ABSTRACT  Neuroimaging techniques have been routinely applied in various studies in neuroscience, which contribute to providing novel insights into brain functions. One of the most important and challenging questions related to data collected from such studies is hypothesis testing for the differences between two samples of networks of brain regions. This is due to the fact that networks constructed from neuroimaging studies, which can be weighted and large, are very complex. Focusing on this problem, a novel hypothesis testing procedure is proposed under a general framework for large weighted networks. The asymptotic null distribution is derived and the power guarantee is also provided theoretically. Simulation experiments and practical application to the real brain networks are carried out to demonstrate the effectiveness of the proposed testing method.

INDEX TERMS  Hypothesis testing, neuroimaging, network data, weighted network, Tracy–Widom distribution.

I. INTRODUCTION
The past decade has witnessed the development of different techniques for analyzing functional neuroimaging data, especially in the study of human brain’s structural and functional systems [1]–[3]. Particularly, there is a strong need to develop neuroimaging-based models to improve our understanding of a diverse span of topics such as cognitive science and mental functions. Recent views in neuroscience claim a variety of implementations via brain networks which comprise different brain regions [4]–[8]. In this setting, for each individual such as a patient, a network is formed by using brain regions of interest (ROIs) as nodes, which are shared across all networks and predefined anatomically. A measure of association between two nodes is represented by the edge between the corresponding ROIs. The way to define this connectivity varies widely with imaging collecting methods. For example, in functional magnetic resonance imaging (fMRI), it is a functional statistical association between each pair of ROIs [9].

The recent advancement in technology and computer prowess has led to a number of important research initiatives in neuroscience, such as Northwestern University Schizophrenia Data and Software Tool (NUSDAST; available at http://schizconnect.org/), International Neuroimaging Data-sharing Initiative (INDI; available at http://fcon_1000.projects.nitrc.org/) [10], [11], and the Child and Adolescent NeuroDevelopment Initiative (CANDI, available at https://www.nitrc.org/). These initiatives make available fMRI datasets collected from around the globe and provide us with the opportunity to perform analysis on many biological studies.

Statistical tasks with brain networks constructed from functional neuroimaging data have received a lot of attentions, such as pattern-based classification, disease states prediction, and performance qualification [4], [12]–[14]. One of the most pressing and important practical problems is hypothesis testing for differences in networks. It covers a
applied to both weighted and binary large-scale network scenarios.

(2) Asymptotic null distribution and power are obtained theoretically. The new statistic is proved to have asymptotic Tracy–Widom distribution under the null hypothesis and full power against alternative models as the nodes number tends to infinity.

(3) The performance of the proposed test is assessed by comparing to the current state-of-the-art methods in simulations and application to real human brain data.

This paper is organized as follows. In Section II, we propose a testing statistic and analyze its asymptotic properties. Simulation studies are presented in Section III. In Section IV, we illustrate our proposed method with a brain connectivity dataset of SCZs and HC. Section V gives some conclusions.

A. NOTATIONS

Throughout this paper, we use lightface letters, boldface lowercase letters, boldface uppercase letters and lightface uppercase letters to denote scalars, vectors, matrices and distributions respectively.

The notation $\mathbb{R}^n$ represents the set of all $n$-dimensional real column vectors. The symbol $\| \cdot \|$ is the Euclidean norm for vector and spectral norm for matrix. For an $n \times n$ matrix $A$, $A_{ij}$ denotes its $(i, j)$ entry. For an $n \times n$ symmetric matrix $A$, $\lambda_i(A)$ denotes its $i$th largest (real) eigenvalue, ordered as $\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A)$, $s_1(A)$ is the largest singular value. For two matrices $A$ and $B$, $A \circ B$ denotes the Hadamard product of $A$ and $B$.

The operations $\mathbb{P}\{ \cdot \}$ and $\mathbb{E}\{ \cdot \}$ denote the probability of some random event and mathematical expectation of random variable respectively. The notation $\text{TW}(1)$ represents Tracy–Widom distribution with index 1. Write $x_n \sim x$ if a sequence of random variables $\{x_n\}$ converges in distribution to random variable $x$ as $n$ tends to infinity.

For two sequences of real numbers $\{a_n\}$ and $\{b_n\}$, $b_n = O_p(a_n)$ means that there exists a constant $c$ such that $b_n/a_n \rightarrow c$ as $n \rightarrow \infty$; $b_n = o_p(a_n)$ means $b_n/a_n \rightarrow 0$ as $n \rightarrow \infty$; $b_n = O_p(a_n)$ means for any $\epsilon > 0$, there exist finite $c > 0$ and $n_0 > 0$ such that $\mathbb{P}\{|b_n/a_n| > c\} < \epsilon$ for any $n > n_0$; and $b_n = o_p(a_n)$ means $\mathbb{P}\{|b_n/a_n| \geq \epsilon\} \rightarrow 0$ as $n \rightarrow \infty$ for any positive $\epsilon$.

II. TWO-SAMPLE TEST FOR LARGE WEIGHTED NETWORKS

A. PROBLEM FORMULATION

Let $F_1 = \{F_{1,i,j}\}_{1 \leq i < j \leq n}$ and $F_2 = \{F_{2,i,j}\}_{1 \leq i < j \leq n}$ be two sequences of distributions. For two samples of weighted networks with $n$ nodes, we assume that $A^{(1)}_1, \ldots, A^{(m)}_1$ are symmetric adjacency matrices independently generated from $\{F_{1,i,j}\}_{1 \leq i < j \leq n}$. Specifically, $A^{(k)}_{1,i,j} \sim F_{1,i,j}$, $k_1 = 1, 2, \ldots, m_1$, $1 \leq i < j \leq n$. For another sample of symmetric adjacency matrices $A^{(1)}_2, \ldots, A^{(m_2)}_2$, we have $A^{(k)}_{2,i,j} \sim F_{2,i,j}$, $k_2 = 1, 2, \ldots, m_2$, $1 \leq i < j \leq n$. In this paper, $\{F_{1,i,j}\}_{1 \leq i < j \leq n}$ and $\{F_{2,i,j}\}_{1 \leq i < j \leq n}$ are assumed.

variety of interesting applications including for instance, the differences in brain structures between networks of males and females [15], and topological changes in gene regulatory networks coming from two kinds of treatments of breast cancer [16]. However, hypothesis testing for populations of networks remains largely understudied especially for large networks (having large number of nodes), which is an increasingly common trend [17]–[20].

