Received: 22 August 2023 | Revised: 9 June 2024 | Accepted: 16 August 2024

DOI: 10.1111/sjos.12750

ORIGINAL ARTICLE

Scandinavian Journal of Statistics

Goodness-of-fit testing based on graph functionals for homogeneous Erdös–Rényi graphs

Barbara Brune[1] | **Jonathan Flossdorf**[2] | **Carsten Jentsch**[2]

[1]Institute of Statistics and Mathematical Methods in Economics, TU Wien, Vienna, Austria

[2]Department of Statistics, TU Dortmund University, Dortmund, Germany

Correspondence
Jonathan Flossdorf, Department of Statistics, TU Dortmund University, Vogelpothsweg 78, 44221 Dortmund, Germany.
Email: flossdorf@statistik.tu-dortmund.de

Funding information
European Union Project: GADMO; Mercator Research Center Ruhr, Grant/Award Number: PR-2019-0019

Abstract

The Erdös–Rényi graph is a popular choice to model network data as it is parsimoniously parameterized, straightforward to interpret and easy to estimate. However, it has limited suitability in practice, since it often fails to capture crucial characteristics of real-world networks. To check its adequacy, we propose a novel class of goodness-of-fit tests for homogeneous Erdös–Rényi models against heterogeneous alternatives that permit nonconstant edge probabilities. We allow for both asymptotically dense and sparse networks. The tests are based on graph functionals that cover a broad class of network statistics for which we derive limiting distributions in a unified manner. The resulting class of asymptotic tests includes several existing tests as special cases. Further, we propose a parametric bootstrap and prove its consistency, which avoids the often tedious variance estimation for asymptotic tests and enables performance improvements for small network sizes. Moreover, under certain fixed and local alternatives, we provide a power analysis for some popular choices of subgraph counts as goodness-of-fit test statistics. We evaluate the proposed class of tests and illustrate our theoretical findings by simulations.

KEYWORDS

asymptotic theory, bootstrap consistency, parametric bootstrap, random graphs, stochastic networks, subgraph counts

1 | INTRODUCTION

Due to the technical progress in recent decades, not only the amount of data is growing but also its complexity. Hence, the analysis of statistical network data has become a popular research objective with applications ranging from social sciences (Carrington et al., 2005; Sarkar & Moore, 2005) or biology (Bassett & Sporns, 2017; Prill et al., 2005) to logistics and transportation processes (Lee & Dong, 2009). Various mathematical and computational methods have been developed to analyze, model, and understand the behavior of networks. As network data is rather complex by nature, it is a typical first step to employ suitable complexity reduction methods. A common approach is to calculate various sorts of summary statistics from the observed network such as, for example, matrix norms or centrality measures that may describe the network structure from different perspectives. However, the general difficulty is to be aware of the limited information that those statistics are able to capture, since they reduce the information of a whole network to a few scalar values (Flossdorf et al., 2023; Flossdorf & Jentsch, 2021; Ofori-Boateng et al., 2021). Another approach is to fit a suitable network model to the observed data in order to analyze graphs in a simplified and controlled setup and to perform statistical inference. One of the simplest network models is the Erdös–Rényi model (Gilbert, 1959) in which the edges are considered as independent Bernoulli random variables with a common probability parameter p. Obviously, this model may have substantial shortcomings in modeling real-world networks. Due to the assumption that all edges form independently with the *same* probability p, it usually fails to capture many of their features. For instance, in social networks, it is plausible that two people are more likely to know each other if they have a common friend. This yields clusters of vertices that have a higher edge probability between each other than to other vertices. That is, there may be several groups in a network, where two members of the same group have a higher probability to be connected than two members from different groups. A possible solution for this is to fit more general random graph models that allow for varying edge probabilities leading to the class of *heterogeneous* Erdös–Rényi models (Ouadah et al., 2020) including, for example, the popularly used stochastic block model (SBM) (Holland et al., 1983) that can also be seen as a special case of graphons (Gao et al., 2015; Lovász, 2012). While these heterogeneous models might match the features of real-world networks noticeably better than their homogeneous counterparts, their rigorous analysis is much more cumbersome from a graph theory point of view. From a practitioner's perspective, it is desirable to choose a parsimonious model with as few parameters as possible that still leads to a good fit to the data in order to simplify the estimation process. Hence, the homogeneous Erdös–Rényi model with only one probability parameter is still a popular benchmark model. Therefore, it is a crucial question if the usage of a parsimonious model is justified or if a more complex model achieves a significantly better fit. This leads to the field of goodness-of-fit procedures that aims to decide whether a specified model fits the underlying data adequately or if a more complex alternative would be a better choice.

Several approaches for goodness-of-fit testing for random graphs have been proposed in the literature. This involves tests for the number of communities in SBMs which are based on the largest singular value of a residual matrix that is obtained by the difference of the estimated block mean effect and the adjacency matrix (Lei, 2016). Jin et al. (2018) developed a testing procedure to decide whether an underlying network has more than one community by counting motifs under the consideration of the problem of degree heterogeneity. Under similar conditions, these authors later established optimality results for a class of tests that are based on signed polygon statistics (Jin et al., 2021). Testing procedures for the existence of communities are also studied in Yuan et al. (2022) using hypergraph motifs. Furthermore, the principal eigenvalue (Bickel & Sarkar, 2016) or the maximum entry (Hu et al., 2021) of a centered and scaled adjacency matrix is used to test whether a random graph is generated by a homogeneous Erdös–Rényi model or by an SBM. For the same hypotheses, Gao and Lafferty (2017) developed an asymptotic test based on specific induced subgraph counts. Subgraph counts as test statistics are also used in Maugis et al. (2017) under the rather restrictive and often unrealistic assumption that multiple independent and identically distributed samples of graphs of the same size are available, and in Ospina-Forero et al. (2019) within a Monte Carlo framework. A more general goodness-of-fit method is proposed in Dan and Bhattacharya (2020), where the goal is to test whether the probability matrix of independent samples of a heterogeneous graph matches a predetermined reference matrix. To do so, optimal minimax sample complexities in various matrix norms such as, for example, the Frobenius norm are derived. In Ouadah et al. (2020), goodness-of-fit tests for Erdös–Rényi-type models have been derived based on the degree variance statistic that serves as an heterogeneity index of a graph (Snijders, 1981). A statistical test for exponential random graph models (Robins et al., 2007) is proposed in Xu and Reinert (2021) with test statistics that are derived from a kernel Stein discrepancy. Furthermore, a goodness-of-fit method for stochastic actor oriented models (Snijders, 1996) is presented in Lospinoso and Snijders (2019) and variants of third-order motifs are used for the alternative of latent space models (Bubeck et al., 2016).

In this paper, we investigate a unified approach for goodness-of-fit testing for homogeneous Erdös–Rényi models. Precisely, we propose tests that aim to tell whether an observed network was either generated by some homogeneous Erdös–Rényi model or by some heterogeneous alternative model. Contrary to the related literature, we do not limit our analysis to just one particular test statistic, but use the concept of graph functionals (see Janson et al., 2011) for goodness-of-fit testing which allows for a unified treatment in theory and practice. Precisely, we propose a class of test statistics that covers a wide range of popular network metrics including, for example, the degree variance statistic, average centrality metrics, as well as raw and centered subgraph counts of arbitrary order and shape among others. Hence, our approach is very flexible and contains various already proposed tests (Gao & Lafferty, 2017; Ouadah et al., 2020) as special cases. Following this approach, we derive general asymptotic theory for the whole proposed class of test statistics in a unified manner. In particular, for newly proposed test statistics that fall into our framework, this avoids the tedious derivation of the individual asymptotics in a case by case manner. Furthermore, while the implementation of asymptotic tests is often tedious as their variances are highly case-dependent, we also propose a simple parametric bootstrap approach and prove a general bootstrap consistency result in order to develop bootstrap versions of the goodness-of-fit tests. This circumvents the tedious variance estimation and enables finite sample improvements.

The paper is organized as follows. It starts off with some general random graph theory and introduces the relevant notation in Section 2. In particular, we recap the concept of graph functionals for Erdös–Rényi models and discuss their representation as weighted sums of (centered) subgraph counts, which enables the derivation of general asymptotic theory for graph functionals.

Based on this theory, we construct a novel class of asymptotic goodness-of-fit tests in Section 3. We discuss some examples of graph functionals, including, for example, the degree variance as well as general notions of centered and raw subgraph counts, and investigate their suitability as goodness-of-fit statistics. Moreover, we provide a power analysis for the popular choices of subgraph counts of triangles and two-stars under certain fixed and local SBM alternatives. In Section 4, to avoid the often tedious variance estimation for the construction of asymptotic tests, we propose a general parametric bootstrap procedure for graph functional-based goodness-of-fit testing and establish bootstrap consistency results that justify the resulting bootstrap critical values. Subsequently, we investigate the performance of asymptotic and bootstrap versions of several graph functional-based goodness-of-fit tests in an extensive simulation study in Section 5. The final Section 6 consists of some concluding remarks. Proofs and additional results are deferred to the Appendix.

2 | PRELIMINARIES: RANDOM GRAPHS AND GRAPH FUNCTIONALS

In this section, we give an overview of the used network generating models and the concept of graph functionals. Particularly, we gather existing limiting distribution theory for this class of statistics for Erdös–Rényi graphs in order to enable the development of unified asymptotic and bootstrap goodness-of-fit tests.

2.1 | Settings

Suppose we observe a random graph G that is defined by a vertex set $V = V(G)$, and an edge set $E = E(G)$. The edge set consists of pairs of vertices $\{v_1, v_2\}$, where $v_1, v_2 \in V(G)$. We denote by $n := n(G) = |V(G)|$ the number of vertices, and by $m := m(G) = |E(G)|$ the number of edges of the graph G, respectively. The vertex set is denoted by $V(G) = \{v_1, \dots, v_n\}$, and each edge in $E(G) = \{e_1, \dots, e_m\}$ connects two of the vertices. We define the empty graph as a graph with no edges and the null graph $\emptyset$ as the graph with no vertices and, thus, no edges. A network can alternatively be represented by an adjacency matrix $A = (A_{ij})$ of dimension $(n \times n)$. In this paper, we concentrate on unweighted and undirected networks without self-loops, which are often called simple graphs. Their adjacency matrices are symmetric with binary entries such that $A_{ij} = 1$ indicates the presence of an edge between two vertices i and j, and $A_{ij} = 0$ if there is no edge between the respective vertices. As self-loops are not allowed, we have $A_{ii} = 0$ for all i. Throughout the paper, we focus on random graphs and use the well-established Erdös–Rényi graph as the benchmark model.

Definition 1 (Erdös–Rényi (ER) graph). Let $G = (V, E)$ be a random graph on n vertices with adjacency matrix A and let $p \in [0, 1]$ be the connection probability. Then, we call G an *Erdös–Rényi graph* $\mathcal{G}(n, p)$, if the edges are realizations of stochastically independent and identically Bernoulli distributed random variables. That is, for $1 \leq i < j \leq n$, we have $A_{ij} \sim \text{Bin}(1, p)$ with $A_{ji} := A_{ij}$, and $A_{ii} := 0$ for all i. We denote the resulting ER model class by

$$\mathcal{G}_{ER}(n) = \{\mathcal{G}(n, p), p \in [0, 1]\}.$$

As the ER model has quite restrictive assumptions, various generalizations have been studied. A quite flexible alternative is the heterogeneous Erdös–Rényi model (Ouadah et al., 2020).

