a_n + b_n}$	$n^2 \big[\frac{(a_n - b_n)^2}{a_n + b_n}\big]^2$
AS2	$\epsilon_n = 1/2$	$\lvert d_n - a_n\rvert \gg \lvert\bar{a}_n - b_n\rvert$	$n^{3/2} \frac{(d_n - a_n)^2}{\bar{a}_n + b_n}$	$n^2 \big[\frac{(d_n - a_n)^2}{\bar{a}_n + b_n}\big]^2$
AS3	$\epsilon_n = 1/2$	$\lvert d_n - a_n\rvert \gg \lvert\bar{a}_n - b_n\rvert$	$n^{3/2} \frac{(d_n - a_n)^2}{\bar{a}_n + b_n}$	$n^2 \big[\frac{(\bar{a}_n - b_n)^2}{\bar{a}_n + b_n}\big]^2$

In Case (S) (it includes the 2-community SBM in [20, 4] as a special case), the χ^2 test loses power due to the fact that $\delta_n = 0$, and the oSQ test and the PE test have full power provided that

$$(a_n - b_n)^2/(a_n + b_n) \;\gg\; n^{-1}.$$

In Case (AS1), suppose $|1 - 2\epsilon_n| \geq c$ for a constant $c > 0$. Then, if

$$n^{-3/2} \;\ll\; (a_n - b_n)^2/(a_n + b_n) \;\ll\; n^{-1},$$

the oSQ test does not have full power but the χ^2 test and the PE test have full power.

Example 2 (*Rank-1 model*). Fix $K > 1$ and consider an MMSBM with $P = \eta\eta'$, for some nonnegative vector $\eta \in \mathbb{R}^K$ such that $\|\eta\|_\infty \leq 1$ and $\eta \not\propto \mathbf{1}_K$. This is an example where K is not identifiable. In

Section 4, we will define the *intrinsic number of communities (INC)*, which is the smallest K_0 such that this model can be written as a K_0-community MMSBM. By calculations in Appendix B [8], the INC for this example is 2 (regardless of K), so the alternative hypothesis holds. Consider a special case of $K = 2$:

$$P = \eta\eta', \qquad \eta = \frac{\sqrt{c_n}}{\sqrt{a_n^2 + b_n^2}} \begin{bmatrix} a_n \\ b_n \end{bmatrix}, \qquad h = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix},$$

where $a_n, b_n > 0$ and $c_n \in (0, 1)$. By direct calculations in Appendix B, when $|a_n - b_n| = O(a_n + b_n)$,

$$\delta_n \asymp n^{3/2} c_n \frac{(a_n - b_n)^2}{(a_n^2 + b_n^2)}, \qquad \tau_n \asymp n^2 c_n^2 \frac{(a_n - b_n)^4}{(a_n^2 + b_n^2)^2} \asymp n^{-1}\delta_n^2.$$

It is seen that $\tau_n \to \infty$ implies $\delta_n \to \infty$. Hence, whenever the oSQ test has full power, the χ^2 test also has full power, so does the PE test. However, when $\delta_n \to \infty$ but $\delta_n \ll \sqrt{n}$, the χ^2 test and the PE test both have full power, but the oSQ test does not have full power.

From these examples, the oSQ test outperforms the χ^2 test sometimes (e.g., in Case (S) of Example 1), and the χ^2 test outperforms the oSQ test sometimes (e.g., in Example 2). Power enhancement allows us to combine the strengths of both tests. Below, we further show that the PE test achieves the optimal phase transition.

2.2. A preview of the phase transition

The PE test can successfully separate two hypotheses if $\max\{\delta_n, \tau_n\} \to \infty$. In Theorem 3.5, we provide a matching lower bound: If $\max\{\delta_n, \tau_n\} \to 0$,

the chi-square distance between the probability densities of two hypotheses $\to 0$.

Hence, there exists no test that can asymptotically distinguish two hypotheses when $\max\{\delta_n, \tau_n\} \to 0$. It gives rise to the following phase transition: Consider the two dimensional phase space for MMSBM, where the x-axis is δ_n which calibrates the signals in node degrees, and the y-axis is τ_n which calibrates the signals in cycle counts. The phase space is divided into two regions (see Figure 1):

- *Region of Impossibility (*$\max\{\delta_n, \tau_n\} \to 0$*).* Any alternative in this region is inseparable from a null. For any test, the sum of Type I and Type II errors tends to 1 as $n \to \infty$.
- *Region of Possibility (*$\max\{\delta_n, \tau_n\} \to \infty$*).* Any alternative hypothesis in this region is separable from any null hypothesis. Specifically, the PE test statistic is able to separate two hypotheses, in the sense that for an appropriate threshold, the sum of Type I and Type II errors of the PE test tends to 0 as $n \to \infty$.

We say that a test is *optimally adaptive* if it can distinguish the null and alternative hypotheses in the *whole* Region of Possibility. The PE test is optimally adaptive, but neither the χ^2 test nor the oSQ test is. The χ^2 test can only distinguish two hypotheses in the sub-region of $\delta_n \to \infty$, and the oSQ test can only distinguish two hypotheses in the sub-region of $\tau_n \to \infty$. See Figure 1.

Remark 1 *(Connection to minimax optimality)*: Our phase transition results are more informative than the standard *minimax* results. The detection boundary $\beta_n(K, P, h)$ in (2.9) is for arbitrary (K, P, h), while the minimax detection boundary is the specific value of $\beta_n(K, P, h)$ at some worst-case (K, P, h)

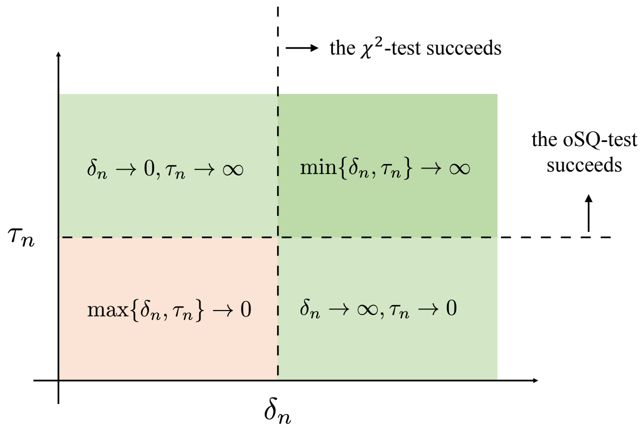


Figure 1: The phase transition of global detection for MMSBM. The light red region is the Region of Impossibility, and the three other green regions constitute the region of Possibility.

in a pre-specified class. Therefore, for a test to be *minimax optimal*, it only requires that the SNR of this test matches $\beta_n(K, P, h)$ at the worst-case (K, P, h); however, for a test to be *optimally adaptive*, its SNR has to match $\beta_n(K, P, h)$ for all (K, P, h). We discuss this more carefully in Section 3.3.

Remark 2 *(Other moment statistics)*: In (2.4), we have defined the *order-m orthodox Signed Polygon statistic* $U_n^{(m)}$, for $m \geq 3$. We now define the *length-m Signed Path* statistic $V_n^{(m)}$ by

$$V_n^{(m)} = \sum_{i_1,\ldots,i_{m+1} \text{ (distinct)}} (A_{i_1 i_2} - \hat{\alpha}_n)(A_{i_2 i_3} - \hat{\alpha}_n) \cdots (A_{i_m i_{m+1}} - \hat{\alpha}_n), \qquad \text{for } m \geq 2. \tag{2.10}$$

The two statistics distinguish two hypotheses from estimating $\mathrm{tr}(\widetilde{\Omega}^m)$ and $\mathbf{1}_n'\widetilde{\Omega}^m\mathbf{1}_n$, respectively. The oSQ statistic is $U_n^{(m)}$ with $m = 4$, and the χ^2 statistic is equivalent to $V_n^{(m)}$ with $m = 2$ (see (2.3)). A natural question is whether we should consider other values of m. In Appendix C [8], we show that under some regularity conditions:

$$\begin{aligned}
\mathrm{SNR}(U_n^{(m)}) &\asymp n^m \alpha_0^{-m} \|P - \alpha_0 \mathbf{1}_K \mathbf{1}_K'\|^m \asymp \tau_n^{m/4}, \\
\mathrm{SNR}(V_n^{(m)}) &\asymp n^{\frac{m+1}{2}} \alpha_0^{-\frac{m}{2}} \|P - \alpha_0 \mathbf{1}_K \mathbf{1}_K'\|^{m-2} \|Ph - \alpha_0 \mathbf{1}_K\|^2 \asymp \tau_n^{(m-2)/4} \delta_n.
\end{aligned}$$

Hence, either $\mathrm{SNR}(U_n^{(m)}) \to \infty$ or $\mathrm{SNR}(V_n^{(m)}) \to \infty$ is a stronger requirement than $\max\{\delta_n, \tau_n\} \to \infty$. In other words, we do not benefit from a better phase transition by considering other values of m.