In this paper, motivated by fMRI data from INDI, we investigate the differences between two populations of interest, e.g., schizophrenics (SCZs) and healthy controls (HCs). To attain this goal, we need to develop a statistical test of hypothesis for two samples of networks. The main challenge is that networks are non-traditional forms and are not Euclidean objects, which makes extension of traditional testing methods hard. Focusing on this problem, [15], [21]–[23] have carried out some new hypothesis tests. Nevertheless, these methods are restricted to binary or random dot product networks, which are not suitable for many situations in practice since many real networks are weighted [15], [24]–[26]. Moreover, most studies above are limited to the assumption of fixed nodes number and large sample size except [22]. For example, the statistic in [15] is constructed by using half-vectorization of the graph Laplacian matrices and the inverse of the corresponding pooled covariance matrix, which is not an easy task especially when the nodes number $n$ is large. This can be seen from a case of $n = 200$, the dimension of the half-vectorization based covariance matrix would be as large as 20100, thus posing challenges for estimating this pooled covariance matrix and its inverse.

The main purpose of this paper is to analyze the differences between networks from neuroimaging data, which therefore can provide diagnostic suggestions. To address this issue, a hypothesis test for two samples of weighted and large networks is developed. Specifically, we expect to investigate asymptotic null distribution and asymptotic power under the alternative theoretically, and further evaluate the proposed method via simulation studies and real functional neuroimaging data studies. We first propose a novel testing statistic based on the singular value. Then, we prove that this statistic asymptotically converges to Tracy–Widom distribution under the null hypothesis and the asymptotic power tends to 1 under the alternative hypothesis.

The main advantage of our approach is that our proposed test adopts a very general framework, allowing the connectivity to be weighted and the number of the nodes to grow to infinity, instead of requiring binary networks or fixed nodes number like in most of the existing methods. Moreover, although our main application in this paper is two-sample test of brain connectivity networks, our method is applicable to any other weighted or unweighted binary networks.

The main contributions of this work can be summarized as follows:

(1) A novel two-sample test for populations of networks is developed. Due to the generality of the connectivity structure on the networks, the proposed test can be...
to be defined on bounded intervals and decided by some parameters. We consider networks that are undirected and with no self-loops, hence $A^{(k_1)}_{i,i} = A^{(k_2)}_{i,i} = 0$ for all $k_1 = 1, 2, \ldots, m_1$, $k_2 = 1, 2, \ldots, m_2$, $i = 1, \ldots, n$. We formulate the problem of distinguishing two groups of weighted networks as the following two-sample hypothesis testing:

$$H_0 : F_1 = F_2 \iff H_1 : F_1 \neq F_2. \quad (1)$$

### B. SINGULAR VALUE BASED TEST

For any $k_1 = 1, 2, \ldots, m_1$, $k_2 = 1, 2, \ldots, m_2$, $i,j = 1, 2, \ldots, n$, and $i \neq j$, denote $\sigma_{i,j} > 0$ and $\sigma_{2,i,j} > 0$ as the standard deviations of $A^{(k_1)}_{i,j}$ and $A^{(k_2)}_{i,j}$ respectively. In addition, $\hat{\sigma}_{i,j} > 0$ and $\hat{\sigma}_{2,i,j} > 0$ are some estimators of $\sigma_{i,j}$ and $\sigma_{2,i,j}$ respectively, $i \neq j$. We first introduce a centralized and re-scaled symmetric matrix $V$ with zeros on diagonal and entries for $1 \leq i \leq j \leq n$ as

$$V_{i,j} = \sqrt{(n-1)} \left( \frac{\tilde{A}_{i,j} - \tilde{A}_{2,i,j}}{\sqrt{(n-1)} \left( \frac{\sigma_{i,j}^2}{m_1} + \frac{{\sigma_{2,i,j}}^2}{m_2} \right)} \right), \quad (2)$$

where $\tilde{A}_{i,j} = \frac{1}{m_u} \sum_{u=1}^{m_u} A^{(k_u)}_{u,i,j}$, $u = 1, 2$.

We propose the following singular value based test statistic:

$$T_{TW} = n^{2/3} (s_1(V) - 2). \quad (3)$$

Moreover, the rejection rule for testing problem (1) under significance of $\alpha \in (0, 1)$ is

Reject $H_0$ if $T_{TW} \geq \tau_{\alpha/2}$, \quad (4)

where $\tau_{\alpha/2}$ is the $\alpha/2$ upper quantile of $TW(1)$.

Define a symmetric $n \times n$ matrix $V_0$ with zero diagonal entries and other elements for $1 \leq i < j \leq n$ as

$$V_{0,i,j} = \sqrt{(n-1)} \left( \frac{\tilde{A}_{i,j} - \tilde{A}_{2,i,j}}{\sqrt{(n-1)} \left( \frac{\sigma_{i,j}^2}{m_1} + \frac{{\sigma_{2,i,j}}^2}{m_2} \right)} \right). \quad (5)$$

Extending the proof of Theorem 3 in [22], we have the following lemma.

**Lemma 1:** For $V_0$ defined by (5), we have

$$n^{2/3} (\lambda_1(V_0) - 2) \sim TW(1), \quad (6)$$
$$n^{2/3} (-\lambda_n(V_0) - 2) \sim TW(1). \quad (7)$$

We now have the following theorem on the asymptotic null distribution of the test.