Definition 2 (Heterogeneous Erdös–Rényi [HER] graph). Let $G = (V, E)$ be a random graph on n vertices with adjacency matrix A and let $\boldsymbol{P} = (p_{ij})_{i,j=1,\dots,n}$ be the symmetric $(n \times n)$ matrix of connection probabilities with $p_{ij} \in [0, 1]$. Then, we call G a *heterogeneous Erdös–Rényi graph* $\mathcal{G}(n, \boldsymbol{P})$, if the edges are realizations of stochastically independent Bernoulli random variables. That is, for $1 \leq i < j \leq n$, we have $A_{ij} \sim \text{Bin}(1, p_{ij})$ with $A_{ji} := A_{ij}$, and $A_{ii} := 0$ for all i. We denote the resulting HER model class by

$$\mathcal{G}_{\text{HER}}(n) = \{\mathcal{G}(n, \boldsymbol{P}), \boldsymbol{P} = (p_{ij}), p_{ii} = 0, p_{ij} = p_{ji}, p_{ij} \in [0, 1] \ \forall\ i, j\}.$$

In a nutshell, the HER model extends the classical ER model by offering the opportunity for individual link probabilities for each edge. This expansion is helpful for the modeling of more flexible scenarios. However, it also increases the complexity. For example, parameter estimation in this model becomes infeasible without further restrictions. Hence, we are interested in testing whether a parsimonious homogeneous ER model is already sufficient to model the underlying data or whether an HER model achieves a significantly better fit. Consequently, having observed a simple graph G of size n, we consider the testing problem

$$\text{H}_0 : \ G \in \mathcal{G}_{\text{ER}}(n) \qquad \text{versus} \qquad \text{H}_1 : \ G \in \mathcal{G}_{\text{HER}}(n) \setminus \mathcal{G}_{\text{ER}}(n). \tag{1}$$

Although $\mathcal{G}_{\text{HER}}(n)$ can be decomposed into disjoint sets, that is, in $\mathcal{G}_{\text{ER}}(n)$ and $\mathcal{G}_{\text{HER}}(n) \setminus \mathcal{G}_{\text{ER}}(n)$ as above, it will not be possible to consistently detect *arbitrary* alternatives as $n \to \infty$, when testing H_0 against H_1. This is because a heterogeneous ER model has to deviate *sufficiently enough* from the homogeneous model to be able to detect it, since we only rely on a single network observation in the goodness-of-fit context. For example, consider an HER model that has the same edge probability p for all except a finite and (for increasing n) fixed number of edges, that have edge probability q with $q \neq p$. Then, as only finitely many edge probabilities deviate from p, such heterogeneous alternatives will be not detectable in the limit as n increases.

2.2 | Graph functionals

We use the concept of graph functionals for the development of a general and rich class of goodness-of-fit tests. In this subsection, we gather existing theory for graph functionals that is required in order to derive general asymptotic theory.

2.2.1 | General concept

Graph functionals cover a wide range of network statistics that can capture various forms of structural patterns within a given network. Hence, the class of graph functionals appears to be favorable for goodness-of-fit testing. Following Janson et al. (2011), we make use of the following definition of graph functionals.

Definition 3 (Graph isomorphism, automorphism, and graph functional). Let G and H be two simple graphs. An isomorphism of graph G onto H is a bijection $\varphi : V(G) \to V(H)$ such that any two vertices $v_1, v_2 \in V(G)$ are adjacent in G if and only if the vertices $\varphi(v_1), \varphi(v_2) \in V(H)$ are adjacent in H. If there exists an isomorphism between G and H, we call G and H isomorphic. A real-valued random variable $X_n = X_n(G)$ that only depends on the isomorphism type of a graph G of size n, that is, with $X_n(G) = X_n(H)$ for all G and H that are isomorphic, is called a *graph functional*. Further, we denote by $\mathrm{aut}(H)$ the number of automorphisms of H, that is, the number of isomorphisms for $G = H$.

In other words, a graph functional is a function computed from a graph G that does not hinge on the vertex labels. We give some examples of special cases next, which will serve as running examples throughout this paper for illustration purposes.

Example 1 (Frobenius norm). A rather simple example for a graph functional is the squared Frobenius norm $\|A\|_F^2$ of the adjacency matrix $A = (A_{ij})$ of a graph $G = (V, E)$. It is defined by

$$\|A\|_F^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij}^2.$$

By summing up all (squared) entries of A, it is easy to see that relabeling the vertices of G does not change $\|A\|_F^2$ such that $\|A\|_F^2$ is a graph functional. As we concentrate on undirected and unweighted graphs, which have a symmetric adjacency matrix A with binary entries A_{ij}, we have

$$\|A\|_F^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij}^2 = 2 \sum_{1 \le i < j \le n} A_{ij} = 2m = \binom{n}{2} \widehat{p}, \tag{2}$$

where $m = |E(G)|$ is the number of edges of G and $\widehat{p} = \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} A_{ij}$ is the natural estimator of p based on the adjacency matrix $A = (A_{ij})$. Since the Frobenius norm counts the number of edges and is thus the same for every permutation of a given graph, it belongs to the class of graph functionals.

Example 2 (Degree variance). The *degree variance* V_n of a (simple) graph $G = (V, E)$ is defined as

$$V_n := \frac{1}{n} \sum_{i=1}^{n} (D_i - \overline{D}_n)^2, \tag{3}$$

where D_i is the degree of vertex v_i and $\overline{D}_n = \frac{1}{n} \sum_{i=1}^{n} D_i$ denotes the average degree of the underlying network G. Note that the individual degrees D_i do actually hinge on the vertex labels, but the sum of (squared) degrees as well as the average degree do not change by relabeling the graph. Hence, V_n is invariant under isomorphism change and, consequently, a graph functional. The degree variance is quite popular as it is an intuitive metric that can handily be computed directly from the adjacency matrix. As pointed out by Snijders (1981), it serves as a certain heterogeneity index of a graph and has been used for goodness-of-fit testing in Ouadah et al. (2020). Further investigated in Flossdorf and Jentsch (2021), it is shown that V_n performs reasonably well in capturing local structures of the graph (e.g., centrality traits), but also that V_n has weaknesses, when it comes to capturing global characteristics like the overall amount of links.

Example 3 (Eigenvector centrality). Eigenvector centrality can be interpreted as a natural extension of the degree centrality. It is measured relatively to the amount of relations a node has to other nodes, and relationships to high-scoring nodes are more valuable than those to low-scoring ones. Thus, not only the number of links is taken into account but also the influences of the nodes to which the links exist. By denoting the score of node i with x_i, the eigenvector centrality is formally defined by

$$x_i = \frac{1}{\lambda}\sum_{j=1}^{n} a_{ij}x_j, \quad i = 1,\dots,n,$$

where λ is a constant. If we define the vector of scores as $x = (x_1,\dots,x_n)$, it is possible to express the equation in matrix form $\lambda x = Ax$. Hence, λ is an eigenvalue of A and x is the corresponding eigenvector which includes the centrality scores. In order to get a metric for the whole network, the *average eigenvector centrality* is defined as

$$\mathrm{E}_n = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Note that the eigenvector values x_i are typically scaled to be able to compare the different scores properly with each other. We use the internal scaling of the eigenvector, that is, we replace x_i by $x_i/\|x\|$, as it has unit length in the Euclidean norm. By the same arguments as used in the above examples, the metric E_n is a graph functional.

While these examples focus on three concrete special cases of graph functionals, we consider also the more general concept of subgraph counts. According to (Janson et al., 2011, p. 165), we define centered and raw versions of subgraph counts as follows.

Definition 4 (Subgraph counts). Let G be a (simple) graph on n vertices, and H be another graph with $n(H) \leq n$. Consider the $(n)_{n(H)}$ different injective mappings φ from vertices of H into $\{1,\dots,n\}$, where, for $x, y \in \mathbb{N}_0$ with $x \leq y$, let $(x)_y = x(x-1)(x-2)\dots(x-y+1)$ denote the descending factorials. Each mapping φ maps H onto a copy $\varphi(H)$ of H in a complete graph with n vertices. Then, the *centered subgraph count* of H in G is defined by

$$S_n(H) = \sum_{\varphi} \prod_{e \in E(\varphi(H))} (\mathbb{1}\{e \in E(G)\} - \mathbb{P}(e \in E(G))),$$

while the *raw subgraph count* of H in G is defined by

$$T_n(H) = \frac{1}{\mathrm{aut}(H)} \sum_{\varphi} \prod_{e \in E(\varphi(H))} \mathbb{1}\{e \in E(G)\}.$$

It is worth noting that a more precise notation would be to use $S_G(H)$ and $T_G(H)$ as the subgraph counts depend on the whole graph G and not only on its size $n = n(G) = |V(G)|$. However, as we apply the subgraph counts always to a graph G, we suppress this dependence and write $S_n(H)$ and $T_n(H)$, respectively.

While, in a relabeled graph, the subgraph of interest may exist between different node combinations compared to the observed graph, the sum of such subgraphs remains the same over the whole network. Hence, centered subgraph counts $S_n(H)$ as well as raw subgraph counts $T_n(H)$ belong to the class of graph functionals and form important subclasses. Although defined similarly, following the expositions in Janson et al. (2011), they mainly differ in three important aspects:

(i) For $H \neq \emptyset$, $S_n(H)$ is a random variable with expectation 0, that is, $E(S_n(H)) = 0$, for models with independently formed edges, as $\mathbb{P}(e \in E(G)) = \mathbb{E}(\mathbb{1}\{e \in E(G)\})$, which is not the case for $T_n(H)$, where $E(T_n(H)) > 0$ typically holds.
(ii) While $T_n(H)$ can be computed exclusively from the graph G (i.e., from its adjacency matrix A), this is not the case for $S_n(H)$, which requires knowledge of $\mathbb{P}(e \in E(G))$, that is, p for Erdös–Rényi graphs.
(iii) While $T_n(H)$ counts a substructure H only once for each vertex combination (formally by dividing by $\mathrm{aut}(H)$), $S_n(H)$ counts all its automorphisms.

Subgraph counts are popular metrics for characterizing networks and are used for various inferential approaches (Bhattacharyya & Bickel, 2015; Gao & Lafferty, 2017; Maugis et al., 2020). However, these methods typically use subgraph counts of a certain simple shape and small order, which makes them quite restrictive in their practical application. Even for simple subgraph structures, the computational complexity quickly becomes large compared to, for example, the degree variance statistic. However, as subgraph counts are always graph functionals, this property allows us to derive asymptotics in a unified way for flexible shapes and orders, as we will show throughout the paper. Popularly used subgraphs are, for example, triangles C_3 ( ), two-stars P_3 ( ) and single edges P_2 ( ) which are depicted in Figure 1a.

Remark 1 (Differing notions of subgraph counts). Note that there are different notions of subgraph counts in the literature. Differences arise from considering the order of the vertices, that is, counting a substructure once for each vertex combination, or counting all automorphisms. Another aspect is whether to consider only induced copies of subgraphs, that is, also preserving nonedges, or all noninduced copies, not necessarily preserving nonedges. Our definitions above are in accordance with those of Janson et al. (2011). For $S_n(H)$, all automorphisms for each vertex combination are considered, as can be seen from the index of the sum. For $T_n(H)$, the prefactor causes us to count a substructure only once for each vertex combination, which is a somewhat more intuitive concept. Both $T_n(H)$ and $S_n(H)$ are defined as noninduced versions. However, under the condition $p = o(1)$ the induced and noninduced versions are asymptotically the same (Janson et al., 2011, p. 78).

Figure 1b illustrates the mechanism of subgraph counts with an example: In a complete graph of order 4, we can find four subgraphs of shape C_3, hence $T_4(C_3) = 4$. For the calculation of the centered versions, we would have to consider each of these $\mathrm{aut}(C_3) = 6$ times, that is, the sum in the definition of S_n would range over $(4)_3 = 4 \cdot 3 \cdot 2 = 24$ mappings φ leading then to $S_4(C_3) = 24$.