Remark 3 *(The regime of a constant SNR)*: Our phase transition covers the case of SNR $\to 0$ and SNR $\to \infty$. It is also interesting to study the regime of a constant SNR. In the special case of symmetric 2-community SBM (see Example 1), this regime has been well understood. If $n(a_n - b_n)^2/(a_n + b_n) < 1$, the two hypotheses are mutually contiguous; if $n(a_n - b_n)^2/(a_n + b_n) > 1$, the two hypotheses are asymptotically singular, and the signed cycle statistic $U_n^{(m)}$, with $m \asymp \log^{1/4}(n)$, has asymptotically full power [20, 4, 10]. Beyond this special case, much less is known. Some partial results were obtained for general structure of P [6], unequal-size communities [24], and mixed memberships [11], where they were primarily interested in getting a good estimate of Π rather than the global testing problem. It is largely unclear whether there exists a test with asymptotically full power in the constant SNR regime for a general MMSBM. Given the expression of β_n in our phase transition and the full-power test for the special 2-community SBM, we conjecture that a full-power test may be constructed from a weighted combination of $\{U_n^{(m)}, V_n^{(m)}\}_{1\le m\le \log(n)}$, where $U_n^{(m)}$ and $V_n^{(m)}$ are the signed cycle statistics and signed path statistics as in Remark 2. We note that our proposed PE test is a combination of $U_n^{(4)}$ and $V_n^{(2)}$.

Remark 4 *(Which test to use at finite n and with knowledge of parameters)*: Our theoretical results are for the asymptotic setting of $n\to\infty$, but our analysis does give the SNRs of these statistics for finite n. Recall that $\Omega=\Pi P\Pi'$ and $\widetilde{\Omega}=\Pi(P-\alpha_0\mathbf{1}_K\mathbf{1}_K')\Pi'$. Let $[\Omega\circ(1-\Omega)]$ be the matrix whose (i,j)th entry is $\Omega_{ij}(1-\Omega_{ij})$. We can obtain that

$$\mathrm{SNR}(X_n)\sim\frac{\mathbf{1}_n'\widetilde{\Omega}^2\mathbf{1}_n}{\sqrt{2(\mathbf{1}_n'[\Omega\circ(1-\Omega)]^2\mathbf{1}_n)}},\qquad \mathrm{SNR}(Q_n)\sim\frac{\mathrm{tr}(\widetilde{\Omega}^4)}{\sqrt{8\mathrm{tr}([\Omega\circ(1-\Omega)]^4)}}.$$

In principle, for finite n, if the parameters of the alternative hypothesis are given, we can compute these precise SNRs and decide which test to use. However, in practice, we always recommend using the PE test. One advantage of PE is its 'adaptivity': It yields a good power uniformly in many settings, so we do not worry about choosing between χ^2 and oSQ.

Remark 5 *(Comparison with DCMM)*: The DCMM model [13, 15] generalizes MMSBM by accommodating degree heterogeneity. It introduces a degree parameter $\theta_i>0$ for each node i and assumes

$$\Omega_{ij}\equiv\mathbb{P}(A_{ij}=1)=\theta_i\theta_j(\pi_i'P\pi_j),\qquad 1\le i<j\le n.$$

Although MMSBM is a sub-class of DCBM by forcing $\theta_i\equiv 1$, the global testing for these two models is quite different. Consider Example 2 in Section 2.1, where $P=\eta\eta'$. By letting $\tilde{\theta}_i=\theta_i\cdot\pi_i'\eta$, we can also write $\Omega_{ij}=\tilde{\theta}_i\tilde{\theta}_j$ for all $1\le i,j\le n$. Then, it becomes a null model under DCBM, although it is still an alternative model under MMSBM (where the intrinsic number of communities is 2; see Section 4). This example shows that restricting to a sub-class of models can change the detection boundary.

Compared with MMSBM, DCMM has many more free parameters, so 'degree matching' [15] is possible: Given any alternative DCMM, there exists a null DCMM such that for each node, its expected degree under the null model is matched with its expected degree under the alternative model. Then, any degree-based test loses power. In contrast, such 'degree matching' is impossible under MMSBM; and we can find many settings where the (degree-based) χ^2 test has superior power. Hence, to achieve the optimal phase transition, it is crucial to use the χ^2 statistic for 'power enhancement'.

3. Main results

Recall that we consider the global testing problem (1.3) under the MMSBM model, where the Bernoulli probability matrix Ω under two hypotheses are as in (1.4). In Section 3.1, we derive the null distributions of the three test statistics (χ^2, oSQ and PE). In Section 3.2, we study the power of the three tests. We provide lower bound arguments and phase transitions in Section 3.3.

3.1. The asymptotic null distributions

Under the null hypothesis, $\Omega = \alpha_n \mathbf{1}_n \mathbf{1}_n'$, with $\alpha_n \in (0,1)$ calibrating the sparsity level of the network. We estimate α_n by $\hat{\alpha}_n = (n-1)^{-1}\bar{d}$, where $\bar{d}$ is the average node degree. Let X_n and Q_n be the χ^2 statistic and the oSQ statistic in (2.2) and (2.5), respectively. The following theorem characterizes the joint null distribution of (X_n, Q_n).

Theorem 3.1 (Asymptotic joint null distribution)**.** *Consider the global testing problem in* (1.3)*, where* $\Omega = \alpha_n \mathbf{1}_n \mathbf{1}_n'$ *under the null hypothesis. Suppose that* $\alpha_n \leq 1 - c_0$ *for a constant* $c_0 \in (0,1)$ *and that* $n\alpha_n \to \infty$ *as* $n \to \infty$*. Then, under the null hypothesis,*

$$\left(\frac{X_n - n}{\sqrt{2n}}, \frac{Q_n}{2\sqrt{2}n^2\hat{\alpha}_n^2}\right) \quad \xrightarrow[n\to\infty]{\mathcal{L}} \quad \mathcal{N}(0, I_2).$$

The proof of Theorem 3.1 can be found in Appendix D [8]. We immediately obtain the asymptotic distributions of the three test statistics.

Corollary 3.1. *Under the conditions of Theorem 3.1, as* $n \to \infty$*, the following statements are true:*

- *The* χ^2 *test statistic satisfies that* $(X_n - n)/\sqrt{2n} \longrightarrow \mathcal{N}(0,1)$ *in distribution.*
- *The oSQ test statistic satisfies that* $Q_n/(2\sqrt{2}n^2\hat{\alpha}_n^2) \longrightarrow \mathcal{N}(0,1)$ *in distribution.*
- *The PE test statistic satisfies that* $S_n \to \chi_2^2(0)$ *in distribution.*

Fix any $\epsilon \in (0,1)$. The level-ϵ degee-based χ^2 test rejects the null hypothesis if

$$(X_n - n)/\sqrt{n} \quad > \quad (1-\epsilon)\text{-quantile of } \mathcal{N}(0,1). \tag{3.1}$$

The level-ϵ oSQ test rejects the null hypothesis if

$$Q_n/\big(2\sqrt{2}n^2\hat{\alpha}_n^2\big) \quad > \quad (1-\epsilon)\text{-quantile of } \mathcal{N}(0,1). \tag{3.2}$$