**Theorem 1:** For any $u = 1, 2$, and $i,j = 1, 2, \ldots, n$, if max $|\tilde{\sigma}_{u,i,j} - \sigma_{u,i,j}| = o_p(n^{-7/6})$, $\sigma_{u,i,j} = O_p(1)$, then under the null hypothesis $H_0$ in (1), we have

$$n^{2/3} (\lambda_1(V) - 2) \sim TW(1), \quad (8)$$
$$n^{2/3} (-\lambda_n(V) - 2) \sim TW(1). \quad (9)$$

**Proof:** Under the null hypothesis, $\sigma_{1,i,j} = \sigma_{2,i,j}$, and it is observed that

$$V_{i,j} = \sqrt{\frac{\sigma_{1,i,j}^2}{m_1} + \frac{\sigma_{2,i,j}^2}{m_2}} V_{0,i,j}. \quad (10)$$

Without loss of generality, assume $\tilde{\sigma}_{1,i,j} \leq \tilde{\sigma}_{2,i,j}$. According to the condition on the errors of standard deviations, we have the following for the numerator in (10),

$$\sqrt{\frac{\sigma_{1,i,j}^2}{m_1} + \frac{\sigma_{2,i,j}^2}{m_2}} \leq \sqrt{\frac{\sigma_{1,i,j}^2}{m_1} + \frac{\sigma_{2,i,j}^2}{m_2}} + \sqrt{\frac{m_1 + m_2}{m_1 m_2}} O_p(n^{-2/3}), \quad (11)$$

where the last inequality comes from as follows. Obviously, there exist a constant $c > 0$ and $n_0 > 0$ such that for any $n > n_0$, $O_p(|\tilde{\sigma}_{1,i,j} - \tilde{\sigma}_{2,i,j}|/(\sigma_{1,i,j} - \tilde{\sigma}_{1,i,j}) \leq c$. Then applying the condition max $|\tilde{\sigma}_{1,i,j} - \tilde{\sigma}_{1,i,j}| = o_p(n^{-7/6})$, for any $\varepsilon > 0$ and $n > n_0$, we have

$$P \left\{ \frac{O_p(|\tilde{\sigma}_{1,i,j} - \tilde{\sigma}_{1,i,j}|)}{n^{-7/6}} \right. \left. \leq \varepsilon \right\} \rightarrow 0, \quad (12)$$

which implies $O_p(|\tilde{\sigma}_{1,i,j} - \tilde{\sigma}_{1,i,j}|) \leq o_p(n^{-7/6})$.

Similarly, we have

$$\sqrt{\frac{\sigma_{1,i,j}^2}{m_1} + \frac{\sigma_{2,i,j}^2}{m_2}} \leq \sqrt{\frac{\sigma_{1,i,j}^2}{m_1} + \frac{\sigma_{2,i,j}^2}{m_2}} + \sqrt{\frac{m_1 + m_2}{m_1 m_2}} O_p(n^{-2/3}). \quad (13)$$

From (11) and (13),

$$\sqrt{\frac{\sigma_{1,i,j}^2}{m_1} + \frac{\sigma_{2,i,j}^2}{m_2}} \leq \sqrt{\frac{\sigma_{1,i,j}^2}{m_1} + \frac{\sigma_{2,i,j}^2}{m_2}} + \sqrt{\frac{m_1 + m_2}{m_1 m_2}} o_p(n^{-2/3}). \quad (14)$$

Combining (14) with (10), we obtain

$$V - V_0 = M \circ V_0, \quad (15)$$

$V - V_0 = M \circ V_0$, \quad (15)
where $M$ is an $n \times n$ matrix with entries $M_{ij} = o_p(n^{-7/6})$. In addition, from [27], $\|V_0\| = O_p(1)$. Then, from the definition of spectral norm,

$$
\|V - V_0\| = \max_{x \in \mathbb{R}^n} \|(M \circ V_0)x\|
= \max_{x \in \mathbb{R}^n} \left\| \sum_{i=1}^{n} \left( \sum_{j=1}^{n} M_{ij} V_{0,i} x_j \right)^2 \right\|
\leq \max_{x \in \mathbb{R}^n} \left\| \sum_{i=1}^{n} \left( \sum_{j=1}^{n} M_{ij}^2 \right) \left( \sum_{j=1}^{n} V_{0,i}^2 x_j^2 \right) \right\|
\leq \max_{i,j} |M_{ij}| \sqrt{n} \|V_0\| \max_{i,j} |M_{ij}|.
\tag{16}
$$

Therefore,

$$
|\lambda_1(V) - \lambda_1(V_0)| \leq \|V - V_0\| \leq o_p(n^{-2/3}).
\tag{17}
$$

From (17) and Lemma 1, we immediately have (8). The asymptotic distribution given by (9) is similar. □

Note that Theorem 1 is rather general and has great practical value in real applications with the advantage of imposing almost no restrictions on the network structure (undirected and with no self-loops). Moreover, the order of the error condition in the theorem seems to be a little strict, but there still exist some appropriate estimators of $\sigma_{u,i,j}$ and we will discuss some standard estimators and compare their properties in the later discussions.

Next, we analyze the size and power for the proposed rejection rule (4).

From

$$
P(T_{TW} \geq \tau_{a/2}) \leq P\left\{ n^{2/3} (\lambda_1(V) - 2) \geq \tau_{a/2} \right\}
+ P\left\{ n^{2/3} (-\lambda_n(V) - 2) \geq \tau_{a/2} \right\}
= \alpha + o_n(1),
\tag{18}
$$

we can conclude that the rejection rule given by (4) has size $\alpha$.

Define a symmetric $n \times n$ matrix $\tilde{V}$ with diagonal elements 0 and for any $1 \leq i < j \leq n$,

$$
\tilde{V}_{ij} = \frac{\mu_{1,i,j} - \mu_{2,i,j}}{\sqrt{(n - 1) \left( \frac{\sigma_{u,i,j}^2}{m_1} + \frac{\sigma_{u,i,j}^2}{m_2} \right)}}
\tag{19}
$$

where $\mu_{u,i,j} = \mathbb{E}[A_{u,i,j}], k_u = 1, 2, \ldots, m_u, u = 1, 2$. The asymptotic power can be presented in the theorem below.

**Theorem 2:** Under the same assumptions in Theorem 1, if $n^{-2/3}(s_1(\tilde{V}) - 4)^{-1} \leq o_n(1)$, then

$$
P(T_{TW} \geq \tau_{a/2}) = 1 - o_n(1).
\tag{20}
$$

**Proof:** Define a symmetric $n \times n$ matrix $W$ with zero diagonal and entries for all $1 \leq i < j \leq n$ as

$$
W_{ij} = \frac{\mu_{1,i,j} - \mu_{2,i,j}}{\sqrt{(n - 1) \left( \frac{\sigma_{u,i,j}^2}{m_1} + \frac{\sigma_{u,i,j}^2}{m_2} \right)}}
\tag{21}
$$

Recall the definitions of $V, V_0$ and $\tilde{V}$ given by (2), (5) and (19) respectively, from (15), it is observed that

$$
V_{ij} = (1 + o_p(n^{-2/3}))(\tilde{V}_{ij} - W_{ij}).
\tag{22}
$$

Thus

$$
(1 + o_p(n^{-2/3}))\tilde{V}_{ij} = V_{ij} + (1 + o_p(n^{-2/3}))W_{ij}.
\tag{23}
$$

The last equation implies

$$
V = \tilde{V} \circ (J + D) - W \circ (J + D),
\tag{24}
$$

where, $J$ is the $n \times n$ matrix with all the elements equal to 1, $D$ is an $n \times n$ matrix with elements $D_{ij} = o_p(n^{-2/3})$. Similarly to the proof of Theorem 1, we have $s_1(\tilde{V} \circ (J + D)) = s_1(\tilde{V})$ and $s_1(W \circ (J + D)) = s_1(W)$ with probability 1 as $n \to \infty$.