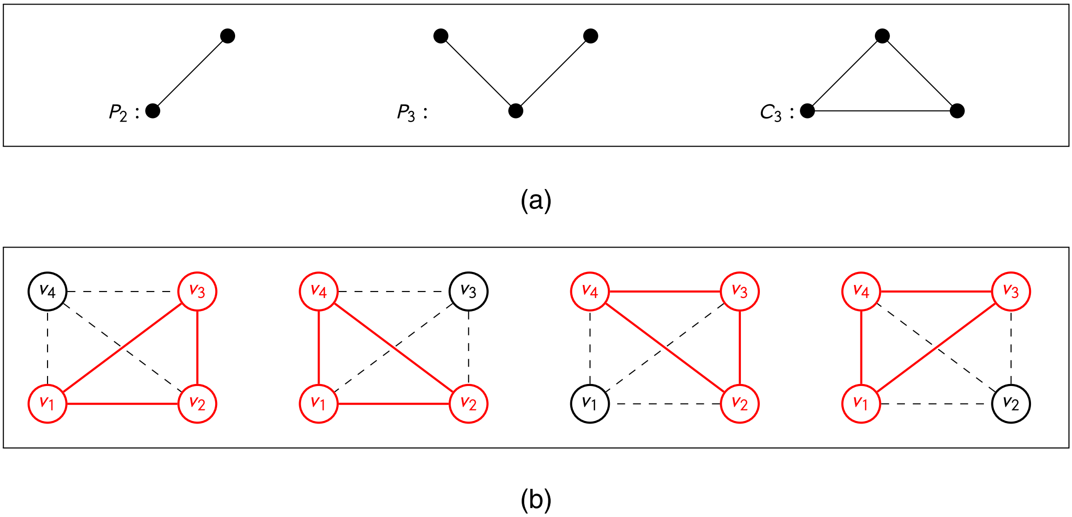

FIGURE 1 Illustration of the mechanism of subgraphs and subgraph counts. (a) Examples for subgraphs: edge P_2, two-star P_3 and circle C_3 (of size 3); (b) All raw subgraphs (circles) of shape C_3 in a complete graph with $n = 4$ vertices leading to $T_4(C_3) = 4$.

2.2.2 | Unified representation

The key foundation of our unified procedure for constructing goodness-of-fit tests lies in a fundamental relationship between graph functionals and *centered* subgraph counts: Any graph functional $X_n = X_n(G)$ can be written as a linear combination of centered subgraph counts $S_n(H)$. This expansion enables the unified derivation of the asymptotic distributions of graph functionals X_n and is referred to as the method of higher projections. In particular, the representation of X_n in terms of $S_n(H)$'s allows to exploit beneficial features of centered subgraph counts that are summarized in the following proposition.

Proposition 1 (Janson et al., 2011, lemma 6.41). *Let H and J, $H, J \neq \emptyset$, be graphs without isolated vertices. Then, we have*

(a) $\mathbb{E}(S_n(H)) = 0$.
(b) $\mathbb{V}\mathrm{ar}(S_n(H)) = \mathrm{aut}(H)(n)_{n(H)}(p(1-p))^{m(H)}$, *where* $(x)_y = x(x-1)(x-2)\cdots(x-y+1)$ *denotes the descending factorials.*
(c) If H and J are nonisomorphic, then $S_n(H)$ and $S_n(J)$ are orthogonal, that is, $\mathbb{C}\mathrm{ov}(S_n(H), S_n(J)) = 0$.

That is, for any nonnull graph H, we have that $S_n(H)$ is a centered random variable, that is, $\mathbb{E}(S_n(H)) = 0$ and its variance $\mathbb{V}\mathrm{ar}(S_n(H))$ can be explicitly stated and has a rather simple form. Moreover, whenever two (nonnull) graphs H and J are not isomorphic, then $S_n(H)$ and $S_n(J)$ are uncorrelated, that is, $\mathbb{C}\mathrm{ov}(S_n(H), S_n(J)) = 0$.

In addition to the nice properties summarized in Proposition 1, the following result states that centered subgraph counts $S_n(H)$ can be used to (linearly) decompose *any* graph functional.

Proposition 2 (Janson et al., 2011, lemma 6.42). *Every graph functional X_n of a $\mathcal{G}(n,p)$ graph has a unique expansion*

$$X_n = \sum_{H \in \mathcal{H}} a_n(H) S_n(H), \tag{4}$$

for some real-valued coefficients $a_n(H) = a_n(H,p)$ depending on n, p and H, where $\mathcal{H}$ contains all unlabeled graphs without isolated vertices and of at most n vertices. Furthermore, the terms in the sum are orthogonal, hence

$$\mathbb{V}\mathrm{ar}(X_n) = \sum_{H \in \mathcal{H}} a_n^2(H) \mathbb{V}\mathrm{ar}(S_n(H)). \tag{5}$$

Note that the expectation of X_n is represented solely by the coefficient $a_n(\emptyset)$, that is, $E(X_n) = a_n(\emptyset)$. This is because the count of null graphs is deterministic and given by $S_n(\emptyset) := 1$ and all other terms $S_n(H)$, $H \neq \emptyset$, are centered by construction. While Proposition 2 guarantees the expansion (4) for arbitrary graph functionals X_n, it may be the case that not all centered subgraph counts $S_n(H)$ in (5) actually contribute (asymptotically) to $\mathbb{V}\mathrm{ar}(X_n)$.

Definition 5 (Dominance of families of subgraphs). For a family of nonnull unlabeled graphs without isolated vertices $\mathcal{H}_0$, we call a graph functional X_n *dominated by* $\mathcal{H}_0$ if, for a given sequence of connection probabilities $p = p(n) \xrightarrow[n\to\infty]{} p_0 \in [0,1]$, it holds

$$\frac{\mathbb{V}\mathrm{ar}(X_n)}{\sum_{H \in \mathcal{H}_0} a_n^2(H) \mathbb{V}\mathrm{ar}(S_n(H))} = \frac{\sum_{H \in \mathcal{H}} a_n^2(H) \mathbb{V}\mathrm{ar}(S_n(H))}{\sum_{H \in \mathcal{H}_0} a_n^2(H) \mathbb{V}\mathrm{ar}(S_n(H))} \xrightarrow[n\to\infty]{} 1.$$

Being dominated by a family of (sub)graphs $\mathcal{H}_0$ means that asymptotically all the variance of the graph functional X_n is explained by those centered subgraphs in the family $\mathcal{H}_0$, which may be smaller than $\mathcal{H}$ as defined in Proposition 2.

2.2.3 | Asymptotics

The representation of graph functionals as linear combinations of centered subgraph counts enables a unified derivation of asymptotic theory for the whole class of graph functionals. While Nowicki and Wierman (1988) and Nowicki (1989) showed that raw subgraph counts $T_n(H)$ follow asymptotic normal distributions if the underlying model is contained in $\mathcal{G}_{\mathrm{ER}}(n)$, we are interested in finding the asymptotic distributions for *centered* subgraph counts $S_n(H)$ in order to get also asymptotic results for graph functionals via Proposition 2. In particular, by combining the statements of Propositions 1 and 2 with the asymptotic results established for $S_n(H)$'s in Thm. 6.43 in Janson et al. (2011), we get the following general asymptotic normality result for graph functionals.

Theorem 1 (Janson et al., 2011, Thm. 6.49). *Let X_n be a graph functional of $\mathcal{G}(n,p)$ with $p = p(n) \xrightarrow[n\to\infty]{} p_0 \in [0,1]$. Suppose X_n is dominated by a family of connected graphs $\mathcal{H}_0$ such that, for all $H \in \mathcal{H}_0$ and $np^{r(H)} \xrightarrow[n\to\infty]{} \infty$ with $r(H) = \max_{J \subseteq H} d(J)$, where $d(J) = m(J)/n(J)$, the coefficients*

$$b(H) = \sup_n \frac{n^{n(H)/2} p^{m(H)/2} a_n(H)}{\sqrt{\mathbb{V}\mathrm{ar}(X_n)}},$$

are finite and satisfy

$$\sum_{H \in \mathcal{H}_0} b(H)^2 \mathrm{aut}(H) < \infty.$$

Then, as $n \to \infty$, it holds

$$R(X_n) := \frac{X_n - \mathbb{E}(X_n)}{\sqrt{\mathbb{V}\mathrm{ar}(X_n)}} \xrightarrow{d} \mathcal{N}(0, 1). \tag{6}$$

Note that the expression $J \subseteq H$ in the above definition of $r(H)$ includes all graphs for which $E(J) \subseteq E(H)$ and $V(J) \subseteq V(H)$. As a direct consequence of Theorem 1, there exists a nonnegative sequence $(c_n)_{n \in \mathbb{N}}$ with $c_n = c(n, p(n))$ such that $c_n^2 \mathbb{V}\mathrm{ar}(X_n) \to V^2 > 0$ as $n \to \infty$, leading to $c_n(X_n - \mathbb{E}(X_n)) \xrightarrow{d} \mathcal{N}(0, V^2)$.

The general representation of graph functionals as linear combinations of centered subgraph counts in Proposition 2 and the limiting distributions derived in Theorem 1 form together a powerful framework that enables a unified treatment and a flexible construction of goodness-of-fit tests for (homogeneous) Erdös–Rényi random graphs.

3 | GOODNESS-OF-FIT TESTING FOR ERDÖS–RÉNYI MODELS

In this section, we formulate our unified approach for goodness-of-fit testing for ER graphs based on graph functionals. That is, having observed a simple graph G of size n and recalling (1), we restate the testing problem

$$\mathrm{H}_0 : \; G \in \mathcal{G}_{\mathrm{ER}}(n) \qquad \text{versus} \qquad \mathrm{H}_1 : \; G \in \mathcal{G}_{\mathrm{HER}}(n) \setminus \mathcal{G}_{\mathrm{ER}}(n). \tag{7}$$

Consequently, under the assumptions of Theorem 1 and in view of the asymptotic normality result in (6), it appears natural to pick a certain graph functional $X_n = X_n(G)$ and to use $R(X_n)$ as the test statistic for testing H_0 against H_1 in (7), where $\mathbb{E}(X_n)$ and $\mathbb{V}\mathrm{ar}(X_n)$ have to be replaced by $\mathbb{E}_{H_0}(X_n)$ and $\mathbb{V}\mathrm{ar}_{H_0}(X_n)$, which denote expectation and variance of X_n under H_0, respectively. However, this is usually not feasible, because neither $\mathbb{E}_{H_0}(X_n)$ nor $\mathbb{V}\mathrm{ar}_{H_0}(X_n)$ are typically known. Hence, instead, we will consider feasible goodness-of-fit test statistics of the form

$$\widehat{R}(X_n) := \frac{X_n - \widehat{\mathbb{E}_{H_0}}(X_n)}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(X_n)}}, \tag{8}$$

where $\widehat{\mathbb{E}_{H_0}}(X_n)$ and $\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(X_n)$ are suitable (consistent) estimators of $\mathbb{E}_{H_0}(X_n)$ and $\mathbb{V}\mathrm{ar}_{H_0}(X_n)$, respectively. Such estimators are naturally obtained by replacing the unknown parameter p in the explicit expressions of $\mathbb{E}_{H_0}(X_n)$ and $\mathbb{V}\mathrm{ar}_{H_0}(X_n)$ by its consistent estimator $\widehat{p}$ defined in (2).

However, note that $\mathbb{E}_{H_0}(X_n)$ and $\mathbb{V}\mathrm{ar}_{H_0}(X_n)$ are generally difficult to derive directly. Making use of Proposition 2 (see also the discussion below Proposition 2), under H_0, we have

$$\mathbb{E}_{H_0}(X_n) = \sum_{H\in\mathcal{H}} a_n(H)\mathbb{E}_{H_0}(S_n(H)) = a_n(\emptyset)\mathbb{E}_{H_0}(S_n(\emptyset)) = a_n(\emptyset),$$
$$\mathbb{V}\mathrm{ar}_{H_0}(X_n) = \sum_{H\in\mathcal{H}} a_n^2(H)\mathbb{V}\mathrm{ar}(S_n(H)).$$

As the coefficients $a_n(H)$ depend on n, p and H, i.e. $a_n(H) = a_n(H, p)$, estimators $\widehat{\mathbb{E}_{H_0}}(X_n)$ and $\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(X_n)$ are naturally obtained by defining

$$\widehat{\mathbb{E}_{H_0}}(X_n) = \hat{a}_n(\emptyset),$$
$$\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(X_n) = \sum_{H\in\mathcal{H}} \hat{a}_n^2(H)\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(S_n(H)),$$

where $\hat{a}_n(H) = a_n(H, \hat{p})$ and $\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(S_n(H)) = \mathrm{aut}(H)(n)_{n(H)}(\hat{p}(1-\hat{p}))^{m(H)}$. With these prerequisites in place, by making use of Theorem 1, we are able to make the following general statement about the limiting distribution of $\hat{R}(X_n)$ defined in (8) under H_0.