The level-ϵ PE test rejects the null hypothesis if

$$S_n \quad > \quad (1-\epsilon)\text{-quantile of } \chi_2^2(0). \tag{3.3}$$

Remark 6: We give a brief explanation of why X_n and Q_n are asymptotically uncorrelated. Let Q_n^* be a proxy of Q_n by replacing $\hat{\alpha}_n$ by α_n in (2.5). Moreover, from (2.3), we can re-write $X_n = n + \frac{1}{n\hat{\alpha}_n(1-\hat{\alpha}_n}\sum_{i,j,k(dist)}(A_{ij} - \hat{\alpha}_n)(A_{jk} - \hat{\alpha}_n)$. Replacing $\hat{\alpha}_n$ by α_n in this expression leads to a proxy of X_n, denoted by X_n^*. Let $W = A - \mathbb{E}[A]$. Under the null hypothesis, $A_{ij} = \alpha_n + W_{ij}$. It follows that

$$X_n^* = n + \frac{1}{n\alpha_n(1-\alpha_n)} \sum_{i,j,k \text{ (dist)}} W_{ij}W_{jk}, \qquad Q_n^* = \sum_{i_1,i_2,i_3,i_4 \text{ (dist)}} W_{i_1i_2}W_{i_2i_3}W_{i_3i_4}W_{i_4i_1}.$$

A key observation is that $\mathbb{E}[(W_{ij}W_{jk})(W_{i_1i_2}W_{i_2i_3}W_{i_3i_4}W_{i_4i_1})] = 0$ for all (i,j,k,i_1,i_2,i_3,i_4) such that (i,j,k) are distinct and (i_1,i_2,i_3,i_4) are distinct. To verify this, it suffices to check all possible cases of $\{i,j,k\} \cap \{i_1,i_2,i_3,i_4\} \neq \emptyset$. For example, when $i = i_1$, $j = i_2$ and $k = i_3$, the expectation is equal to $\mathbb{E}[W_{ij}^2 W_{jk}^2 W_{ki_4} W_{i_4i}] = 0$. Other cases are similar. Therefore, X_n^* and Q_n^* are uncorrelated. Since $X_n \approx X_n^*$ and $Q_n \approx Q_n^*$, we can show that X_n and Q_n are asymptotically uncorrelated.

3.2. Power analysis

Under the alternative hypothesis, $\Omega = \Pi P \Pi'$. We notice that the parameters are not identifiable. There may exist $K^* \neq K$ and (Π^*, P^*) such that $\Omega = \Pi^* P^* (\Pi^*)'$ also holds. To address this issue, we follow Occam's razor to choose the parameters associated with the smallest possible K. This K is called the Intrinsic Number of Communities (INC). [2] In this subsection, we always assume that (K, P, Π) are the parameters associated with INC. The detailed discussion of INC is deferred to Section 4.

Write

$$h = \frac{1}{n}\sum_{i=1}^{n} \pi_i, \qquad G = \frac{1}{n}\sum_{i=1}^{n} \pi_i \pi_i', \qquad \alpha_0 = \alpha_0(h, P) = h' P h.$$

We assume there exists a constant $C > 0$ such that

$$\frac{\max_{1\leq k\leq K} h_k}{\min_{1\leq k\leq K} h_k} \leq C, \qquad \text{and} \qquad \|G^{-1}\| \leq C. \tag{3.4}$$

For a constant $c \in (0,1)$, we assume

$$\alpha_0 \leq c, \qquad \text{and} \qquad n\alpha_0 \geq c^{-1}. \tag{3.5}$$

These conditions are mild. Condition (3.4) is about balance of communities. This is easier to see in the case of no mixed membership (i.e., π_i only takes values in $\{e_1, \ldots, e_K\}$). In this case, G is a diagonal matrix, and both h_k and G_{kk} are equal to the fraction of nodes in community k. Then, (3.4) says that the fraction of nodes in each community is bounded away from zero, which is a mild condition (e.g., [23] used a similar condition). Condition (3.5) is about network sparsity. Under our model, the average node degree is at the order of $n\alpha_0$; therefore, (3.5) allows the average node degree to range from $O(1)$ to $O(n)$, which covers a wide range of sparsity.

First, we study the (degree-based) χ^2 test.

Theorem 3.2 (Power of the χ^2 test)**.** *Consider the global testing problem in* (1.3)*, where* (3.4)-(3.5) *hold under the alternative hypothesis. Let*

$$\delta_n = n^{3/2}\alpha_0^{-1}\|Ph - \alpha_0 \mathbf{1}_K\|^2.$$

Suppose $\delta_n \geq C$. There exist a constant $c_1 > 0$ such that, under the alternative hypothesis,

$$\mathbb{E}\left[\frac{X_n - n}{\sqrt{2n}}\right] \geq c_1\delta_n - O\big(n^{-1/2}\alpha_0^{-1/2}\big),$$

$$\mathrm{Var}\left(\frac{X_n - n}{\sqrt{2n}}\right) = O\Big(1 + n^{-1/2}\delta_n + n^{-2}\alpha_0^{-1}\delta_n^2 \log(n)\Big).$$

The proof of Theorem 3.2 can be found in Appendix E [8]. We are interested in the scenario of $\delta_n \to \infty$. Write $\psi_n^{(1)} = (X_n - n)/\sqrt{2n}$ for short. By Theorem 3.2 and the assumption $n\alpha_n \geq c^{-1}$, we have:

$$\mathbb{E}[\psi_n^{(1)}] \geq C\delta_n, \qquad \mathrm{SD}(\psi_n^{(1)}) \leq \begin{cases} C\big(1 + \delta_n\sqrt{\log(n)/n}\big), & \text{if } \delta_n = O(\sqrt{n}), \\ C\big(n^{-1/4}\sqrt{\delta_n} + \delta_n\sqrt{\log(n)/n}\big), & \text{if } \delta_n \gg \sqrt{n}. \end{cases}$$

[2]We will show in Sectioin 4 that $\mathrm{INC} = 1$ if and only if $\Omega \propto \mathbf{1}_n\mathbf{1}_n'$, which is compatible with the null model in (1.4).

It follows that the SNR is

$$\geq \begin{cases} C\min\{\delta_n,\ \sqrt{n/\log(n)}\}, & \text{if } \delta_n = O(\sqrt{n}), \\ C\min\{n^{1/4}\sqrt{\delta_n},\ \sqrt{n/\log(n)}\}, & \text{if } \delta_n \gg \sqrt{n}. \end{cases}$$

In either case, the SNR tends to ∞. We expect that the χ^2 test successfully separates the null from the alternative. This gives the following corollary:

Corollary 3.2. *Consider a pair of hypotheses as in* (1.3)-(1.4)*, where* $\alpha_n \leq 1 - c_0$ *and* $n\alpha_n \to \infty$ *under the null hypothesis and the conditions* (3.4)-(3.5) *hold under the alternative hypothesis. Suppose* $\delta_n \to \infty$ *under the alternative hypothesis. For any fixed* $\epsilon \in (0,1)$*, consider the level-*ϵ χ^2 *test in* (3.1)*. Then, the level of the test tends to* ϵ *and the power of the test tends to* 1.

The proof of Corollary 3.2 can be found in Appendix G [8]. Next, we study the oSQ test.