Applying the triangle inequality of spectral norm, we have

$$
s_1(V) \geq s_1(\tilde{V}) - s_1(W)
\tag{25}
$$

with probability 1 as $n$ tends to infinity. Noting that $W$ is a mean zero matrix whose singular value can be bounded by using the TW(1) asymptotic distribution. Hence, for any $\beta \in (0, 1),$

$$
P\left\{ s_1(W) \leq 2 + n^{-2/3} \tau_{\beta} \right\}
= 1 - P\left\{ s_1(W) > 2 + n^{-2/3} \tau_{\beta} \right\}
\geq 1 - \left( P\left\{ \lambda_1(W) > 2 + n^{-2/3} \tau_{\beta} \right\}
+ P\left\{ -\lambda_n(W) > 2 + n^{-2/3} \tau_{\beta} \right\} \right)
= 1 - 2\beta + o_n(1).\n\tag{26}
$$

Set $\tau_{\beta} = n^{2/3}(s_1(\tilde{V}) - 4) - \tau_{a/2}$, and plug this in (26), then we have

$$
1 - 2\beta + o_n(1)
\leq P\left\{ s_1(W) \leq 2 + n^{-2/3} \left( n^{2/3}(s_1(\tilde{V}) - 4) - \tau_{a/2} \right) \right\}
= P\left\{ 2 + n^{-2/3} \tau_{a/2} \leq s_1(\tilde{V}) - s_1(W) \right\}
\leq P\left\{ 2 + n^{-2/3} \tau_{a/2} \leq s_1(W) \right\}
= P\left\{ n^{2/3}(s_1(V) - 2) \geq \tau_{a/2} \right\}
= P(T_{TW} \geq \tau_{a/2}).\n\tag{27}
$$

If $n^{-2/3}(s_1(\tilde{V}) - 4)^{-1} \leq o_n(1)$, for a fixed $\alpha \in (0, 1)$, we have $\tau_{\beta}^{-1} = o_n(1)$, that is $\beta = o_n(1)$. Therefore,

$$
P(T_{TW} \geq \tau_{a/2}) = 1 + o_n(1).\n\tag{28}
$$

This theorem thus holds. □

**Remark 1:** It is noticeable that a test between two large graphs for $m_1 = m_2 = 1$ has been discussed in [22], in which the test statistic also has the TW(1) asymptotic distribution. However, rather than considering single network...
in each population, our testing statistic focuses on two-sample test on two populations of multiple networks, which has become increasingly more common in practice. Besides, [22] proves the asymptotic Tracy–Widom law under the binary random network setting, whereas in this paper, we are interested in weighted networks, and simultaneously can yield a two-sample test for binary networks when \( \{F_{1, i, j}\}_{1 \leq i, j \leq n} \) and \( \{F_{2, i, j}\}_{1 \leq i, j \leq n} \) are Bernoulli distributions, which will be discussed later. Finally, our theoretical results are based on appropriate estimates of the element-wise standard deviations instead of the unknown true parameters presented in Theorem 3 in [22]. Moreover, compared with the method in [15], which is based on geometric characterization between networks and involves the inverse of covariance matrix, our method is easier to compute.

C. COMPARISON OF DIFFERENT ESTIMATORS OF STANDARD DEVIATION

Various estimators for standard deviations of adjacency matrix entries can be used as the plugging estimates in the test statistic and we will analyze their properties in this subsection.

The most natural and simple estimator is the sample standard deviation applying to each sample group. We denote our proposed test based on sample standard deviation as SAMPT and the corresponding estimator by \( \hat{\sigma}_{\text{SAMPT}, u, i, j} \). Next, we consider its estimate error.

Define two \( n \times n \) matrices \( G_1 \) and \( G_2 \), with elements \( G_{u, i, j} = E[(A_{u, i, j})^2] \) for \( u = 1, 2 \), then

\[
|\hat{\sigma}_{\text{SAMPT}, u, i, j}^2 - \sigma_{u, i, j}^2| = \frac{\sum_{k_u = 1}^{m_u} (A_{u, i, j})^2 - m_u \bar{A}_{u, i, j})^2}{m_u - 1} - G_{u, i, j} + \sigma_{u, i, j}^2
\]

\[
= \frac{\sum_{k_u = 1}^{m_u} (A_{u, i, j})^2}{m_u} - G_{u, i, j} + \frac{\sum_{k_u = 1}^{m_u} (A_{u, i, j})^2}{m_u (m_u - 1)}
\]

\[
= \left( \sum_{k_u = 1}^{m_u} (A_{u, i, j})^2 \right) \left( G_{u, i, j} + \frac{\sum_{k_u = 1}^{m_u} (A_{u, i, j})^2}{m_u (m_u - 1)} \right)
\]

\[
\leq \frac{\sum_{k_u = 1}^{m_u} (A_{u, i, j})^2}{m_u} - G_{u, i, j} + \max_{k_u = 1, 2, \ldots, m_u} \left( \frac{(A_{u, i, j})^2}{m_u - 1} \right)
\]

\[
= O_p(m_u^{-1/2}) + \frac{\max_{k_u = 1, 2, \ldots, m_u} (A_{u, i, j})^2}{m_u - 1}
\]

\[
+ O_p(m_u^{-1/2}) \frac{\| \bar{A}_{u, i, j} \|_2 + \| \mu_{u, i, j} \|_2}{m_u - 1}
\]

\[
= O_p(m_u^{-1/2}), \tag{29}
\]

where the second to last equality comes from Bernstein’s inequality [28] and the last equality is due to the boundedness of distributions \( \{F_{u, i, j}\}_{1 \leq i, j \leq n} \) in our framework. It is observed that \( \hat{\sigma}_{\text{SAMPT}, u, i, j} \) and \( \sigma_{u, i, j} \) are bounded, so \( |\hat{\sigma}_{\text{SAMPT}, u, i, j} + \sigma_{u, i, j} | \) is also bounded, then max \( |\hat{\sigma}_{\text{SAMPT}, u, i, j} - \sigma_{u, i, j} | = O_p(m_u^{-1/2}) \).