Proposition 3. *Let $X_n = X_n(G)$ be a graph functional computed from an Erdös–Rényi graph G and suppose the assumptions of Theorem 1 are fulfilled. Additionally, let*

$$\text{(A)} \quad \frac{\mathbb{V}\mathrm{ar}_{H_0}(X_n)}{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(X_n)} \xrightarrow{P} 1 \quad \text{and} \quad \text{(B)} \quad \frac{\mathbb{E}_{H_0}(X_n) - \widehat{\mathbb{E}_{H_0}}(X_n)}{\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(X_n)}} = \frac{a_n(\emptyset) - \hat{a}_n(\emptyset)}{\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(X_n)}} \xrightarrow{P} 0, \qquad (9)$$

hold. Then, as $n \to \infty$, we have

$$\hat{R}(X_n) = \frac{X_n - \widehat{\mathbb{E}_{H_0}}(X_n)}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(X_n)}} \xrightarrow{d} \mathcal{N}(0,1).$$

An asymptotic level-α test for testing the null hypothesis H_0 of an underlying homogeneous Erdös–Rényi graph against heterogeneous alternatives H_1 in (1) is given by the decision rule

$$\phi_\alpha(X_n) = \begin{cases} 1, & \text{if } \left|\hat{R}(X_n)\right| > z_{1-\alpha/2}, \\ 0, & \text{else,} \end{cases}$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution.

The additional conditions (A) and (B) imposed in (9) are important to guarantee the limiting standard normal distribution of the *feasible* test statistic $\hat{R}(X_n)$ in (8). While condition (A) is likely to hold in relevant cases of graph functionals used in practice, condition (B) is actually more delicate. As we will see in the following illustrating examples, it is case dependent, whether this condition holds for a certain graph functional or not. Next, we pick up Example 1 from above and show that the Frobenius norm $\|A\|_F^2$ is a graph functional, that does *not* satisfy property (9) as condition (B) fails to hold. Additionally, we discuss why $\|A\|_F^2$ is not suitable at all for testing H_0 against H_1.

Example 1 (Frobenius norm). On the one hand, according to (2), we have $\|A\|_F^2 = \binom{n}{2}\hat{p}$. On the other hand, we have $\hat{p} = \binom{n}{2}^{-1} \sum_{1\le i<j\le n} A_{ij} = \binom{n}{2}^{-1} T_n(P_2)$ with

$$\begin{aligned}
T_n(P_2) &= \frac{1}{\mathrm{aut}(P_2)}\sum_{\varphi}\prod_{e\in E(\varphi(P_2))} \mathbb{1}\{e\in E(G)\} = \frac{1}{\mathrm{aut}(P_2)}\sum_{\varphi}\prod_{e\in E(\varphi(P_2))} (\mathbb{1}\{e\in E(G)\} - p + p)\,. \\
&= \frac{1}{\mathrm{aut}(P_2)}\left(\sum_{\varphi}\prod_{e\in E(\varphi(P_2))} (\mathbb{1}\{e\in E(G)\} - p) + \sum_{\varphi}\prod_{e\in E(\varphi(P_2))} p\right) \\
&= \frac{1}{\mathrm{aut}(P_2)}(S_n(P_2) + n(n-1)p) \\
&= \frac{1}{2}S_n(P_2) + \binom{n}{2}p,
\end{aligned}$$

where we used that the product $\prod_{e\in E(\varphi(P_2))}$ consists only of one factor, as P_2 has only one edge, and $p = \mathbb{P}(e\in E(G))$. Altogether, when considering $\|A\|_F^2$ as a graph functional X_n, we have the expansion

$$\|A\|_F^2 = \binom{n}{2}\hat{p} = T_n(P_2) = \frac{1}{2}S_n(P_2) + \binom{n}{2}p = \frac{1}{2}S_n(P_2) + \binom{n}{2}pS_n(\emptyset). \tag{10}$$

That is, the (squared) Frobenius norm can likewise be expressed in terms of the estimated edge probability $\hat{p}$, by using the raw subgraph count $T_n(P_2)$ or the centered subgraph count $S_n(P_2)$. According to Proposition 2, (10) gives exactly the representation of $X_n = \|A\|_F^2$ in terms of centered subgraph counts in (4) with $a_n(P_2) = \frac{1}{2}$ and $a_n(\emptyset) = \binom{n}{2}p$. Hence, according to Theorem 1, we have

$$\frac{\|A\|_F^2 - \mathbb{E}(\|A\|_F^2)}{\sqrt{\mathbb{V}\mathrm{ar}(\|A\|_F^2)}} = \frac{\frac{1}{2}S_n(P_2)}{\sqrt{\mathbb{V}\mathrm{ar}\left(\frac{1}{2}S_n(P_2) + \binom{n}{2}p^3\right)}} = \frac{S_n(P_2)}{\sqrt{\mathrm{aut}(P_2)n(n-1)p(1-p)}} \xrightarrow{d} \mathcal{N}(0,1).$$

However, as we have

$$\frac{a_n(\emptyset) - \hat{a}_n(\emptyset)}{\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(\|A\|_F^2)}} = \frac{\binom{n}{2}p - \binom{n}{2}\hat{p}}{\sqrt{\mathbb{V}\mathrm{ar}_{H_0}\left(\binom{n}{2}\hat{p}\right)}} = \frac{\frac{1}{2}S_n(P_2)}{\sqrt{\mathbb{V}\mathrm{ar}_{H_0}\left(\frac{1}{2}S_n(P_2)\right)}} \xrightarrow{d} \mathcal{N}(0,1),$$

again, condition (B) in (9) does *not* hold. Moreover, as $\|A\|_F^2 = \widehat{E_{H_0}}(\|A\|_F^2)$ holds, we even have

$$\hat{R}(\|A\|_F^2) = \frac{\|A\|_F^2 - \widehat{\mathbb{E}_{H_0}}(\|A\|_F^2)}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(\|A\|_F^2)}} = 0.$$

Consequently, for $X_n = \|A\|_F^2$ and under the null of an ER graph, in contrast to the infeasible test statistic $R(\|A\|_F^2)$, the feasible goodness-of-fit test statistic $\hat{R}(\|A\|_F^2)$ does not converge to a standard normal distribution. Actually, the feasible test statistic does also not converge to some other nondegenerate distribution, it becomes exactly zero such that this test statistic is not suitable at all for testing H_0 against H_1. But this makes a lot of sense, because just counting the number of edges in a graph G (this is what the (squared) Frobenius norm does), does not allow at all to distinguish between homogeneous and heterogeneous Erdös–Rényi graphs.

This example demonstrates that condition (B) can actually fail for (too) simple graph functionals. In the next subsection, we illustrate the whole machinery of graph functional-based goodness-of-testing and the derivation of asymptotic theory in detail for the example of the degree variance V_n defined in (3).

3.1 | A deep example: degree variance goodness-of-fit testing

Firstly, we would like to note that a goodness-of-fit test for V_n has already been proposed by Ouadah et al. (2020). Eventually, we will see that—although using a different concept of proof—we come to the same result for our asymptotic version which underlines the unified characteristic of our approach that contains this test as a special case. As stated in Corol. 1 in Ouadah et al. (2020), it yields $\mathbb{E}_{H_0}(V_n) = n^{-1}(n-1)(n-2)p(1-p)$ and $\mathbb{V}\mathrm{ar}_{H_0}(V_n) = n^{-3}2(n-1)(n-2)^2p(1-p)(1+(n-6)p(1-p))$. In order to formulate a feasible test, we make use of Proposition 3 above.

Proposition 4. *Let $G \in \mathcal{G}_{\mathrm{ER}}$ be a homogeneous ER graph $\mathcal{G}(n,p)$ with $p = p(n) \xrightarrow[n\to\infty]{} p_0 \in [0,1)$ such that $np \xrightarrow[n\to\infty]{} \infty$. Then, the degree variance V_n standardized by its estimated moments is asymptotically standard normal. That is, we have*

$$\widehat{R}(V_n) = \frac{V_n - \widehat{\mathbb{E}_{H_0}}(V_n)}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(V_n)}} \xrightarrow{d} \mathcal{N}(0,1),$$

where $\widehat{\mathbb{E}_{H_0}}(V_n) = n^{-1}(n-1)(n-2)\widehat{p}(1-\widehat{p})$ and $\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(V_n) = n^{-3}2(n-1)(n-2)^2\widehat{p}(1-\widehat{p})(1+(n-6)\widehat{p}(1-\widehat{p}))$. An asymptotic level-$\alpha$ test for testing the null hypothesis H_0 of an underlying homogeneous Erdös–Rényi graph against heterogeneous alternatives H_1 in (1) is then given by the decision rule

$$\phi_\alpha(V_n) = \begin{cases} 1, & \text{if } \left|\widehat{R}(V_n)\right| > z_{1-\alpha/2}, \\ 0, & \text{else.} \end{cases}$$

In Section 2.2.1, we already argued that V_n is a graph functional. In order to find its representation following Proposition 2, we consider the centered version $V_n - \mathbb{E}_{H_0}(V_n)$, where $\mathbb{E}_{H_0}(V_n) = n^{-1}(n-1)(n-2)p(1-p)$, and use the Hoeffding decomposition (Hoeffding, 1992; Lee, 2019) as in Ouadah et al. (2020). Then, under the null of a homogeneous ER graph, we get

$$\begin{aligned} V_n - \mathbb{E}_{H_0}(V_n) = {} & \frac{2(n-2)}{n^2}(1-2p)\sum_{1\le i<j\le n}\widetilde{A}_{ij} + \frac{2(n-4)}{n^2}\sum_{1\le i<j<k\le n}\left(\widetilde{A}_{ij}\widetilde{A}_{ik} + \widetilde{A}_{ij}\widetilde{A}_{jk} + \widetilde{A}_{ik}\widetilde{A}_{jk}\right) \\ & - \frac{8}{n^2}\sum_{1\le i<j<k<l\le n}\left(\widetilde{A}_{ij}\widetilde{A}_{kl} + \widetilde{A}_{ik}\widetilde{A}_{jl} + \widetilde{A}_{il}\widetilde{A}_{jk}\right) =: V_n^{(A)} + V_n^{(B)} + V_n^{(C)}, \end{aligned} \tag{11}$$

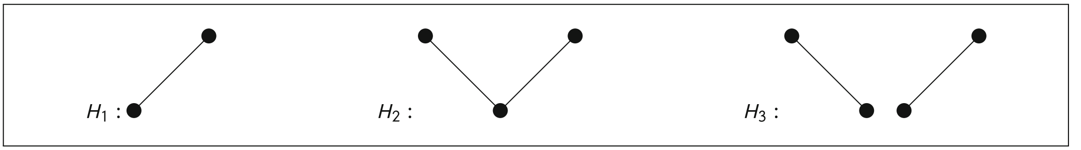

FIGURE 2 Subgraphs in the decomposition of the degree variance V_n, where $H_1 = P_2$ and $H_2 = P_3$.

where $\widetilde{A}_{ij} = A_{ij} - p$ are the centered entries of the adjacency matrix. While V_n does not depend on the usually unknown probability p, $\mathbb{E}_{H_0}(V_n)$ does. However, according to Proposition 4, we can use $\widehat{\mathbb{E}}_{H_0}(V_n)$, which is obtained by replacing p by $\hat{p}$ in $\mathbb{E}_{H_0}(V_n)$, leading to the same asymptotics. The components of the above decomposition are uncorrelated. This simplifies the derivation of the moments of V_n, especially of its variance, leading to

$$\mathbb{V}\mathrm{ar}_{H_0}(V_n) = \mathbb{V}\mathrm{ar}_{H_0}(V_n^{(A)}) + \mathbb{V}\mathrm{ar}_{H_0}(V_n^{(B)}) + \mathbb{V}\mathrm{ar}_{H_0}(V_n^{(C)}).$$