Theorem 3.3 (Power of the oSQ test)**.** *Consider the global testing problem in* (1.3)*, where* (3.4)-(3.5) *hold under the alternative hypothesis. Let*

$$\tau_n = n^2\alpha_0^{-2}\|P - \alpha_0 \mathbf{1}_K\mathbf{1}_K'\|^4.$$

Suppose $\tau_n \geq C$*. There exists a constant* $c_2 > 0$ *such that, under the alternative hypothesis,*

$$\mathbb{E}\left[\frac{Q_n}{2\sqrt{2}n^2\hat{\alpha}_n}\right] \geq c_2\tau_n - o\left(1 + n^{-1/2}\tau_n^{3/4}\right),$$

$$\mathrm{Var}\left(\frac{Q_n}{2\sqrt{2}n^2\hat{\alpha}_n^2}\right) = O\Big(1 + n^{-1}\tau_n^{3/2} + n^{-2}\alpha_0^{-1}\tau_n^2\log(n)\Big).$$

The proof of Theorem 3.3 can be found in Appendix F [8]. We are interested in the cases where $\tau_n \to \infty$. Write $\psi_n^{(2)} = Q_n/\big(2\sqrt{2}n^2\hat{\alpha}_n\big)$ for short. By Theorem 3.3 and the assumption $n\alpha_n \geq c^{-1}$, we have:

$$\mathbb{E}[\psi_n^{(2)}] \geq C\tau_n, \qquad \mathrm{SD}(\psi_n^{(2)}) \leq \begin{cases} C\big(1 + \tau_n\sqrt{\log(n)/n}\big), & \text{if } \tau_n = O(n^{2/3}), \\ C\big(n^{-1/2}\tau_n^{3/4} + \tau_n\sqrt{\log(n)/n}\big), & \text{if } \tau_n \gg n^{2/3}. \end{cases}$$

It follows that the SNR is

$$\geq \begin{cases} C\min\{\tau_n,\ \sqrt{n/\log(n)}\}, & \text{if } \tau_n = O(n^{2/3}), \\ C\min\{n^{1/2}\tau_n^{1/4},\ \sqrt{n/\log(n)}\}, & \text{if } \tau_n \gg n^{2/3}. \end{cases}$$

In either case, the SNR tends to ∞. We expect that the oSQ test can successfully separate two hypotheses, as stated in the following corollary:

Corollary 3.3. *Consider a pair of hypotheses as in* (1.3)-(1.4)*, where* $\alpha_n \to 0$ *and* $n\alpha_n \to \infty$ *under the null hypothesis and* (3.4)-(3.5) *hold under the alternative hypothesis. Suppose* $\tau_n \to \infty$ *under the alternative hypothesis. For any fixed* $\epsilon \in (0,1)$*, consider the level-*ϵ *oSQ test in* (3.2)*. Then, as* $n \to \infty$*, the level of the test tends to* ϵ *and the power of the test tends to* 1.

The proof of Corollary 3.3 can be found in Appendix H [8]. Last, we study the PE test, where the test statistic is $S_n = (\psi_n^{(1)})^2 + (\psi_n^{(2)})^2$.

Theorem 3.4 (Power of the PE test). *Consider a pair of hypotheses as in (1.3)-(1.4), where $\alpha_n \to 0$ and $n\alpha_n \to \infty$ under the null hypothesis and (3.4)-(3.5) hold under the alternative hypothesis. Suppose $\max\{\delta_n, \tau_n\} \to \infty$ under the alternative hypothesis. Then, under the alternative hypothesis,*

$$S_n \xrightarrow{\mathbb{P}} \infty.$$

Furthermore, for any fixed $\epsilon \in (0,1)$, consider the level-ϵ PE test in (3.3). Then, as $n \to \infty$, the level of the test tends to ϵ and the power of the test tends to 1.

By Theorem 3.4, the PE test successfully distinguishes two hypotheses as long as $\delta_n^2 + \tau_n^2 \to \infty$ (or equivalently, $\max\{\delta_n, \tau_n\} \to \infty$). The proof of Theorem 3.4 can be found in Appendix I [8].

3.3. The lower bounds and phase transitions

To obtain lower bounds, we switch to the random-membership MMSBM (this follows the convention: If we use a non-random Π, only trivial lower bounds can be obtained, which is uninteresting). Fix $K \geq 2$ and consider a sequence of (P_n, F_n), indexed by n, where $P_n \in \mathbb{R}^{K \times K}$ is an eligible community matrix and F_n is a distribution on the probability simplex of $\mathbb{R}^K$. Let $h_n = \mathbb{E}_{\pi \sim F_n}[\pi] \in \mathbb{R}^K$. We often drop the subscript n in (P_n, F_n, h_n) to simplify the notations. The (randomized) alternative hypothesis is

$$H_1^{(n)}: \qquad \Omega = \Pi P \Pi', \qquad \text{where} \quad \Pi = [\pi_1, \pi_2, \ldots, \pi_n]', \quad \pi_i \overset{iid}{\sim} F. \tag{3.6}$$

We pair this alternative hypothesis with the null hypothesis below:

$$H_0^{(n)}: \qquad \Omega = \alpha_0 \mathbf{1}_n \mathbf{1}_n', \qquad \text{where} \quad \alpha_0 = h' P h. \tag{3.7}$$

Let $f_1(A)$ and $f_0(A)$ be the probability densities associated with (3.6) and (3.7), respectively. The χ^2-distance between two hypotheses is defined as $\int [\frac{f_1(A)}{f_0(A)} - 1]^2 f_0(A) dA$. Two hypotheses are asymptotically indistinguishable if the χ^2-distance $\to 0$. The following theorem is proved in Appendix J [8].

Theorem 3.5 (Lower bound). *Consider a sequence of hypothesis pairs (3.6)-(3.7) indexed by n. Let*

$$\beta_n = \beta_n(K, P_n, h_n) = \max\{n^{3/2}\alpha_0^{-1}\|Ph - \alpha_0 \mathbf{1}_K\|^2, \quad n^2 \alpha_0^{-2} \|P - \alpha_0 \mathbf{1}_K \mathbf{1}_K'\|^4\}.$$

If $\beta_n \to 0$, then the chi-square distance between two hypotheses converges to 0 as $n \to \infty$.

The proof of Theorem 3.5 can be found in Appendix J [8]. We now combine Theorem 3.5 with Theorem 3.4 and obtain the phase transitions:

- *Region of Impossibility*. When $\beta_n = \max\{\delta_n, \tau_n\} \to 0$, by Theorem 3.5, the two hypotheses are asymptotically inseparable, where for any test, the sum of type I and type II errors tends to 1 as $n \to \infty$.
- *Region of Possibility*. When $\beta_n = \max\{\delta_n, \tau_n\} \to \infty$, by Theorem 3.4, the PE test can successfully separate the two hypotheses: for properly chosen $\epsilon_n \to 0$, the sum of type I and type II errors of the level-ϵ_n PE test tends to 0 as $n \to \infty$.

We conclude that the PE test is optimally adaptive. As we have mentioned in Section 1, none of the previously existing tests are optimally adaptive.

Theorem 3.4 is more informative than the standard minimax lower bound. To prove a minimax lower bound, we only need to pick one 'worst-case' configuration of (K, P_n, h_n), but Theorem 3.4 is for all configurations of (K, P_n, h_n). The test that works well for the 'worst-case' configuration may have unsatisfactory performances for other configurations. For example, a commonly studied configuration in the literature is $P_n = (a_n - b_n)I_K + b_n \mathbf{1}_K \mathbf{1}_K'$ and $h_n = (1/K)\mathbf{1}_K$. For this configuration, $\delta_n = 0$, and the oSQ test alone is optimal. However, when (P_n, h_n) deviate from this configuration, the PE test can outperform the oSQ test (as seen in the simulations in Section 5).

Theorem 3.5 can be used to derive the minimax lower bounds for different parameter classes. In standard minimax arguments, we adopt the original MMSBM with a non-random Π, but we will consider the worst case performance for a class of parameters. Let $\{\alpha_n\}_n$ be a positive sequence in $[0, 1]$ and $\{\gamma_n\}_n$ be another positive sequence. Given (K, P, Π), write $h = h(\Pi) = n^{-1}\sum_{i=1}^n \pi_i$ and $\alpha_0 = \alpha_0(P, \Pi) = h'Ph$. We introduce the following classes of Bernoulli probability matrices:

$$\mathcal{M}_{0n}(c_0, \alpha_n) = \big\{\Omega = b\mathbf{1}_n\mathbf{1}_n' :\ b \leq 1 - c_0,\ b \geq \alpha_n\big\},$$

$$\mathcal{M}_{1n}(t_0; K, c, C, \alpha_n, \gamma_n) = \left\{\begin{array}{ll}\Omega = \Pi P \Pi' : & (\Pi, P) \text{ satisfies (3.4)-(3.5) for } c, C > 0, \\ & \alpha_0 \equiv \alpha_0(P, \Pi) \geq \alpha_n/2, \\ & \|Ph(\Pi) - \alpha_0 \mathbf{1}_K\|_F \geq t_0 \|P - \alpha_0 \mathbf{1}_K \mathbf{1}_K'\|, \\ & \text{and } 2\alpha_0^{-1}\|P - \alpha_0 \mathbf{1}_K \mathbf{1}_K'\| \geq \gamma_n \end{array}\right\}.$$