It is noted that the SAMPT estimator is more suitable for cases of large sample size. Specifically, to satisfy the estimate error condition in Theorem 1, the sample size needs to be \( m_u = O_p(m_u^{1/2}) \) and \( \sigma_u > 7/3 \). This implies a faster increasing rate of \( m_u \) than that of nodes number \( n \). However, this is hard to hold in some scenarios like for fMRI networks. Actually, as pointed out in [29], it can be hard to obtain adequate number of samples in some neuroimaging studies.

Now, we consider an estimator of \( \sigma_{u, i, j} \) under the stochastic block model (SBM), which endows the networks with planted clusters [30]. This estimate method is inspired by [22]. First, assume the networks are SBMs, or approximate them with SBMs according to a weaker version of Szemerédi’s regularity lemma [31]. Then perform the community detection algorithms proposed in [32] to obtain the estimated membership vector \( v_u \in \{1, \ldots, l_u\} \) and the community set \( D_{u, i} = \{l_u \leq i \leq n, v_u_i = l_u \} \), where \( l_u \) is the estimated community number, \( l = 1, 2, \ldots, l_u \) and \( v_u_i \) is the \( i \)th element of \( v_u \). Then \( \sigma_{u, i, j} \) is consequently approximated by the sample standard deviation of all the entries \( A_{u, i', j'} \) for \( i' \in D_{u, v_u_i}, j' \in D_{u, v_u_j} \), and \( k_u = 1, 2, \ldots, m_u \). We refer our proposed test under this SBM-based standard deviation estimate as SBMT.

If we assume that each community number is at least proportional to \( n/l_u \), where \( l_u \) is the true community size, then similarly with SAMPT method, max \( |\hat{\sigma}_{\text{SBMT}, u, i, j} - \sigma_{u, i, j} | = O_p(l_u m_u^{-1/2} n^{-1}) \). This implies that when \( m_u = O_p(n^{1/2}), l_u = O(n^{1/2}), 0 < \gamma_u < \alpha_u/2 - 1/6 \), the estimate error condition in Theorem 1 can be satisfied. This condition further means \( \alpha_u > 1/3 + 2\gamma_u \) and is much less restrictive than that in SAMPT method especially when \( \gamma_u \) is controlled to be relatively small. Hence, the SAMPT method can generally achieve better performance, which is also demonstrated in simulations in Section III.

Although the theoretical results are based on weighted networks, our proposed method can also be applied to unweighted binary networks, where \( F_{u, i, j} = \text{Bernoulli}(P_{u, i, j}) \), with \( P_{u, i, j} \) the link probability matrix of the \( u \)th group and hence \( A_{u, i, j} \sim \text{Bernoulli}(P_{u, i, j}) \). For this case, we only need to estimate every mean parameter \( P_{u, i, j} \) instead of standard deviation \( \sigma_{u, i, j} \) since \( \sigma_{u, i, j} = P_{u, i, j}(1 - P_{u, i, j}) \). In a similar way, one can adopt the two estimate methods mentioned above but in terms of sample mean rather than sample standard deviation, and both of the properties remain the same. On the other hand, under the binary framework, [33] proposes a modified neighborhood smoothing (MNBS) method to estimate \( P_{u, i, j} \). Specifically, for each node, that paper defines its own neighborhood by some distance measure between nodes and employ the average network information restricted in that neighborhood. We refer this method as MNBS-based test (MNST).
III. SIMULATION STUDY

In this section, we evaluate the performance of the proposed TW(1)-type tests based on different standard deviation and mean value estimators. We compare SAMPT and SBMT with the test based on $T_2$ in [15] using geometric characterization of the space of graph (namely GEOT) for weighted network scenarios, and compare SAMPT, SBMT, and MNBST with the test based on $T_{[0]}$ in [22] involving some estimate of Frobenius distance between two network distributions (namely FROT) for binary network scenarios. For the consideration of covariance matrix inverse in [15], we only implement relatively small $n$ for GEOT in our simulations. To conduct simulations related to MNBS for binary network cases, a tuning parameter in [33] is required to be specified and we set it to be 3 as default. Moreover, the significance level is set at $\alpha = 0.05$ throughout this section. The testing performance is evaluated by the empirical significance level and empirical power through 5000 Monte Carlo simulations.

Three experiments are operationalized in this section. In the first experiment, we generate two groups of weighted networks with adjacency matrices $\{A^{(k_1)}_{1}, A^{(k_2)}_{2}\}$ from Beta distributions, and set $n$ growing from 10 to 100 with step 10 and $m_1 = m_2 = m = 80$. Each network has two communities with the membership $N(i) = I(1 \leq i \leq n/2)+2I(n/2+1 \leq i \leq n)$ for the $i$th node, where $I(\cdot)$ denotes the indicator function. If $N(i) = N(j)$, the nodes $i$ and $j$ are in the same community, then $A^{(k_1)}_{1,ij} \sim \text{Beta}(x_1, x_2)$ and $A^{(k_2)}_{2,ij} \sim \text{Beta}(y_1+y_2, x_2+y_2)$; otherwise, $A^{(k_1)}_{1,ij} \sim \text{Beta}(y_1, y_2)$ and $A^{(k_2)}_{2,ij} \sim \text{Beta}(x_1+y_1, x_2+y_2)$, where, $x_1 = 2, x_2 = 4, y_1 = 1, y_2 = 9, 1 \leq i < j \leq n, k_u = 1, 2, \ldots, m_u, u = 1, 2$. We take $\eta = 0$ for the null hypothesis and $\eta = 0.1$ for the alternative hypothesis. The results are reported in Figure 1.

To explore the performance of the proposed test on large-scale networks, we consider the second experiment, which is similar with the first one but with different parameters. We vary $n$ growing from 100 to 1000 with step 100, $m = 20, 30, x_1 = 2, x_2 = 4, y_1 = 1, y_2 = 8$, and set $\eta$ as 0 and 0.04 for the null hypothesis and the alternative hypothesis respectively. The results are shown in Figure 2.