Consequently, the variance $\mathbb{V}\mathrm{ar}_{H_0}(V_n)$ decomposes into three parts $\mathbb{V}\mathrm{ar}_{H_0}(V_n^{(A)})$, $\mathbb{V}\mathrm{ar}_{H_0}(V_n^{(B)})$ and $\mathbb{V}\mathrm{ar}_{H_0}(V_n^{(C)})$ that correspond to the components that are depicted in Figure 2 and can be calculated by

$$\begin{aligned}\mathbb{V}\mathrm{ar}_{H_0}(V_n) = {} & \frac{2(n-1)(n-2)^2}{n^3}(1-2p)^2 p(1-p) + \frac{2(n-1)(n-2)(n-4)^2}{n^3}(p(1-p))^2 \\ & + \frac{8(n-1)(n-2)(n-3)}{n^3}(p(1-p))^2.\end{aligned}$$

For the derivation of the limiting normal distribution in Proposition 4, we assumed $np \to \infty$. In view of Theorem 1, the weaker assumption $np^{2/3} \to \infty$ seems to be sufficient. However, a closer inspection of the three parts of $\mathbb{V}\mathrm{ar}(V_n)$ yields that the first one is of order $O(p)$, the second one $O(np^2)$ and the third one $O(p^2)$. Hence, if $np^{2/3} \to \infty$, we have to distinguish three cases: if $np \to \infty$, the second term $V_n^{(B)}$ becomes the leading term. If $np = O(1)$, both the first and the second term, that is $V_n^{(A)}$ and $V_n^{(B)}$ become the leading terms. If $np \to 0$, the first term $V_n^{(A)}$ becomes the leading term. That is, whenever, np does not diverge to ∞, the degree variance V_n does asymptotically depend on the first term which corresponds to $S_n(P_2)$ as described below, which is demonstrated to not lead to a useful goodness-of-fit test statistic.

Contrary to the presentations in Ouadah et al. (2020), it is not $V_n^{(A)}$, but the second part $V_n^{(B)}$, that makes the main contribution to the variance of V_n, if $np \to \infty$ as $n \to \infty$. Intuitively, this allocation of the variances makes sense. The information contained in part $V_n^{(A)}$ is the centered number of edges in the observed network. This does not tell us much about the structure of the graph. However, part $V_n^{(B)}$ counts the number of edges that share a common node and are active—it is a rescaled version of the centered count of two-stars P_3 (paths on three vertices with length two). This part of the decomposition contains information on the local structure of the observed network. Part $V_n^{(C)}$ counts the number of pairs of *disjoint* edges that are both active, again an information that does not tell us much about the global or local structure. Hence, it is reasonable that $V_n^{(B)}$ contains most of the information of the network and for that explains most of the variance. The decomposition into $V_n^{(A)}$, $V_n^{(B)}$, and $V_n^{(C)}$ is the linear combination of centered counts of the substructures $H_1 = P_3$, $H_2 = P_2$ and H_3, illustrated in Figure 2. Thus, alternatively, we can represent (11) as

$$V_n - \mathbb{E}(V_n) = \frac{2(n-2)}{n^2}(1-2p)\frac{1}{\mathrm{aut}(H_1)}S_n(H_1) + \frac{2(n-4)}{n^2}\frac{1}{\mathrm{aut}(H_2)}S_n(H_2) - \frac{8}{n^2}\frac{1}{\mathrm{aut}(H_3)}S_n(H_3)$$
$$= \frac{(n-2)}{n^2}(1-2p)S_n(H_1) + \frac{(n-4)}{n^2}S_n(H_2) - \frac{1}{n^2}S_n(H_3).$$

Note that compared to (11), we need the prefactors $\frac{1}{\mathrm{aut}(H_i)}$ for $i \in \{1,2,3\}$ because the definition of $S_n(H)$ considers all automorphisms of a given substructure, whereas the sum indices in (11) include each node combination only once. Embedded into Proposition 2, as $\mathrm{aut}(H_1) = 2$, $\mathrm{aut}(H_2) = 2$, and $\mathrm{aut}(H_3) = 8$, we have

$$a_n(\emptyset) = \mathbb{E}_{H_0}(V_n),\ a_n(H_1) = \frac{(n-2)}{n^2}(1-2p),\ a_n(H_2) = \frac{(n-4)}{n^2},\ \text{and } a_n(H_3) = -\frac{1}{n^2}.$$

as well as

$$\mathbb{V}\mathrm{ar}_{H_0}(S_n(H_1)) = 2n(n-1)p(1-p), \mathbb{V}\mathrm{ar}_{H_0}(S_n(H_2)) = 2n(n-1)(n-2)(p(1-p))^2, \text{and}$$
$$\mathbb{V}\mathrm{ar}_{H_0}(S_n(H_3)) = 8n(n-1)(n-2)(n-3)(p(1-p))^2.$$

Eventually, Theorem 1 can be applied. For this, we need to check if its assumptions are fulfilled. Firstly, a dominating family of connected graphs is required. As $np \to \infty$, both $a_n^2(H_1)\mathbb{V}\mathrm{ar}_{H_0}(S_n(H_1))$ and $a_n^2(H_3)\mathbb{V}\mathrm{ar}_{H_0}(S_n(H_3))$ are of lower order than $a_n^2(H_2)\mathbb{V}\mathrm{ar}_{H_0}(S_n(H_2))$, we set $\mathcal{H}_0 := \{H_2\}$. Then, we can verify that $\mathcal{H}_0$ is a dominating family of connected graphs for V_n by checking the corresponding condition. Indeed, we have

$$\frac{\mathbb{V}\mathrm{ar}_{H_0}(V_n)}{\sum_{H\in\mathscr{H}_0} a_n^2(H)\mathbb{V}\mathrm{ar}_{H_0}(S_n(H))}$$
$$= \frac{\frac{2(n-1)(n-2)^2}{n^3}p(1-p)(1+(n-6)p(1-p))}{2\frac{(n-4)^2}{n^4}n(n-1)(n-2)(p(1-p))^2} = \frac{n-2+(n^2-8n+12)}{(n^2-8n+16)} \xrightarrow[n\to\infty]{} 1.$$

Consequently, the variance of V_n is explained by the dominating family $\mathcal{H}_0$. Asymptotically, it is solely driven by part $V_n^{(B)}$ of the Hoeffding decomposition. As a second step, we need to check whether the coefficient

$$b(H_2) = \sup_{n\geq 4}\frac{n^{3/2}p\ \ \frac{(n-4)}{n^2}}{\sqrt{\frac{2(n-1)(n-2)^2}{n^3}p(1-p)(1+(n-6)p(1-p))}},$$

is finite. Indeed, the finiteness of the supremum above is implied by the convergence

$$\lim_{n\to\infty}\frac{(n^2-8n+16)n^2p}{2(n-1)(n^2-4n+4)(1-p)(1+(n-6)p(1-p))} = \frac{1}{2(1-p_0)^2} < \infty \quad \forall\, p_0 \in [0,1),$$

Furthermore, as the dominating family $\mathcal{H}_0$ is finite, this obviously leads to

$$b^2(H_2)\mathrm{aut}(H_2) = \frac{1}{(1-p_0)^2} < \infty \quad \forall\, p_0 \in [0,1). \tag{12}$$

Thus, from (11) to (12), all conditions of Theorem 1 are fulfilled, yielding the asymptotic normality of $V_n - \mathbb{E}(V_n)$. That is, for $n \to \infty$, we have

$$\frac{V_n - \mathbb{E}(V_n)}{\sqrt{\mathbb{V}\mathrm{ar}(V_n)}} \xrightarrow{d} \mathcal{N}(0,1). \tag{13}$$

In particular, for $c_n^2 = (np^2(1-p)^2)^{-1}$, we have $c_n^2 \mathbb{V}\mathrm{ar}(X_n) \to 2$ and $c_n(X_n - \mathbb{E}(X_n)) \xrightarrow{d} \mathcal{N}(0,2)$. As mentioned above, the true underlying parameter p is usually not known in practice. An implementable result for the shown asymptotics is given in Proposition 4. Details on the embedding of $\hat{p}$ in this context can be found in Appendix A.

3.2 | Goodness-of-fit testing based on centered subgraph counts

In view of the general decomposition (4) for arbitrary graph functionals X_n given in Proposition 2, it is natural to make use of centered subgraph counts $S_n(H)$ as building blocks for our unified goodness-of-fit testing approach. However, although the $S_n(H)$'s are graph functionals according to Definition 3, they are not yet feasible for goodness-of-fit testing as they require the knowledge of the edge probability p. Consequently, the $S_n(H)$'s cannot be computed from the adjacency matrix A of a graph G alone. As it is natural to get feasible versions of the $S_n(H)$'s by replacing p by $\hat{p}$, we define

$$\hat{S}_n(H) := \sum_{\varphi} \prod_{e \in E(\varphi(H))} (\mathbb{1}\{e \in E(G)\} - \hat{p}),$$

Moreover, under the assumptions of Theorem 1 (see also Thm. 6.43 in Janson et al., 2011), it is guaranteed that $S_n(H)$ converges to a standard normal distribution $\mathcal{N}(0,1)$ when divided by the square root of its variance $\mathbb{V}\mathrm{ar}_{H_0}(S_n(H))$. In particular, this holds for any nonnull subgraph H. However, according to Example 1, in addition to the null graph $\emptyset$, also P_2 causes problems and is not suitable for goodness-of-fit testing, because it is a too simplistic subgraph structure. Fortunately, as stated in the following proposition, all subgraph structures H that are more complex than P_2 allow to use $\hat{S}_n(H)$ divided by $\sqrt{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(S_n(H))}$ instead of $S_n(H)$ divided by $\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(S_n(H))}$ without changing the limiting distribution.

Proposition 5. *Let $G \in \mathcal{G}_{\mathrm{ER}}$ be a homogeneous ER graph $\mathcal{G}(n,p)$ with $p = p(n) \xrightarrow[n\to\infty]{} p_0 \in [0,1)$. Further, let H be a nonnull connected subgraph with $H \neq P_2$ and assume $np^{r(H)} \xrightarrow[n\to\infty]{} \infty$. Then, we have*

$$\hat{S}_n(H) = \sum_{\substack{J \subseteq H \\ J \text{ induced}}} \left(\prod_{h=n(H)-k(J)}^{n(H)-1} (n-h) \right) S_n(\tilde{J})(p-\hat{p})^{m(H)-m(J)}, \tag{14}$$

where $k(J)$ denotes the number of isolated vertices in J, $\tilde{J}$ denotes the graph obtained from J after removing all isolated vertices. This leads to

$$\frac{\hat{S}_n(H)}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(S_n(H))}} = \frac{\hat{S}_n(H)}{\sqrt{\mathrm{aut}(H)(n)_{n(H)}(\hat{p}(1-\hat{p}))^{m(H)}}} \xrightarrow{d} \mathcal{N}(0,1).$$

We then get an asymptotic level-α test for testing the null hypothesis H_0 of an underlying homogeneous Erdös–Rényi graph against heterogeneous alternatives H_1 in (7) by the decision rule

$$\phi_\alpha(S_n(H)) := \begin{cases} 1, & \text{if } \left| \frac{\widehat{S}_n(H)}{\sqrt{\operatorname{aut}(H)(n)_{n(H)}(\widehat{p}(1-\widehat{p}))^{m(H)}}} \right| > z_{1-\alpha/2}, \\ 0, & \text{else.} \end{cases}$$

That is, according to Proposition 5, using feasible centered subgraph counts $\widehat{S}_n(H)$ instead of infeasible ones $S_n(H)$ and dividing by the square root of the estimated variance of $S_n(H)$ (under H_0) does not change the asymptotic distribution, which is still standard normal $\mathcal{N}(0,1)$. Note that we do not subtract an estimated mean $\widehat{\mathbb{E}_{H_0}}(S_n(H))$ here, because $\mathbb{E}(S_n(H)) = 0$ holds. Instead, we are adjusting $S_n(H)$ itself to get $\widehat{S}_n(H)$, which then can be computed from A alone. This leads to the formulation of asymptotic goodness-of-fit tests for statistics of shape $S_n(H)$.