We abbreviate the two classes as M_{0n} and $M_{1n}(t_0)$, respectively. We note that when $t_0 > 0$, an additional constraint on (P, Π) is imposed. Define the minimax testing risk as (below, the infimum is taken over all possible tests ψ)

$$Risk_n^*(t_0) = \inf_{\psi}\Big\{\sup_{\Omega \in \mathcal{M}_{0n}} \mathbb{P}(\psi = 1) + \sup_{\Omega \in \mathcal{M}_{1n}(t_0)} \mathbb{P}(\psi = 0)\Big\}. \tag{3.8}$$

The following theorem is proved in Appendix K [8]:

Theorem 3.6 (Minimax lower bound). *Fix $K \geq 2$. Suppose $\alpha_n \to 0$, $\gamma_n \to 0$ and $n\alpha_n \to \infty$.*

- *Fix $t_0 = 0$. If $n^2\alpha_n^2\gamma_n^4 \to 0$, then* $\lim_{n\to\infty}\{Risk_n^*(t_0)\} = 1$.
- *Fix $0 < t_0 < \sqrt{\frac{(K-1)(K+3)}{16K}}$. If $n^{3/2}\alpha_n\gamma_n^2 \to 0$, then* $\lim_{n\to\infty}\{Risk_n^*(t_0)\} = 1$.

Theorem 3.6 implies that the minimax lower bound changes with the parameter class. When $t_0 = 0$, it is a very broad class, including the symmetric cases in Example 1 of Section 2.2 where the χ^2 test loses power. In this broad class, the minimax lower bound is governed by $n^2\alpha_n^2\gamma_n^4$ (corresponding to the previous τ_n), and the oSQ test is minimax optimal. When $t_0 > 0$, we restrict to a narrower class, with those extremely symmetrical settings excluded. The minimax lower bound is governed by $n^{3/2}\alpha_n\gamma_n^2$ (corresponding to the previous δ_n), and the χ^2-test is minimax optimal. In comparison, for both classes, the PE test is minimax optimal.

4. The identifiability of K

To our best knowledge, the identifiability of MMSBM has not yet been carefully studied. We present our results, which are of independent interest.

***Definition* 4.1.** Fix $n \geq K$. We call $\Pi \in \mathbb{R}^{n\times K}$ an eligible membership matrix if each row is a weight vector and the $K \times K$ identity matrix I_K is a sub-matrix of Π. We call $P \in [0,1]^K$ eligible if it is entry-wise non-negative.

In MMSBM, $\Omega = \Pi P \Pi'$, for some K and eligible (Π, P). However, there may exist $K^* \neq K$ and eligible (Π^*, P^*) such that $\Omega = \Pi^* P^* (\Pi^*)'$ also holds. Below is an example.

Example 3. Let $K = 4$ and $K^* = 2$. Fix any mixed membership matrix $\Pi = [\pi_1, \pi_2, \ldots, \pi_n]' \in \mathbb{R}^{n\times 4}$. Introduce P, P^* and $\Pi^* = [\pi_1^*, \pi_2^*, \ldots, \pi_n^*]' \in \mathbb{R}^{n\times 2}$ as follows:

$$P = 0.01 \times \begin{bmatrix} 1 & 2 & 1.8 & 3 \\ 2 & 4 & 3.6 & 6 \\ 1.8 & 3.6 & 3.24 & 5.4 \\ 3 & 6 & 5.4 & 9 \end{bmatrix}, \quad P^* = 0.01 \times \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix}, \quad \pi_i^* = \begin{bmatrix} 0.5\pi_i(2) + 0.4\pi_i(3) + \pi_i(4) \\ \pi_i(1) + 0.5\pi_i(2) + 0.6\pi_i(3) \end{bmatrix}.$$

We connect (K, Π, P) and (K^*, Π^*, P^*). Let $\eta^* = (0.3, 0.1)'$, $h_1 = (0, 0.5, 0.4, 1)'$, $h_2 = (1, 0.5, 0.6, 0)'$, and $H = [h_1, h_2] \in \mathbb{R}^{4\times 2}$. It is straightforward to see that $P^* = \eta^*(\eta^*)'$ and $\Pi^* = \Pi H$. Furthermore, we can verify that $P = (H\eta^*)(H\eta^*)'$. It follows that

$$\Pi^* P^* (\Pi^*)' = (\Pi H)[\eta^*(\eta^*)'](H'\Pi') = \Pi P \Pi' = \Omega.$$

We can view this example as an MMSBM with $K = 4$ communities or an MMSBM with $K^* = 2$ communities. At the same time, the rank of Ω is $r = 1$. Which of (K, K^*, r) shall we use as the correct definition of 'number of communities'?

To address this issue, we follow Occam's razor to define the number of communities as the smallest K that is compatible with the matrix Ω.

***Definition* 4.2.** The *Intrinsic Number of Communities (INC)* of Ω, denoted as k_Ω, is the smallest integer K such that $\Omega = \Pi P \Pi'$ for some eligible $\Pi \in \mathbb{R}^{n\times K}$ and $P \in \mathbb{R}^{K\times K}$.

Proposition 4.1 (Identifiability of parameters of MMSBM). *Suppose $\Omega = \Pi_0 P_0 \Pi_0'$ for some eligible (Π_0, P_0, K_0) as in Definition 4.1. Recall that the INC k_Ω is as in Definition 4.2.*

- *There exists a pair of eligible $\Pi \in \mathbb{R}^{n,k_\Omega}$ and $P \in \mathbb{R}^{k_\Omega,k_\Omega}$ such that $\Omega = \Pi P \Pi'$.*
- *If there is another pair of eligible $\Pi^* \in \mathbb{R}^{n,k_\Omega}$ and $P^* \in \mathbb{R}^{k_\Omega,k_\Omega}$ such that $\Omega = \Pi^* P^* (\Pi^*)'$, then there must exist a permutation matrix $D \in \mathbb{R}^{k_\Omega,k_\Omega}$ such that $P^* = DPD'$. Therefore, when $K_0 = k_\Omega$, P is identifiable up to permutation.*
- *If, in addition, $\mathrm{rank}(P) = k_\Omega$, then both P and Π are identifiable up to permutation.*
- *It holds that $K_0 \geq k_\Omega \geq \mathrm{rank}(P_0)$. If P_0 is non-singular, then $k_\Omega = \mathrm{rank}(P_0) = K_0$.*

Proposition 4.1 is proved in Appendix L.1 [8]. Throughout this paper, we assume that the K in the alternative hypothesis is the INC defined above.

The definition of INC is natural, but it is not easy to compute. We introduce an alternative formula, which connects k_Ω with the geometry associated with the eigenvectors of Ω. This equivalent definition is much more convenient to use.

Proposition 4.2 (Equivalent definition of INC). *Fix K and $n \geq K$. Suppose $\Omega = \Pi P \Pi'$ for some eligible $\Pi \in \mathbb{R}^{n\times K}$ and $P \in \mathbb{R}^{K\times K}$ as in Definition 4.1. Let $r = \mathrm{rank}(\Omega)$, and let $\xi_1, \xi_2, \ldots, \xi_r \in \mathbb{R}^n$*

be the eigenvectors of Ω associated with nonzero eigenvalues. Write $\Xi = [\xi_1, \xi_2, \ldots, \xi_r]$. Let $C(\Xi) \subset \mathbb{R}^r$ be the convex hull of the n rows of Ξ. Then, $C(\Xi)$ is a polytope and k_Ω is equal to the number of vertices of this polytope.

Proposition 4.2 is proved in Appendix L.2 [8]. We apply Proposition 4.2 to get k_Ω in Example 3. In that example, we have seen that Ω is a rank-1 matrix, $\Omega = (\Pi H\eta^*)(\Pi H\eta^*)'$, implying that $k_\Omega \geq 1$. Additionally, $\Xi = \xi_1 \propto \Pi H\eta^*$, and $C(\Xi)$ is an interval in $\mathbb{R}$ (a simplex with 2 vertices). It follows immediately that $k_\Omega = 2$.