The third experiment is for binary networks. We create networks with two communities from Bernoulli distributions. The settings of $n, m$, and membership for each node are the same as those in the second experiment. Under the null hypothesis, for $1 \leq i < j \leq n, k_u = 1, 2, \ldots, m_u, u = 1, 2, A^{(k_u)}_{u,ij} \sim \text{Bernoulli}(0.7)$ if $N(i) = N(j)$, i.e., the nodes $i$ and $j$ are in the same community; and $A^{(k_u)}_{u,ij} \sim \text{Bernoulli}(0.3)$ otherwise. Under the alternative hypothesis, the only distribution changes are $A^{(k_1)}_{2,ij} \sim \text{Bernoulli}(0.7+\eta)$, $\eta = 0.05, 1 \leq i < j \leq n/2$. The corresponding results are plotted in Figure 3.

The comparison results of the three experiments in Figures 1–3 reveal a superior performance of SBMT than that of SAMPT, GEOT and FROT in both significance level and power for all the cases and also the effectiveness of MNBST in binary setting cases. Specifically, for weighted network cases, Figure 1 and Figure 2 show the undesirable behaviors of GEOT and SAMPT since with increasing number of nodes $n$, the empirical significance levels of both tests grow quickly away from the significance level of $\alpha = 0.05$ and can become too large to be used in practice. This is due to that both GEOT and SAMPT require large sample sizes compared with the number of nodes. Figure 2 further shows that the inflated null rejection can be alleviated with larger sample sizes. It is also noted that the powers of GEOT and SAMPT are larger than those of SBMT, but with small differences. In contrast, the performance of SBMT is much better, the empirical significance levels are stable and smaller than $\alpha = 0.05$, and the powers also improve to 1 as $n$ grows.
As for unweighted network situations shown in Figure 3, SBMT and MNBST both perform well with empirical significance levels close to 0.05 and empirical powers increasing fast, while SAMPT suffers a high type I error and FROT has a large type II error with powers to be zero. The bad performance of FROT in power is partly caused by the minor differences between \( \{A(1)_{k1}\}_{k1=1}^{m1} \) and \( \{A(2)_{k2}\}_{k2=1}^{m2} \) we set, and the poor behaviors of GEOT and SAMPT are also due to the strong requirement of large sample size compared with \( n \). In addition, SBMT and MNBST fully utilize the inherent connections of networks and thus are more efficient even for large-scale networks.

Overall, for the tests based on three kinds of methods we propose to estimate standard deviation, SAMPT is more suitable for cases of large sample size, whereas SBMT is more robust, especially when the networks are SBM constructed, and MNBST also generally has good performance, but only can be applied to unweighted networks.

IV. SCHIZOPHRENIA DATA ANALYSIS

Schizophrenia is a complex mental illness characterized by abnormal brain function and is difficult to diagnose [34], [35]. Therefore, it is of great interest to compare SCZ and HC groups and automatically test whether there are significant differences between these two groups. A diagnostic system aided by hypothesis testing between their neuroimaging data would lead to a more objective approach to distinct schizophrenia and healthy people.

Consider a dataset contributed by the Centers of Biomedical Research Excellence (COBRE) included in INDI. It contains raw anatomical and functional scans from 147 subjects of 72 SCZs and 75 HCs, and can be downloaded from the public database (http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html). This connectomics dataset has been processed in [36]. A set of 264 ROIs from the predefined parcellation in [37] are used as nodes to construct the brain networks from fMRI measurements. These ROIs are derived through several
FIGURE 3. Empirical significance levels (top) and powers (bottom) of SAMPT, SBMT, MNBST and FROT for unweighted networks with Bernoulli-distribution binary edges and large-scale number of nodes $n$ from 100 to 1000. The sample sizes $m_1$ and $m_2$ of two groups are both equal to $m$ varying at $m = 20$ (left) and $m = 30$ (right).

meta-analyses of task fMRI data and with the fc-Mapping technique applied to rs-fcMRI data [37], [38]. This ROI definition has been widely used in neuroimaging studies because it is functional based and can provide better validity of networks than structural atlases or arbitrarily defined sampling grids [39]–[41]. The main pre-processing steps for all the scans sequentially contain slice time correction, anatomic normalization, frame-wise displacement analysis, extraction of cerebral spinal volume and white matter compartments, and nuisance regression to the time series of fMRI activity. Finally, the edge weights are measured by the Pearson $r$-values between the time series in corresponding ROIs, which is a common method to measure functional connectivity [42], [43]. More pre-processing and registration details can be found from supplemental materials in [36].

After a series of pre-processing procedures, 263 ROIs are chosen as nodes since node 75 in COBRE is missing, 54 SCZ and 70 HC subjects are kept and time course adjacency matrices of size $263 \times 263$ are extracted.

Figure 4 shows the mean connectivity networks over SCZ and HC groups, which initially visualizes the differences between them. In Figure 4, for convenience of visualization, the values of the Fisher $z$-transformed correlations between the nodes have been binarized with respect to the 75th percentile of the overall distribution of the entries in the full COBRE dataset. It can be seen from Figure 4 that there are differences between SCZ and HC groups, but not very distinct.

Now we test for health differences between two groups of brain networks extracted from the COBRE data using the proposed TW(1)-type two-sample test. Using the data pre-processed in [36], which are weighted networks, we analyze the graphical properties of these brain functional networks from two aspects. First, we test health differences between SCZs and HCs. Then, we assess the ability of our test to get
objects into the same category if they are in the same group (here we choose the group HC), which is helpful for automatic diagnosis. We do not compare our method with GEOT in this case because by half-vectorization of the Laplacian matrix, the dimension of the covariance matrix in the statistic in [15] reaches 34716 which is so large and difficult to be implemented in practice. In addition, the MNBST and FROT methods do not work in this case since they are designed for binary networks and are not applied to the weighted networks here.

In the first situation, the health differences between SCZ and HC groups are tested by SAMPT and SBMT and they both reject the null hypothesis with zero \( p \)-value. This implies significant differences between brain networks of SCZs and HCs. In the second situation, we randomly subsample 60 subjects from HCs and compare them with the original HC data. As expected, both tests fail to reject the null hypothesis under significant level of 0.05, since the samples are extracted from the same population. However, when we decrease the subsample size to 50, SBMT still fail to reject the null hypothesis while SAMPT shows a reject result. This confirms that good standard deviation estimators of adjacency entry standard deviations contribute to better testing performance.