We illustrate the above result for feasible centered subgraph counts $\widehat{S}_n(H)$ for the special case of a triangle $H = C_3$, that is, $\widehat{S}_n(C_3)$, in the following example. Raw subgraph counts $T_n(H)$ will be discussed in Section 3.3, where we also give an example of the corresponding raw subgraph count $T_n(C_3)$.

Example 4 (Asymptotics for centered counts of triangles). Under the assumptions of Theorem 1, for the infeasible centered subgraph count $S_n(C_3)$, we have

$$\frac{S_n(C_3) - \mathbb{E}_{H_0}(C_3)}{\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(C_3)}} = \frac{S_n(C_3)}{\sqrt{\operatorname{aut}(C_3)n_{n(C_3)}(p(1-p))^{m(C_3)}}} = \frac{S_n(C_3)}{\sqrt{6n(n-1)(n-2)(p(1-p))^3}} \xrightarrow{d} \mathcal{N}(0,1),$$

and, according to Proposition 5, for feasible subgraph counts $\widehat{S}_n(C_3)$, we have the decomposition (see (A1) in the Appendix)

$$\begin{aligned} \widehat{S}_n(C_3) &= \sum_{\substack{J \subseteq C_3 \\ J \text{induced}}} \left(\prod_{h=n(C_3)-k(J)}^{n(C_3)-1} (n-h) \right) S_n(\widetilde{J})(p-\widehat{p})^{m(C_3)-m(J)} \\ &= S_n(C_3) + 3\left(\prod_{h=n(C_3)-0}^{n(C_3)-1} (n-h) \right) S_n(P_3)(p-\widehat{p}) + 3\left(\prod_{h=n(C_3)-1}^{n(C_3)-1} (n-h) \right) S_n(P_2)(p-\widehat{p})^2 \\ &\quad + \left(\prod_{h=0}^{n(C_3)-1} (n-h) \right)(p-\widehat{p})^3 \\ &= S_n(C_3) + 3S_n(P_3)(p-\widehat{p}) + 3(n-2)S_n(P_2)(p-\widehat{p})^2 + n(n-1)(n-2)(p-\widehat{p})^3. \end{aligned}$$

Here, we have used that there are also three ways to drop one edge from C_3 leading to the same substructure and there are three ways to drop two edges leading to the same substructure, respectively. When dropping two edges, we end up with a graph consisting of three nodes and one edge such that one node is isolated. Hence, $k(J) = 1$ and $\widetilde{J} = P_2$ in this case, where P_2 consist of two vertices and one edge. Dropping all edges leads to three isolated vertices, that is, $k(J) = 3$ and $\widetilde{J}$ becomes the null graph $\varnothing$ leading to $S_n(\varnothing) = 1$. When dividing by $\sqrt{\operatorname{aut}(C_3)n_{n(C_3)}(\widehat{p}(1-\widehat{p}))^{m(C_3)}} = \sqrt{6n(n-1)(n-2)(\widehat{p}(1-\widehat{p}))^3}$, we get

$$\frac{3S_n(P_3)(p-\hat{p})}{\sqrt{6n(n-1)(n-2)(\hat{p}(1-\hat{p}))^3}} = O_P\left(\frac{\sqrt{n^3(p(1-p))^2(\frac{\sqrt{p}}{n})}}{\sqrt{n^3(p(1-p))^3}}\right)$$
$$= O_P\left(\frac{1}{n\sqrt{1-p}}\right) = O_P\left(\frac{1}{n}\right) = o_P(1),$$
$$\frac{3(n-2)S_n(P_2)(p-\hat{p})^2}{\sqrt{6n(n-1)(n-2)(\hat{p}(1-\hat{p}))^3}} = O_P\left(\frac{n\sqrt{n^2(p(1-p))}(\frac{\sqrt{p}}{n})^2}{\sqrt{n^3(p(1-p))^3}}\right)$$
$$= O_P\left(\frac{1}{(1-p)^2 n^{3/2}}\right) = O_P\left(\frac{1}{n^{3/2}}\right) = o_P(1),$$
$$\frac{n(n-1)(n-2)(p-\hat{p})^3}{\sqrt{6n(n-1)(n-2)(\hat{p}(1-\hat{p}))^3}} = O_P\left(\frac{n^3(\frac{\sqrt{p}}{n})^3}{\sqrt{n^3(p(1-p))^3}}\right)$$
$$= O_P\left(\frac{1}{n^{3/2}(1-p)^{3/2}}\right) = O_P\left(\frac{1}{n^{3/2}}\right) = o_P(1).$$

Altogether, according to Proposition 5, for the feasible centered subgraph count $\hat{S}_n(C_3)$, we have

$$\frac{\hat{S}_n(C_3)}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}}_{H_0}(C_3)}} = \frac{\hat{S}_n(C_3)}{\sqrt{6n(n-1)(n-2)(\hat{p}(1-\hat{p}))^3}} \xrightarrow{d} \mathcal{N}(0,1).$$

Hence, $S_n(C_3)$ divided by $\sqrt{\mathrm{aut}(C_3)(n)_3(p(1-p))^3}$ and $\hat{S}_n(C_3)$ divided by $\sqrt{\mathrm{aut}(C_3)(n)_3(\hat{p}(1-\hat{p}))^3}$ converge to the same limiting (standard normal) distribution.

3.3 | Goodness-of-fit testing based on raw subgraph counts

In contrast to centered subgraph counts $S_n(H)$, which form the building blocks of general graph functionals according to (4), but are not feasible as they require the knowledge of p, *raw* subgraph counts $T_n(H)$ also appear to be a natural and, in particular, feasible goodness-of-fit test statistic. Moreover, as the $T_n(H)$'s are graph functionals by construction, under the assumptions of Theorem 1, we have

$$\frac{T_n(H) - \mathbb{E}_{H_0}(T_n(H))}{\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(T_n(H))}} \xrightarrow{d} \mathcal{N}(0,1).$$

However, both $\mathbb{E}_{H_0}(T_n(H))$ and $\mathbb{V}\mathrm{ar}_{H_0}(T_n(H))$ will generally depend on the unknown parameter p and have to be replaced by suitable estimators $\widehat{\mathbb{E}_{H_0}}(T_n(H))$ and $\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(T_n(H))$ obtained by replacing p by $\hat{p}$. But, as the $T_n(H)$'s are graph functionals as well, for studying their properties, we can expand it in terms of certain centered subgraph counts using (4). Hence, following exactly the approach to decompose $\hat{S}_n(H)$ in terms of (lower order) centered subgraph counts $S_n(\tilde{J})$'s leading to (14) in Proposition 5, we get the following result for raw subgraph counts.

Proposition 6. *Let $G \in \mathcal{G}_{\mathrm{ER}}$ be a homogeneous ER graph $\mathcal{G}(n,p)$ and let H be a nonnull connected subgraph. Then, we have*

$$T_n(H) = \frac{1}{\mathrm{aut}(H)} \sum_{\substack{J \subseteq H \\ J \text{ induced}}} \left(\prod_{h=n(H)-k(J)}^{n(H)-1} (n-h) \right) S_n(\widetilde{J}) p^{m(H)-m(J)}, \tag{15}$$

where $k(J)$ denotes the number of isolated vertices in J, $\widetilde{J}$ denotes the graph obtained from J after removing all isolated vertices.

For the case of $H = C_3$, we illustrate the consequences of (15) for the dominating family of subgraphs in the expansion (4) of the graph functional $T_n(C_3)$.

Example 5 (Asymptotics for raw counts of triangles). If the assumptions of Theorem 1 hold, for the raw subgraph count $T_n(C_3)$, we have

$$R(T_n(C_3)) = \frac{T_n(C_3) - \mathbb{E}_{H_0}(T_n(C_3))}{\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(T_n(C_3))}} \xrightarrow{d} \mathcal{N}(0,1).$$

Hence, to check the conditions imposed in Theorem 1, we have to study the decomposition of $T_n(C_3)$ in terms of $S_n(\widetilde{J})$'s. According to Proposition 6, for $H = C_3$, we get

$$\begin{aligned}
T_n(C_3) &= \frac{1}{\mathrm{aut}(C_3)} \sum_{J \subseteq C_3} \left(\prod_{h=n(C_3)-k(J)}^{n(C_3)-1} (n-h) \right) S_n(\widetilde{J}) p^{m(C_3)-m(J)} \\
&= \frac{1}{\mathrm{aut}(C_3)} \left(S_n(J) + 3 \left(\prod_{h=n(C_3)-0}^{n(C_3)-1} (n-h) \right) S_n(P_3) p + 3 \left(\prod_{h=n(C_3)-1}^{n(C_3)-1} (n-h) \right) S_n(P_2) p^2 \right. \\
&\quad \left. + \left(\prod_{h=0}^{n(C_3)-1} (n-h) \right) p^3 \right) \\
&= \frac{1}{\mathrm{aut}(C_3)} \left(S_n(C_3) + 3 S_n(P_3) p + 3(n-2) S_n(P_2) p^2 + n(n-1)(n-2) p^3 \right).
\end{aligned}$$

Consequently, using Proposition 1, for the variance of $T_n(C_3)$, we get

$$\begin{aligned}
&\mathbb{V}\mathrm{ar}_{H_0}(T_n(C_3)) \\
&= \frac{1}{\mathrm{aut}(C_3)^2} \left(\mathbb{V}\mathrm{ar}_{H_0}(S_n(C_3)) + 9 \mathbb{V}\mathrm{ar}_{H_0}(S_n(P_3)) p^2 + 9(n-2)^2 \mathbb{V}\mathrm{ar}_{H_0}(S_n(P_2)) p^4 \right) \\
&= \frac{1}{\mathrm{aut}(C_3)^2} \left(\mathrm{aut}(C_3)(n)_3 (p(1-p))^3 + 9 \mathrm{aut}(P_3)(n)_3 (p(1-p))^2 p^2 \right. \\
&\quad \left. + 9(n-2)^2 (n)_2 \mathrm{aut}(P_2) p(1-p) p^4 \right) \\
&= \frac{(n)_3}{\mathrm{aut}(C_3)} p^3 (1-p)^3 + \frac{9 \mathrm{aut}(P_3)(n)_3}{\mathrm{aut}(C_3)^2} p^4 (1-p)^2
\end{aligned}$$

$$+\frac{9(n-2)^2(n)_2\mathrm{aut}(P_2)}{\mathrm{aut}(C_3)^2}p^5(1-p)$$

$$=\binom{n}{3}p^3(1-p)^3+3\binom{n}{3}p^4(1-p)^2+3(n-2)\binom{n}{3}p^5(1-p)$$

$$=\mathbb{V}\mathrm{ar}^{(1)}_{H_0}(T_n(C_3))+\mathbb{V}\mathrm{ar}^{(2)}_{H_0}(T_n(C_3))+\mathbb{V}\mathrm{ar}^{(3)}_{H_0}(T_n(C_3)),$$

with an obvious notation for $\mathbb{V}\mathrm{ar}^{(i)}_{H_0}(T_n(C_3))$, $i=1,2,3$. Taking a closer look at these three terms, if $p=p(n)\to p_0\in[0,1)$, we see that $\mathbb{V}\mathrm{ar}^{(1)}_{H_0}(T_n(C_3))=O(n^3p^3)$, $\mathbb{V}\mathrm{ar}^{(2)}_{H_0}(T_n(C_3))=O(n^3p^4)$ and $\mathbb{V}\mathrm{ar}^{(3)}_{H_0}(T_n(C_3))=O(n^4p^5)$. Hence, there are different cases to distinguish:

(i) If $p=p(n)\to p_0\in(0,1)$, we have that $\mathbb{V}\mathrm{ar}^{(3)}_{H_0}(T_n(C_3))$, that is, $S_n(P_2)$ dominates.