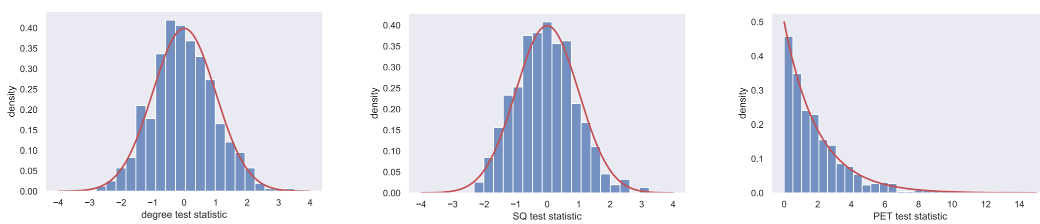
Using Proposition 4.2, we can also easily see that the definition of k_Ω is compatible with the form of Ω for the null hypothesis. If $\Omega = \alpha_n \mathbf{1}_n \mathbf{1}_n'$, then it is obvious that $k_\Omega = 1$. If $k_\Omega = 1$, the last bullet point of Proposition 4.1 implies $\Omega = \lambda_1 \xi_1 \xi_1'$ and $C(\Xi) \subset \mathbb{R}$. By Proposition 4.2, $C(\Xi)$ has to be a singleton, i.e., all the entries of ξ_1 are equal; hence, $\Omega \propto \mathbf{1}_n \mathbf{1}_n'$.

5. Simulations

We conduct numerical experiments to investigate the behavior of the degree-based χ^2 test, the orthodox Signed Quadrilateral (oSQ) test and the newly proposed Power Enhancement (PE) test. Codes used to perform the simulations are available in the supplementary material [8] and in Github[3].

Experiment 1: The asymptotic null distributions. We study how well the asymptotic null distributions in Corollary 3.1 fit the simulated data, for a moderately large n. In these experiments, we generate networks from the Erdös-Rényi model, where $\Omega = \alpha \mathbf{1}_n \mathbf{1}_n'$.

In Experiment 1.1, we fix $n = 200$, $\alpha = 0.1$ and generate 500 networks from the Erdös-Rényi model with $\Omega = \alpha \mathbf{1}_n \mathbf{1}_n'$. The histograms of the three test statistics are shown in Figure 2. They fit the limiting null distributions well.



(a) Degree-based χ^2 test (b) oSQ test (c) PE test

Figure 2: Histograms of the three test statistics under the null hypothesis ($n = 200$). The red curves are the limiting null distributions in Corollary 3.1.

In Experiment 1.2, we focus on the PE test and evaluate the type I error when the target level is set at 5%. Given (n, α), we generate 500 networks, apply the level-5% PE test, and compute the empirical type I error. We let n range in $\{100, 200, 500, 1000\}$ and α range in $\{0.1, 0.2, 0.3, 0.4\}$. The results are shown in Table 1. It suggests that the type I errors are controlled satisfactorily.

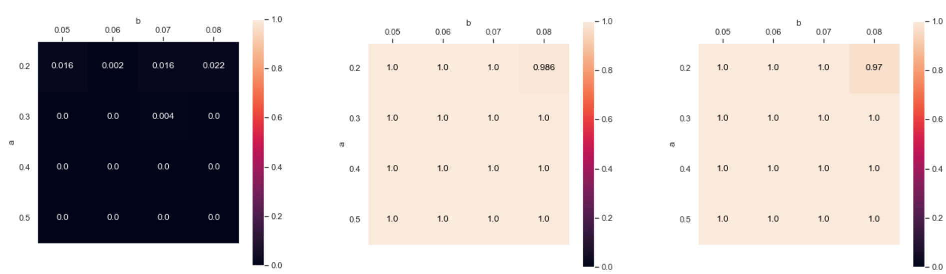
[3] All the simulation codes can be found at: `https://github.com/louiscam/SBM_phase_transition.git`

Table 1. Empirical type I error of the level-5% PE test (calculated based on 500 repetitions).

	$\alpha = 0.1$	$\alpha = 0.2$	$\alpha = 0.3$	$\alpha = 0.4$
$n = 100$	3.2%	3.4%	4%	3.2%
$n = 200$	4%	5.4%	6.2%	3.4%
$n = 500$	5.8%	3.8%	5.2%	4.2%
$n = 1000$	5%	6%	5%	5.6%

Experiment 2: Power comparison of the three tests. We examine the power of the three tests and demonstrate the numerical advantage of PET over χ^2 and oSQ. We will consider two settings, adapted from the examples in Section 2.1. In the first setting, $\delta_n = 0$, hence, only the oSQ test has non-trivial power. In the second setting, $\tau_n \gg \delta_n$, hence, the χ^2 test has a much larger SNR.

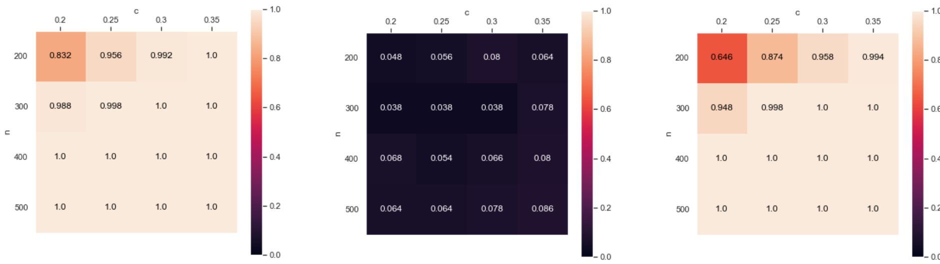
In Experiment 2.1, we let $P = (a-b)I_K + b\mathbf{1}_K\mathbf{1}_K'$, same as in Example 1 of Section 2.1. We let all nodes be pure, with the same number of nodes in each community; this corresponds to $h = K^{-1}\mathbf{1}_K$. By direct calculations, $Ph = \alpha_0\mathbf{1}_K$, with $\alpha_0 = h'Ph = K^{-1}[a + (K-1)b]$. In this setting, $\delta_n = 0$, so that the χ^2 test loses power. We fix $(n, K) = (300, 5)$ and let a range in $\{0.2, 0.3, 0.4, 0.5\}$ and b range in $\{0.05, 0.06, 0.07, 0.08\}$. For each (a, b), we generate 500 networks from the alternative model as above and apply the three tests for a target level 5%; then, we report the proportion of rejections. The results are shown in Figure 3. We observe that the SQ test clearly outperforms the degree-based χ^2 test across these configurations, and that the PE test also benefits from this desirable behavior.



(a) Degree-based χ^2 test (b) SQ test (c) PE test

Figure 3: Empirical power of the three tests in Experiment 2.1 ($n = 300$, $K = 5$). The diagonals of P are equal to a and the off-diagonals are equal to b. Different combinations of (a, b) are considered. In this setting, $\delta_n = 0$ always holds. Therefore, only the oSQ test has a non-trivial power.

In Experiment 2.2, we let $P = c\eta\eta'$, for a vector $\eta \not\propto \mathbf{1}_K$ (similar to Example 2 of Section 2.1). We fix $K = 2$. Let all nodes be pure, with the same number of nodes in each community; hence, $h = K^{-1}\mathbf{1}_K$. We parametrize $\eta = (a/\sqrt{a^2+b^2}, b/\sqrt{a^2+b^2})'$, where $b = 1$ and $a = 1 + n^{-1/4}$. By direct calculations, $\delta_n \sim cn/4$ and $\tau_n \sim c^2 n/4$. We then let n range in $\{200, 300, 400, 500\}$ and c range in $\{0.2, 0.25, 0.3, 0.35\}$. For these values of c, the SNR of the χ^2 test is considerably larger than that of the oSQ test. Similarly as in Experiment 2.1, we set the target level at 5% and calculate the empirical power based on 500 repetitions. The results are shown in Figure 4. We observe that the degree-based χ^2 test clearly outperforms the SQ test across these configurations, and that the PE test also benefits from this desirable behavior.