V. CONCLUSION
Neuroimaging-based model plays a vital role in increasing our understanding of many clinical conditions. Motivated by the applications to neuroimaging, we have investigated hypothesis testing on whether two populations of weighted networks are from the same distribution. In this paper, we have proposed a novel singular value based hypothesis testing statistic which simultaneously can tackle weighted and large-scale networks. The asymptotic distribution of the null statistic and asymptotic power under alternative both have been derived. The test statistic utilizes some plug-in estimates for standard deviations of adjacency matrix entries, and properties of the resulting tests with various estimates have been discussed theoretically and numerically with both simulated and real brain network data. Moreover, the computation is easy, making it strongly practicable. From the example results, the proposed TW(1)-type tests based on SBM and MNBS estimates perform well.

In this paper, our hypothesis testing is based on the asymptotic extreme eigenvalue behaviors for the adjacency matrices of undirected networks with no self-loops. Although our setting is very general compared to the existing literature [15], [21]–[23], functional neuroimaging can be much more sophisticated than the setting in this work. For instance, networks from these data can have self-loops [44], [45]. Besides, directed influence of brain regions on each other can provide effective functional connectivity, however, neuroimaging investigations on the directed networks are still not studied thoroughly because of the technical constraints surrounding the inference [46]–[48]. Such complex networks can capture deeper connectivity information than normal networks. Therefore, in the future work, it is worth extending the hypothesis testing to networks with self-loops and directed structures. Actually, in this work, networks including self-loops can be tackled by removing the assumption \( A_{1,i,i}^{(k_1)} = 0 \) and modifying the scale factor \( n - 1 \) to \( n \) in the statistic (2) and Theorem 2 would hold for this case. For directed network, maybe we need to build a new test under this framework.

ACKNOWLEDGMENT
This work was initialized while Li Chen visited at University of Notre Dame.
REFERENCES

[1] J. Dubois, P. Galdi, L. K. Paul, and R. Adolphs, “A distributed brain network predicts general intelligence from resting-state human neuroimaging data,” *Phil. Trans. Roy. Soc. B, Biol. Sci.*, vol. 373, no. 1756, Sep. 2018, Art. no. 20170284.

[2] S. Bharath, H. Joshi, J. P. John, R. Balachandar, S. Sadanand, J. Saini, K. J. Kumar, and M. Varghese, “A multimodal structural and functional neuroimaging study of anesthetic mild cognitive impairment,” *Amer. J. Geriatric Psychiatry*, vol. 25, no. 2, pp. 158–169, Feb. 2017.

[3] R. L. Golubh, I. Kirsch, N. Maleki, A. D. Wasan, R. R. Edwards, Y. Tu, T. J. Kaptchuk, and J. Kong, “A functional neuroimaging study of expectancy effects on pain response in patients with knee osteoarthritis,” *J. Pain*, vol. 19, no. 5, pp. 515–527, May 2018.

[4] A. J. O’Toole, F. Jiang, H. Abdi, N. Pénard, J. P. Dunlop, and M. A. Parent, “Theoretical, statistical, and practical perspectives on pattern-based classification approaches to the analysis of functional neuroimaging data,” *J. Cognit. Neurosci.*, vol. 19, no. 11, pp. 1735–1752, Nov. 2007.

[5] E. Rykhlevskaia, G. Gratton, and M. Fabiani, “Combining structural and functional neuroimaging data for studying brain connectivity: A review,” *Psychophysiology*, vol. 45, no. 2, pp. 173–187, Mar. 2008.

[6] V. D. Calhoun, J. Sui, K. Kiehl, J. Turner, E. Allen, and G. Pearlson, “Exploring the psychosis functional connectome: Aberrant intrinsic networks in schizophrenia and bipolar disorder,” *Frontiers Psychiatry*, vol. 2, 2012, Art. no. 75.

[7] C. Calmels, M. Foutre, and C. J. Stam, “Beta functional connectivity modulation during the maintenance of motion information in working memory: Importance of the familiarity of the visual context,” *Neuroscience*, vol. 212, pp. 49–58, Jun. 2012.

[8] D. Fraiman, G. Saunier, E. F. Martins, and C. D. Vargas, “Biological motion coding in the brain: Analysis of visually driven EEG functional networks,” *PLoS ONE*, vol. 9, no. 1, Jan. 2014, Art. no. e84612.

[9] E. Bullmore and O. Sporns, “Complex brain networks: Graph theoretical analysis of structural and functional systems,” *Nature Rev. Neurosci.*, vol. 10, no. 3, pp. 186–199, 2009.

[10] B. B. Biswal et al., “Toward discovery science of human brain function,” *Proc. Nat. Acad. Sci. USA*, vol. 107, no. 10, pp. 4734–4739, 2010.

[11] M. Nennes, B. B. Biswal, F. X. Castellanos, and M. P. Milham, “Making data sharing work: The FCP/INDI experience,” *NeuroImage*, vol. 82, pp. 683–691, Nov. 2013.

[12] M. A. Lindquist, “The statistical analysis of fMRI data,” *L. Jollans, R. Boyle, E. Artiges, T. Banaschewski, S. Desrivières, VOLUME 8, 2020*

[13] D. Fraiman, G. Saunier, E. F. Martins, and C. D. Vargas, “Biological motion coding in the brain: Analysis of visually driven EEG functional networks,” *PLoS ONE*, vol. 9, no. 1, Jan. 2014, Art. no. e84612.

[14] E. Bullmore and O. Sporns, “Complex brain networks: Graph theoretical analysis of structural and functional systems,” *Nature Rev. Neurosci.*, vol. 10, no. 3, pp. 186–199, 2009.

[15] B. B. Biswal et al., “Toward discovery science of human brain function,” *Proc. Nat. Acad. Sci. USA*, vol. 107, no. 10, pp. 4734–4739, 2010.

[16] M. Nennes, B. B. Biswal, F. X. Castellanos, and M. P. Milham, “Making data sharing work: The FCP/INDI experience,” *NeuroImage*, vol. 82, pp. 683–691, Nov. 2013.