(ii) If $p=p(n)\to 0$, we have that $\mathbb{V}\mathrm{ar}^{(2)}_{H_0}(T_n(C_3))=o(\mathbb{V}\mathrm{ar}^{(1)}_{H_0}(T_n(C_3)))$, that is, $S_n(P_3)$ is dominated by $S_n(C_3)$.

- (iia) If $p=p(n)\to 0$ such that $np^2\to 0$, we have $\mathbb{V}\mathrm{ar}^{(3)}_{H_0}(T_n(C_3))=o(\mathbb{V}\mathrm{ar}^{(1)}_{H_0}(T_n(C_3)))$ and $S_n(C_3)$ dominates.
- (iib) If $p=p(n)\to 0$ such that $np^2\to K$ for some constant $K>0$, we have that $\mathbb{V}\mathrm{ar}^{(3)}_{H_0}(T_n(C_3))$ and $\mathbb{V}\mathrm{ar}^{(1)}_{H_0}(T_n(C_3))$ are of the same order and both $S_n(C_3)$ and $S_n(P_2)$ dominate.
- (iic) If $p=p(n)\to 0$ such that $np^2\to\infty$, we have that $\mathbb{V}\mathrm{ar}^{(1)}_{H_0}(T_n(C_3))=o(\mathbb{V}\mathrm{ar}^{(3)}_{H_0}(T_n(C_3)))$ and $S_n(P_2)$ dominates.

Note that the condition $np^{r(H)}\to\infty$ as $n\to\infty$ imposed in Theorem 1 for the dominating family of subgraphs $H\in\mathcal{H}$ is trivially fulfilled in cases (i), (iib), and (iic), while case (iia) requires also $np^{r(C_3)}=np\to\infty$. Moreover, in view of Example 1, cases (i) and (iic) lead to a goodness-of-fit test dominated by P_2, which is *not* suitable at all to detect deviations from the H_0 of a homogeneous Erdös–Rényi graph. Case (iib) is a mixed case that makes use of P_2 and C_3. Hence, (iib) will be not efficient, while case (iia) has the largest potential to detect deviations from H_0.

However, at this point, it is important to note that $R(T_n(C_3))$ above defines an *infeasible* test statistic as it requires the knowledge of p, because both $\mathbb{E}_{H_0}(T_n(C_3))=\binom{n}{3}p^3$ and $\mathbb{V}\mathrm{ar}_{H_0}(T_n(C_3))$ as derived above are functions of p. That is, according to our construction principle, instead, we have to replace $\mathbb{E}_{H_0}(T_n(C_3))$ and $\mathbb{V}\mathrm{ar}_{H_0}(T_n(C_3))$ by $\widehat{\mathbb{E}_{H_0}}(T_n(C_3))$ and $\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(T_n(C_3))$, respectively.

Example 5 (Asymptotics for raw counts of triangles). Based on the raw subgraph count $T_n(C_3)$, a feasible goodness-of test statistic is obtained by

$$\widehat{R}(T_n(C_3))=\frac{T_n(C_3)-\widehat{\mathbb{E}_{H_0}}(T_n(C_3))}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(T_n(C_3))}}$$

$$=\frac{T_n(C_3)-\binom{n}{3}\widehat{p}^3}{\sqrt{\binom{n}{3}\widehat{p}^3(1-\widehat{p})^3+3\binom{n}{3}\widehat{p}^4(1-\widehat{p})^2+3(n-2)\binom{n}{3}\widehat{p}^5(1-\widehat{p})}}.$$

Let us have a closer look at the numerator. First, using (10), we can rewrite $\binom{n}{3}\hat{p}^3$ to get

$$\binom{n}{3}\hat{p}^3 = \binom{n}{3}\left(\frac{1}{2\binom{n}{2}}S_n(P_2) + p\right)^3$$
$$= \binom{n}{3}\left(\left(\frac{1}{2\binom{n}{2}}\right)^3 S_n^3(P_2) + 3\left(\frac{1}{2\binom{n}{2}}\right)^2 S_n^2(P_2)p + 3\frac{1}{2\binom{n}{2}}S_n(P_2)p^2 + p^3\right)$$
$$= \frac{\binom{n}{3}}{\left(2\binom{n}{2}\right)^3}S_n^3(P_2) + \frac{3\binom{n}{3}}{\left(2\binom{n}{2}\right)^2}S_n^2(P_2)p + \frac{3\binom{n}{3}}{2\binom{n}{2}}S_n(P_2)p^2 + \binom{n}{3}p^3.$$

Subtracting the last right-hand side from $T_n(C_3)$ and plugging-in the decomposition for $T_n(C_3)$, leads to

$$T_n(C_3) - \widehat{\mathbb{E}_{H_0}}(T_n(C_3)) = \frac{1}{\text{aut}(C_3)}\left(S_n(C_3) + 3S_n(P_3)p + 3(n-2)S_n(P_2)p^2 + n(n-1)(n-2)p^3\right)$$
$$- \left(\frac{\binom{n}{3}}{\left(2\binom{n}{2}\right)^3}S_n^3(P_2) + \frac{3\binom{n}{3}}{\left(2\binom{n}{2}\right)^2}S_n^2(P_2)p + \frac{3\binom{n}{3}}{2\binom{n}{2}}S_n(P_2)p^2 + \binom{n}{3}p^3\right)$$
$$= \frac{1}{\text{aut}(C_3)}(S_n(C_3) + 3S_n(P_3)p) - \left(\frac{\binom{n}{3}}{\left(2\binom{n}{2}\right)^3}S_n^3(P_2) + \frac{3\binom{n}{3}}{\left(2\binom{n}{2}\right)^2}S_n^2(P_2)p\right).$$

The latter two terms on the last right hand are of order $p^{3/2}$ and np^2, which are both slower than the orders of the first two terms, which are $(np)^{3/2}$ and $n^{3/2}p^2$. Hence, the last expression is of leading order $(np)^{3/2}$. Altogether, picking-up the different cases introduced in Example 5, we get

$$\frac{T_n(C_3) - \widehat{E_{H_0}}(T_n(C_3))}{\sqrt{\widehat{\mathbb{V}\text{ar}_{H_0}}(T_n(C_3))}} = \frac{T_n(C_3) - \binom{n}{3}\hat{p}^3}{\sqrt{\binom{n}{3}\hat{p}^3(1-\hat{p})^3 + 3\binom{n}{3}\hat{p}^4(1-\hat{p})^2 + 3(n-2)\binom{n}{3}\hat{p}^5(1-\hat{p})}} \xrightarrow{P} 0,$$

for (i) and (iic), because $(np)^{3/2}/n^2p^{5/2} = \sqrt{\frac{1}{np^2}} \to 0$ in both cases. For (iib), we get

$$\frac{T_n(C_3) - \widehat{E_{H_0}}(T_n(C_3))}{\sqrt{\widehat{\mathbb{V}\text{ar}_{H_0}}(T_n(C_3))}} = \frac{T_n(C_3) - \binom{n}{3}\hat{p}^3}{\sqrt{\binom{n}{3}\hat{p}^3(1-\hat{p})^3 + 3\binom{n}{3}\hat{p}^4(1-\hat{p})^2 + 3(n-2)\binom{n}{3}\hat{p}^5(1-\hat{p})}}$$
$$\xrightarrow{d} \mathcal{N}(0, \sigma^2),$$

where

$$\sigma^2 := \lim_{n\to\infty} \frac{\binom{n}{3}p^3(1-p)^3 + 3\binom{n}{3}p^4(1-p)^2}{\binom{n}{3}p^3(1-p)^3 + 3\binom{n}{3}p^4(1-p)^2 + 3(n-2)\binom{n}{3}p^5(1-p)} = \frac{1}{1+3K},$$

where $K = \lim_{n\to\infty} np^2$. Finally, for (iia), we get

$$\frac{T_n(C_3) - \widehat{E_{H_0}}(T_n(C_3))}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(T_n(C_3))}} = \frac{T_n(C_3) - \binom{n}{3}\hat{p}^3}{\sqrt{\binom{n}{3}\hat{p}^3(1-\hat{p})^3 + 3\binom{n}{3}\hat{p}^4(1-\hat{p})^2 + 3(n-2)\binom{n}{3}\hat{p}^5(1-\hat{p})}}$$
$$\xrightarrow{d} \mathcal{N}(0,1).$$

Hence, using standard normal quantiles for implementing a goodness-of-fit test based on $T_n(C_3)$ will be asymptotically valid only for the case of $p = p(n) \to 0$ such that $np^2 \to 0$. That is, in the case of (iia), where $S_n(C_3)$ dominates. This observation coincides with the findings in Ex. 6.50 in Janson et al. (2011).

In view of the examples discussed above and the assumptions (A) and (B) imposed in Proposition 3, the construction principle of goodness-of-fit tests is flexible and allows for a unified treatment based on the theory on graph functionals gathered in Section 2.2. Nevertheless, one has to be careful as condition (B) may be hurt such that $\hat{R}(X_n)$ turns out to have a different asymptotic behavior than $R(X_n)$, which may lead to tests of incorrect size or tests that have (asymptotically) no power at all.

3.4 | Power analysis for SBM alternatives

For any given graph functional X_n, the feasible test $\hat{R}(X_n)$ (that makes use of standard normal quantiles) requires conditions (A) and (B) in (9) to hold under H_0. The same test will have non-trivial power for (certain) alternatives under H_1, if

$$\left| \frac{\mathbb{E}_{H_1}(X_n) - \widehat{\mathbb{E}_{H_0}}(X_n)}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(X_n)}} \right| \xrightarrow{P} \infty \quad \text{as} \quad n \to \infty,$$

holds. As the class of graph functionals is very broad, there are a lot of different possible test statistics to choose from. Obviously, each possible graph functional might be sensitive to different network information and thus might be more or less suitable for a particular data scenario. On the one hand, this leads to a large flexibility in constructing goodness-of-fit tests for various application fields. On the other hand, the derivation of general consistency results for a goodness-of-fit test $\hat{R}(X_n)$ for arbitrary graph functionals X_n and general alternatives under H_1 will not be possible. However, the question of which graph functional to use in which situation is pretty crucial in practice. Because general statements are obviously very difficult to derive, we provide a concrete theory-driven analysis for centered as well as raw subgraph counts that rely on triangles, denoted as C_3, and two-stars, denoted as P_3, below. Further examples are then investigated in a simulation study in Section 5.

For now, to generate stochastic networks under the alternative, we consider the popular case of SBMs that particularly plays an important role in modeling social network behaviors and patterns. In general, it assumes that the set of n nodes can be divided into K blocks. Typically, nodes of the same block have a higher probability to share an edge than nodes of different blocks. Hence, SBMs are low-parametrized special cases of $\mathcal{G}_{\mathrm{HER}}(n)$ models. Throughout this subsection, to make explicit calculations and derivations tractable, we consider only SBMs with $K = 2$ blocks, equal block sizes (i.e., $n/2$), and equal intra-group probabilities p_{intra} for both blocks. The edge probability between the groups are denoted by p_{inter} such that $p_{\mathrm{intra}} \geq p_{\mathrm{inter}}$. For $p_{\mathrm{intra}} = p_{\mathrm{inter}}$, the SBM becomes a $\mathcal{G}_{\mathrm{ER}}(n)$ model. To rule out detection just based on the total number of edges, we generate models from both model classes such that their mean connectivity parameters p_{mean} (which are allowed to depend on n) coincide. Note that, obviously, $p_{\mathrm{mean}} = p$ for the $\mathcal{G}_{\mathrm{ER}}(n)$ model class. Let us first focus on the performance of C_3.