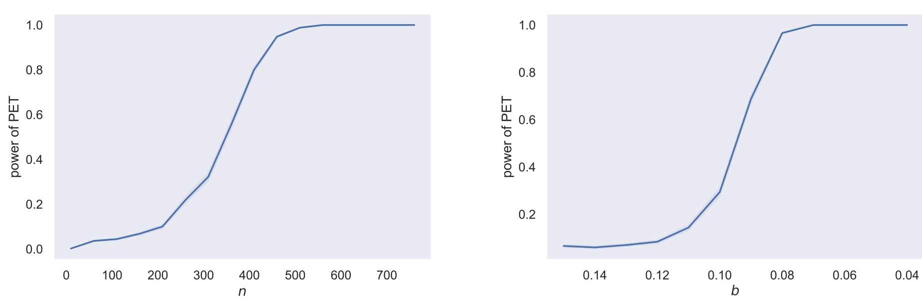


(a) Degree-based χ^2 test (b) SQ test (c) PE test

Figure 4: Empirical power of the three tests in Experiment 2.2 ($K = 2$), where $P = c\eta\eta'$, for a vector $\eta \not\propto \mathbf{1}_K$. Different combinations of (n, c) are considered. The vector η is chosen such that δ_n is always much larger than τ_n. Therefore, the χ^2 test has a much higher SNR.

Experiment 3: Phase transitions for PE. We focus on the PE test and examine its power when the SNR $\beta_n = \max\{\delta_n, \tau_n\}$ gradually increases. This reveals the phase transitions associated with PE.

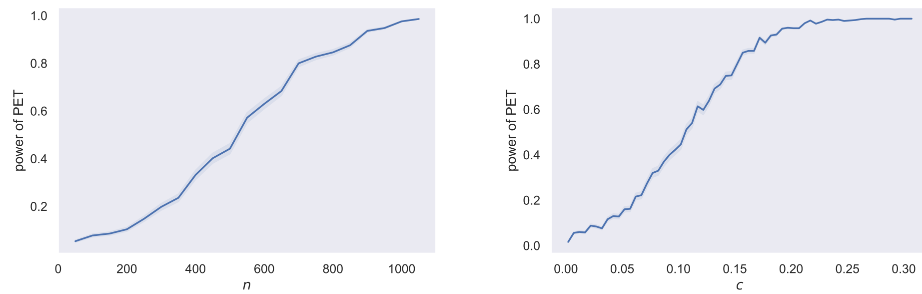
In Experiment 3.1, we use the same model as in Experiment 2.1, where $K = 5$ and $P = (a-b)I_K + b\mathbf{1}_K\mathbf{1}_K'$. By direct calculations, $\beta_n = n^2K^2[a + (K-1)b]^{-2}(a-b)^4$. We fix $a = 0.2$. Then, β_n is monotone increasing with n and monotone decreasing with b. In Experiment 3.1(a), we fix $b = 0.1$ and let n vary from 10 to 760 with a step size of 50. In Experiment 3.1(b), we fix $n = 300$ and let b vary from 0.04 to 0.15 with a step size of 0.01. We report simulation results in Figure 5, where power estimates for each configuration are obtained by averaging the number of rejections over 500 repetitions. The phase transition is visible as we move from vanishing power at low SNR to full power at high SNR.



(a) $b = 0.1$, n varies (b) $n = 300$, b varies

Figure 5: Empirical power of the PE test in Experiment 3.1. In this setting, β_n increases as n increases or b decreases.

In Experiment 3.2, we use the same model as in Experiment 2.2, where $K = 2$ and $P = c\eta\eta'$, with $\eta = (a, b)'/\sqrt{a^2 + b^2}$. We fix $b = 1$ and $a = 1 + n^{1/4}$. Then, $\beta_n \sim \max\{cn/4,\, c^2n/4\}$, which increases with both n and c. In Experiment 3.2(a), we fix $c = 0.06$ and let n vary from 50 to 1100 with a step size of 50. In Experiment 3.2(b), we fix $n = 300$ and let c vary from 0.002 to 0.3 with a step size of 0.005. We report simulation results in Figure 6. It also reveals the phase transition, from the vanishing power at low SNR to full power at high SNR.



(a) $c = 0.06$, n varies

(b) $n = 300$, c varies

Figure 6: Empirical power of the PE test in Experiment 3.2. In this setting, β_n increases as n increases or c increases.

Experiment 4: Comparison with other testing ideas. Other common ideas of global testing include the eigenvalue-based tests and the likelihood-ratio tests. For eigenvalue-based tests, we consider the one in Lei [18]. The test statistic is a function of the largest and smallest eigenvalues of $A - \hat{\alpha}_n \mathbf{1}_n \mathbf{1}_n$. [18] showed that the test statistic converges to a Tracy-Widom distribution under the null hypothesis. We use this null distribution to set the rejection region. The use of likelihood-ratio tests has been limited to SBM (i.e., there is no mixed membership) and requires information on the unknown K in the alternative hypothesis. We instead applied the model selection approach in Bickel and Wang [21], which obtains $\hat{K}$ by successively computing the likelihood ratio between K and $(K+1)$, for $K = 1, 2, \ldots$. We reject the null hypothesis if $\hat{K} > 1$. In this approach, computing the likelihood ratios involves a sum over all possible community labels, and we followed [21] to use the EM algorithm with an initialization by spectral clustering. More details on our implementation of these methods are in our GitHub repository.

In Experiment 4.1, we study SBM with $K = 2$. We consider three models in the alternative hypothesis: (i) The symmetric SBM: $P = (a-b)I_2 + b\mathbf{1}_2\mathbf{1}_2'$, with $a = 0.2$ and $b = 0.05$; the two communities have equal size. (ii) The asymmetric SBM: $P = (a-b)I_2 + b\mathbf{1}_2\mathbf{1}_2'$, with $a = 0.2$ and b drawn from $\mathrm{Uniform}[0.125, 0.175]$; π_i's are i.i.d. drawn from $\mathrm{Multinomial}(1, (0.2, 0.8)')$. (iii) The rank-1 SBM: $P = \eta\eta'$, where $\eta = (a/\sqrt{a^2+b^2}, b/\sqrt{a^2+b^2})'$, $b = 1$ and $a = 1 + n^{-1/2}$; the two communities have equal size. Additionally, we consider the Erdös-Rényi model $\Omega = \alpha_0 \mathbf{1}_n \mathbf{1}_n'$, with $\alpha_0 = 0.2$, as the null hypothesis. The χ^2, oSQ, PE and eigenvalue tests have known null distributions, and we set the rejection region by controlling the level at 5%. For the likelihood ratio test, as mentioned, we reject the null hypothesis if $\hat{K} > 1$. For each model, we fix $n = 500$, generate 100 networks, and measure the power of each test by the fraction of rejections over these 100 repetitions. In Experiment 4.2, we extend Models (i)-(iii) from SBM to MMSBM. For each model, P is the same as before, except that $a = 1 + n^{-1/5}$ in Model (iii); π_i's are i.i.d. generated from Dirichlet$(0.1, 0.1)$ in Model (i), Dirichlet$(0.2, 0.8)$ in Model (ii), and Dirichlet$(0.4, 0.6)$ in Model (iii). The results are in Table 2.

Table 2. Comparison with an eigenvalue-based test and a likelihood-ratio test. For each test, we report the empirical power over 100 repetitions. The settings are described in Experiments 4.1-4.2.