[17] M. A. Lindquist, “The statistical analysis of fMRI data,” *Statist. Sci.*, vol. 23, no. 4, pp. 439–464, Nov. 2008.

[18] L. Jollans, R. Boyle, E. Artiges, T. Banaschewski, S. Desrivières, A. Grigis, J.-L. Martinot, T. Paas, M. N. Smolka, H. Walter, G. Schumann, H. Garavan, and R. Whelan, “Quantifying performance of machine learning methods for neuroimaging data,” *NeuroImage*, vol. 199, pp. 351–365, Oct. 2019.

[19] N. Josephs, L. Lin, S. Rosenberg, and E. D. Kolaczyk, “Bayesian classification, anomaly detection, and survival analysis using neuroimage inputs with application to the microbiome,” *Appl. Stat.*, Apr. 2020, arXiv:2004.04765. [Online]. Available: http://arxiv.org/abs/2004.04765

[20] C. E. Ginestet, J. Li, P. Balachandran, S. Rosenberg, and E. D. Kolaczyk, “Hypothesis testing for network data in functional neuroimaging,” *Ann. Appl. Statist.*, vol. 11, no. 2, pp. 725–750, Jun. 2017.

[21] B. Zhang, H. Li, R. Riggins, M. Zhan, J. Xuan, Z. Zhang, K. E. Hoffman, R. Clarke, and Y. Wang, “Differential dependency network analysis to identify condition-specific topological changes in biological networks,” *Bioinformatics*, vol. 25, no. 4, pp. 526–532, Feb. 2009.

[22] V. D. Schmittmann, S. Jahafar, D. Borsboom, A. O. Savi, and L. J. Waldorp, “Making large-scale networks from fMRI data,” *PLoS ONE*, vol. 10, no. 9, Sep. 2015, Art. no. e0129074.

[23] I. Bojak, T. F. Oostendorp, A. T. Reid, and R. Kötter, “Connecting mean field models of neural activity to EEG and fMRI data,” *Brain Topography*, vol. 23, no. 2, pp. 139–149, Jun. 2010.

[24] N. A. Crossley, A. F. Stones, M. J. Everest, A. T. Reid, R. Kötter, and C. E. Ginestet, “Network structure and the community status of the human brain functional connectivity network,” *Proc. Nat. Acad. Sci. USA*, vol. 110, no. 28, pp. 11583–11588, Jul. 2013.

[25] S. Ryali, T. Chen, K. Supekar, and V. Menon, “Estimation of functional connectivity in fMRI data using stability selection-based sparse partial correlation with elastic net penalty,” *NeuroImage*, vol. 59, no. 4, pp. 3852–3861, Feb. 2012.
L. Chen et al.: Hypothesis Testing for Large Weighted Networks With Applications to Functional Neuroimaging Data

[43] C. Bachmann, H. I. L. Jacobs, P. Porta Mana, K. Dillen, N. Richter, B. von Reutern, J. Dronse, O. A. Onur, K.-J. Langen, G. R. Fink, J. Kukolja, and A. Morrison, “On the extraction and analysis of graphs from resting-state fMRI to support a correct and robust diagnostic tool for Alzheimer’s disease,” *Frontiers Neurosci.*, vol. 12, Sep. 2018, Art. no. 528.

[44] F. Cona, M. Zavaglia, M. Massimini, M. Rosanova, and M. Ursino, “A neural mass model of interconnected regions simulates rhythm propagation observed via TMS-EEG,” *NeuroImage*, vol. 57, no. 3, pp. 1045–1058, Aug. 2011.

[45] S. Schwab, R. Harbord, V. Zerbi, L. Elliott, S. Afyouni, J. Q. Smith, M. W. Woolrich, S. M. Smith, and T. E. Nichols, “Directed functional connectivity using dynamic graphical models,” *NeuroImage*, vol. 175, pp. 340–353, Jul. 2018.

[46] F. V. Farahani, W. Karwowski, and N. R. Lighthall, “Application of graph theory for identifying connectivity patterns in human brain networks: A systematic review,” *Frontiers Neurosci.*, vol. 13, Jun. 2019, Art. no. 585.

[47] G. A. Cecchi, A. R. Rao, M. V. Centeno, M. Baliki, A. V. Apkarian, and D. R. Chialvo, “Identifying directed links in large scale functional networks: Application to brain fMRI,” *BMC Cell Biol.*, vol. 8, no. S1, Jul. 2007, Art. no. S5.

[48] E. J. Friedman, K. Young, G. Tremper, J. Liang, A. S. Landsberg, N. Schuff, and Alzheimer’s Disease Neuroimaging Initiative, “Directed network motifs in Alzheimer’s disease and mild cognitive impairment,” *PLoS ONE*, vol. 10, no. 4, 2015, Art. no. e0124453.

**LI CHEN** received the B.S. degree in mathematics and applied mathematics from Sichuan Normal University, Chengdu, China, in 2013, and the M.S. degree in statistics from Sichuan University, Chengdu, in 2016, where she is currently pursuing the Ph.D. degree in applied statistics with the College of Mathematics. From 2018 to 2019, she was a Visiting Ph.D. Student with the University of Notre Dame, South Bend, IN, USA, which is supported by the China Scholarship Council. Her research interests include high-dimensional data analysis, statistical inference, and network analysis.

**LIZHEN LIN** received the Ph.D. degree in mathematics from the University of Arizona, Tucson, AZ, USA, in 2012. From 2012 to 2014, she was a Postdoctoral Associate with Duke University, Durham, NC, USA. From 2014 to 2016, she was an Assistant Professor with The University of Texas at Austin, Austin, TX, USA. She is currently an Associate Professor with the University of Notre Dame, South Bend, IN, USA. Her research interests include Bayesian asymptotics, Bayesian nonparametrics, big data analysis, statistics on manifolds, network analysis, probabilistic graphical models, and topological data analysis.

**JIE ZHOU** received the B.S. and M.S. degrees in pure mathematics and the Ph.D. degree in probability theory and mathematical statistics from Sichuan University, Chengdu, China, in 1989, 1992, and 2003, respectively. Since 1992, he has been with the College of Mathematics, Sichuan University, where he is currently a Professor. He has authored or coauthored three books and over 100 journal and conference proceeding papers. His current research interests include high-dimensional statistical inference, random matrices, statistical learning, network analysis, information geometry, information fusion, statistics of random processes, nonlinear filtering, and target tracking.

* * *