Test	Erdös-Rényi	SBM			MMSBM		
		Symmetric	Asymmetric	Rank-1	Symmetric	Asymmetric	Rank-1
χ^2	0.04	0	0.96	0.95	0.09	0.87	0.99
oSQ	0.05	1	0.33	0.04	1	0.06	0.03
PE	0.06	1	0.92	0.88	1	0.76	0.98
Eigenvalue [18]	0.06	1	0.56	0.02	1	0.10	0.31
Likelihood [21]	0.43	1	0.53	0.48	0.53	0.59	0.49

The 'Symmetric' and 'Asymmetric' models correspond to Case (S) and Case (AS1) in Example 1 of Section 2.1, and the 'Rank-1' models correspond to Example 2. In Section 2.1, the SNRs of χ^2, oSQ and PE have been analyzed, and their empirical powers here agree with the theoretical results. We now focus on comparing the eigenvalue test and the likelihood ratio test with the PE test. In all six models for the alternative hypothesis, the PE test outperforms the eigenvalue test and the likelihood ratio test. The eigenvalue test has a full power in symmetric SBM and symmetric MMSBM, but its performance is unsatisfactory in the other models. In fact, using the results in [18], we can derive that the SNR of the eigenvalue test is $\asymp \tau_n^{1/4}$; in comparison, the SNR of the PE test is $\max\{\delta_n, \tau_n\}$. Therefore, when $\delta_n \to \infty$ but $\tau_n \to 0$, the PE test has asymptotically full power but the eigenvalue test loses power (for example, in the rank-1 SBM model, $\delta_n \asymp n^{3/2}(a-b)^2 \asymp n^{1/2}$, but $\tau_n^{1/4} \asymp n^{1/2}|a-b| \asymp 1 \ll \delta_n$). The likelihood ratio test has better power than the eigenvalue test in the asymmetric and rank-1 models, but worse in the symmetric models. The likelihood ratio test also uniformly underperforms the PE test. For the SBM settings, the likelihood ratio test is supposed to have the best power, provided that K is given and the likelihood is precisely computed. However, these requirements are practically infeasible. We had to use the model selection criteria [21] to avoid specifying K and to compute the likelihood approximately, so its numerical performance should be inferior to the precise likelihood ratio test. For the MMSBM settings, the likelihood ratio is misspecified, which explains the unsatisfactory numerical performance. In terms of computing time, PE is also the fastest, especially for large n.

6. Discussion

We consider the global testing problem for MMSBM. First, we study the (degee-based) χ^2 test and the oSQ test. These two tests existed in the literature, but their performances under MMSBM had never been studied. We derive their asymptotic null distributions and characterize their powers under the alternative. We discover that, for some parameter regimes, the χ^2 test has a better performance; for some other parameter regimes, the oSQ test has a better performance. It motivates us to combine the strengths of both tests. Next, we propose the Power Enhancement (PE) test. We show that the PE test has a tractable null distribution and outperforms both the χ^2 test and the oSQ test. Last, we study the phase transitions in global testing: We identify a quantity $\beta_n(K,P,h)$, such that two hypotheses are asymptotically inseparable if $\beta_n(K,P,h) \to 0$, and perfectly separable by the PE test if $\beta_n(n,P,h) \to \infty$. This holds for arbitrary (K,P,h) that satisfy mild regularity conditions. Therefore, the PE test is optimally adaptive.

Most existing works on global testing focused on a symmetric SBM [20, 4], which corresponds to a special choice of (K,P,h) in our setting. The optimal test (e.g., the oSQ test or an eigenvalue-based test) for this special case may have unsatisfactory power for other choices of (K,P,h). This motivates our study of phase transitions and optimal adaptivity, where we seek to understand the statistical limits for arbitrary (K, P, h) and find a test that is both optimal and adaptive.

In the hypothesis testing literature, it is not uncommon to combine multiple tests to attain the optimal detection boundary across the whole parameter range. For example, [2] combines the χ^2 test with a scan test for optimal detection of a planted clique, and [17] combines a simple aggregation test and a sparse aggregation test for optimal global testing in a clustering model. However, these are Bonferroni combinations, i.e., the combined test rejects the null hypothesis if any of the tests rejects. The simple Bonferroni combination does not support *p*-value calculation. In contrast, our power enhancement test is based on the joint asymptotic null distribution of two test statistics. As a result, the PE test has a tractable null distribution and enables the calculation of a *p*-value.

In [22], the authors studied the global testing problem in hypergraphs under different sparsity settings. They introduced novel powerful statistical tests in the bounded degree regime and the dense regime. Potential avenues for future research include extending our power enhancement framework to the global testing problem in mixed-membership hypergraphs.

Supplementary Material

Supplementary Proofs and Simulation Code
The supplementary material contains the proofs of all main theorems and lemmas and the Python codes used in the simulation studies.

References

[1] AIROLDI, E., BLEI, D., FIENBERG, S. and XING, E. (2008). Mixed membership stochastic blockmodels. *J. Mach. Learn. Res.* **9** 1981–2014.

[2] ARIAS-CASTRO, E. and VERZELEN, N. (2014). Community detection in dense random networks. *Ann. Statist.* **42** 940–969.

[3] BANERJEE, D. (2018). Contiguity and non-reconstruction results for planted partition models: the dense case. *Electron. J. Probab.* **23** 485–512.

[4] BANERJEE, D. and MA, Z. (2017). Optimal hypothesis testing for stochastic block models with growing degrees. *arXiv:1705.05305*.

[5] BANERJEE, D. and MA, Z. (2022). Asymptotic normality and analysis of variance of log-likelihood ratios in spiked random matrix models. *Ann. Statist. (to appear)*.

[6] BANKS, J., MOORE, C., NEEMAN, J. and NETRAPALLI, P. (2016). Information-theoretic thresholds for community detection in sparse networks. In *Conference on Learning Theory* 383–416. PMLR.

[7] BUBECK, S., DING, J., ELDAN, R. and RÁCZ, M. Z. (2016). Testing for high-dimensional geometry in random graphs. *Random Struct. Algor.* **49** 503–532.

[8] CAMMARATA, L. and KE, Z. T. (2022). Supplementary material for “Power enhancement and phase transitions for global testing of the mixed membership stochastic block model".

[9] GAO, C. and LAFFERTY, J. (2017). Testing for global network structure using small subgraph statistics. *arXiv:1710.00862*.

[10] GAO, C. and MA, Z. (2021). Minimax rates in network analysis: Graphon estimation, community detection and hypothesis testing. *Statist. Sci.* **36** 16–33.

[11] HOPKINS, S. B. and STEURER, D. (2017). Efficient Bayesian estimation from few samples: community detection and related problems. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)* 379–390. IEEE.

[12] JI, P., JIN, J., KE, Z. T. and LI, W. (2022). Co-citation and co-authorship networks of statisticians. *J. Bus. Econom. Statist.* **40** 469–485.
[13] JIN, J., KE, Z. T. and LUO, S. (2017). Estimating network memberships by simplex vertex hunting. *arXiv:1708.07852*.
[14] JIN, J., KE, Z. T. and LUO, S. (2018). Network global testing by counting graphlets. In *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018* 2338–2346.
[15] JIN, J., KE, Z. T. and LUO, S. (2021). Optimal adaptivity of signed-polygon statistics for network testing. *Ann. Statist.* **49** 3408–3433.
[16] JIN, J., KE, Z. T., LUO, S. and WANG, M. (2022). Optimal estimation of the number of network communities. *J. Amer. Statist. Assoc.* 1–16.
[17] JIN, J., KE, Z. T. and WANG, W. (2017). Phase transitions for high dimensional clustering and related problems. *Ann. Statist.* **45** 2151–2189.
[18] LEI, J. (2016). A goodness-of-fit test for stochastic block models. *Ann. Statist.* **44** 401–424.
[19] LU, C. and SEN, S. (2020). Contextual stochastic block model: Sharp thresholds and contiguity. *arXiv:2011.09841*.
[20] MOSSEL, E., NEEMAN, J. and SLY, A. (2015). Reconstruction and estimation in the planted partition model. *Probab. Theory Relat. Fields* **162** 431–461.
[21] WANG, Y. R. and BICKEL, P. J. (2017). Likelihood-based model selection for stochastic block models. *Ann. Statist.* **45** 500–528.
[22] YUAN, M., LIU, R., FENG, Y. and SHANG, Z. (2022). Testing community structure for hypergraphs. *Ann. Statist.* **50** 147–169.
[23] ZHANG, A. Y. and ZHOU, H. H. (2016). Minimax rates of community detection in stochastic block models. *Ann. Statist.* **44** 2252–2280.
[24] ZHANG, P., MOORE, C. and NEWMAN, M. (2016). Community detection in networks with unequal groups. *Phys. Rev. E* **93** 012303.