Theorem 2 (Power of $S_n(C_3)$ and $T_n(C_3)$ under fixed and local alternatives). *Suppose the network G is either generated by a homogeneous ER model $\mathcal{G}(n,p)$ with edge probability p or by an SBM with $K = 2$ blocks, equal block sizes $n/2$, and equal intra-group probability p_{intra} for both blocks such that*

$$p_{\mathrm{intra}} - p_{\mathrm{inter}} = \epsilon \quad \text{and} \quad p_{\mathrm{mean}} = p,$$

where

$$p_{\mathrm{mean}} := \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} p_{ij} = \frac{2\binom{n/2}{2} p_{\mathrm{intra}} + \left(\frac{n}{2}\right)^2 p_{\mathrm{inter}}}{\binom{n}{2}} = \frac{n-2}{2n-2} p_{\mathrm{intra}} + \frac{n}{2n-2} p_{\mathrm{inter}},$$

denotes the mean edge probability (mean connectivity) p_{mean} of the SBM. Further, let $\mathbb{E}_{\mathrm{ER}}(\cdot) = \mathbb{E}_{H_0}(\cdot)$ and $\mathbb{V}\mathrm{ar}_{\mathrm{ER}}(\cdot) = \mathbb{V}\mathrm{ar}_{H_0}(\cdot)$ denote expectation and variance, respectively, under the ER model and let $\mathbb{E}_{\mathrm{SBM}}(\cdot) = \mathbb{E}_{H_1}(\cdot)$ denote the expectation under the SBM. Then, for the centered subgraph count $S_n(C_3)$ and the raw subgraph count $T_n(C_3)$, respectively, for $n \to \infty$, the following holds:

(i) **Fixed alternatives:** *If $p \to p_0 \in [0, 1)$ with $np \to \infty$ and $\epsilon \sim p$, we have*

$$(a) \quad \frac{\mathbb{E}_{\mathrm{SBM}}(S_n(C_3)) - \mathbb{E}_{\mathrm{ER}}(S_n(C_3))}{\sqrt{\mathbb{V}\mathrm{ar}_{\mathrm{ER}}(S_n(C_3))}} = \begin{cases} O(n^{3/2}), & \text{if } p \to p_0 \in (0,1), \\ O((np)3/^2), & \text{if } p \to 0, \end{cases}$$

and

$$(b) \quad \frac{\mathbb{E}_{\mathrm{SBM}}(T_n(C_3)) - \mathbb{E}_{\mathrm{ER}}(T_n(C_3))}{\sqrt{\mathbb{V}\mathrm{ar}_{\mathrm{ER}}(T_n(C_3))}} = \begin{cases} O(n), & \text{if } p \to p_0 \in (0,1), \\ O(n\sqrt{p}), & \text{if } p \to 0, np^2 \to \infty, \\ O((np)3/^2), & \text{if } p \to 0, np^2 \to K \geq 0. \end{cases}$$

Hence, if $np \to \infty$, we get divergence to $+\infty$ in all cases.

(ii) **Local alternatives:** *If $p \to p_0 \in [0,1)$ with $np \to \infty$ and $\epsilon = o(p)$, we have*

$$(a) \quad \frac{\mathbb{E}_{\mathrm{SBM}}(S_n(C_3)) - \mathbb{E}_{\mathrm{ER}}(S_n(C_3))}{\sqrt{\mathbb{V}\mathrm{ar}_{\mathrm{ER}}(S_n(C_3))}} = \begin{cases} O(n^{3/2}\epsilon^3), & \text{if } p \to p_0 \in (0,1), \\ O(n^{3/2}\epsilon^3 p^{-3/2}), & \text{if } p \to 0, \end{cases}$$

and

$$(b) \quad \frac{\mathbb{E}_{\mathrm{SBM}}(T_n(C_3)) - \mathbb{E}_{\mathrm{ER}}(T_n(C_3))}{\sqrt{\mathbb{V}\mathrm{ar}_{\mathrm{ER}}(T_n(C_3))}} = \begin{cases} O(n\epsilon^3 p^{-5/2}), & \text{if } n\epsilon \to \infty, np^2 \to \infty, \\ O(\epsilon p^{-5/2}), & \text{if } n\epsilon \to K_\epsilon \geq 0, np^2 \to \infty, \\ O(n^{3/2}\epsilon^3 p^{-3/2}), & \text{if } n\epsilon \to \infty, np^2 \to K \geq 0, \\ O(n^{1/2}\epsilon p^{-3/2}), & \text{if } n\epsilon \to K_\epsilon \geq 0, np^2 \to K \geq 0. \end{cases}$$

Whenever, in one of the above cases, $\epsilon \to 0$ such that the corresponding expression in $O(\cdot)$ notation is bounded, local power is obtained. Whenever $\epsilon \to 0$ at any faster rate, we get $o(1)$.

In the context of our work, we are more interested in statements (a) about the behavior of the centered subgraph count $S_n(C_3)$. However, for both $S_n(C_3)$ and $T_n(C_3)$, the theorem gives conditions such that C_3 is sensitive to distinguishing between the investigated models. Similar to these derivations, we can derive a corresponding result for P_3.

Theorem 3 (Power of $S_n(P_3)$ and $T_n(P_3)$ under fixed and local alternatives). *With the same setup as for Theorem 2, the following holds for $n \to \infty$:*

(i) **Fixed alternatives:** *If $p \to p_0 \in [0,1)$ with $np^{2/3} \to \infty$ and $\epsilon \sim p$, we have*

$$(a) \quad \frac{\mathbb{E}_{\mathrm{SBM}}(S_n(P_3)) - \mathbb{E}_{\mathrm{ER}}(S_n(P_3))}{\sqrt{\mathbb{V}\mathrm{ar}_{\mathrm{ER}}(S_n(P_3))}} = \begin{cases} O(n^{1/2}), & \text{if } p \to p_0 \in (0,1) \\ O(n^{1/2}p), & \text{if } p \to 0 \end{cases},$$

and

$$(b) \quad \frac{\mathbb{E}_{\mathrm{SBM}}(T_n(P_3)) - \mathbb{E}_{\mathrm{ER}}(T_n(P_3))}{\sqrt{\mathbb{V}\mathrm{ar}_{\mathrm{ER}}(T_n(P_3))}} = \begin{cases} O(1), & \text{if } p \to p_0 \in (0,1) \\ O(\sqrt{p}), & \text{if} p \to 0, np \to \infty \\ O(\sqrt{np}), & \text{if } p \to 0, np \to K \geq 0 \end{cases}.$$

Hence, if $np^2 \to \infty$, we get divergence to $+\infty$ for $S_n(P_3)$, but not for $T_n(P_3)$.

(ii) **Local alternatives:** *If $p \to p_0 \in [0,1)$ with $np^{2/3} \to \infty$ and $\epsilon = o(p)$, we have*

$$(a) \quad \frac{\mathbb{E}_{\mathrm{SBM}}(S_n(P_3)) - \mathbb{E}_{\mathrm{ER}}(S_n(P_3))}{\sqrt{\mathbb{V}\mathrm{ar}_{\mathrm{ER}}(S_n(P_3))}} = \begin{cases} O(n^{1/2}\epsilon^2), & \text{if } p \to p_0 \in (0,1) \\ O(n^{1/2}\epsilon^2 p^{-1}), & \text{if } p \to 0 \end{cases},$$

and

$$(b)\quad \frac{\mathbb{E}_{\mathrm{SBM}}(T_n(P_3)) - \mathbb{E}_{\mathrm{ER}}(T_n(P_3))}{\sqrt{\mathbb{V}\mathrm{ar}_{\mathrm{ER}}(T_n(P_3))}} = \begin{cases} O(\epsilon), & \text{if } p \to p_0 \in (0,1) \\ O(\epsilon p^{-1/2}), & \text{if } p \to 0,\ np \to \infty \\ O(n^{1/2}\epsilon), & \text{if } p \to 0,\ np \to K \geq 0 \end{cases}.$$

Hence, if $n\epsilon^4 = O(1)$ for the case $p \to p_0 \in (0,1)$ or if $n\epsilon^4 p^{-2} = O(1)$ for $p \to 0$, local power is obtained for $S_n(P_3)$. Whenever $\epsilon \to 0$ at any faster rate, we get $o(1)$. For $T_n(P_3)$, we do not have power.

Considering the rates, the triangle structure is more sensitive to detecting these types of SBMs. Intuitively, this makes sense: The conditions for triangles are more restrictive than for two-stars, since all links between the three involved nodes have to exist. Obviously, this is easier to achieve in the dense blocks of an SBM than in an ER graph due to $p_{\mathrm{intra}} > p_{\mathrm{mean}}$. These results are confirmed by the simulation study in Section 5, where we also evaluate the behavior for more flexible SBM setups with different parameters.

4 | BOOTSTRAP THEORY FOR GRAPH FUNCTIONALS

While knowledge of the presented limiting distributions for the class of graph functionals generally allows the construction and implementation of a powerful testing procedure, such asymptotic tests might have issues when it comes to small sample sizes, i.e. for small network sizes n. Additionally, the derivation of the centered subgraph representation following Proposition 2 is not trivial and can be tedious. In order to provide a more feasible solution and to potentially profit from finite sample improvements, we propose to use parametric bootstrapping to approximate the distribution of the test under the null.

4.1 | Bootstrap scheme

Having observed a simple random graph G, the bootstrap algorithm to estimate the distribution of a graph functional X_n under the null of an ER-graph is defined as follows:

Step 1. Estimate the connection probability p by $\hat{p} = \binom{n}{2}^{-1} \sum_{1\leq i<j\leq n} A_{ij}$.
Step 2. Conditional on $\boldsymbol{A} = (A_{ij})_{1\leq i,j\leq n}$, generate an ER graph G^* by drawing a symmetric adjacency matrix $\boldsymbol{A}^* = (A^*_{ij})_{1\leq i,j\leq n}$ from $\mathcal{G}(n,\hat{p})$. That is, conditional on $\boldsymbol{A} = (A_{ij})_{1\leq i,j\leq n}$, we have $A^*_{ij} \sim \mathrm{Bin}(1,\hat{p})$ for $1 \leq i < j \leq n$ with $A^*_{ji} := A^*_{ij}$, and $A^*_{ii} := 0$ for all i.
Step 3. Calculate the bootstrap graph functional $X^*_n = X_n(G^*)$.
Step 4. Repeat Steps 2 and 3 B times, where B is large, to obtain $X_n^{*(b)}$, $b = 1,\ldots,B$.
Step 5. Approximate the distribution of the centered graph functional $X_n - \mathbb{E}(X_n)$ by the empirical distribution of the centered bootstrap graph functionals $X_n^{*(1)} - \mathbb{E}^*(X^*_n),\ldots,X_n^{*(B)} - \mathbb{E}^*(X^*_n)$ (Hall), where $\mathbb{E}^*(\cdot)$ denotes the bootstrap expectation conditional on the original network or, alternatively, the standardized distribution of $\frac{X_n-\mathbb{E}(X_n)}{\sqrt{\mathbb{V}\mathrm{ar}(X_n)}}$ by the empirical distribution of the bootstrap graph functionals $\frac{X_n^{*(1)}-\mathbb{E}^*(X^*_n)}{\sqrt{\mathbb{V}\mathrm{ar}^*(X^*_n)}}, \ldots, \frac{X_n^{*(B)}-\mathbb{E}^*(X^*_n)}{\sqrt{\mathbb{V}\mathrm{ar}^*(X^*_n)}}$ (studentized Hall), where $\mathbb{V}\mathrm{ar}^*(\ldots)$ denotes the bootstrap variance conditional on the original network.

4.2 | Bootstrap theory

In the following theorem, we provide asymptotic theory for the bootstrap procedure and prove its consistency in the framework of Theorem 1 under the null and the alternative.

Theorem 4. *Suppose $G \in \mathcal{G}_{\text{HER}}$ with mean connectivity $p_{\text{mean}} = \binom{n}{2}^{-1} \sum_{1\le i<j\le n} p_{ij}$ such that $p_{\text{mean}} = p_{\text{mean}}(n) \xrightarrow[n\to\infty]{} p_0 \in [0,1)$ with $\sup_{i,j} p_{ij} = o(1)$ if $p_0 = 0$ and $\sup_{n;i,j} p_{ij} < 1$ and $\inf_{n;i,j} p_{ij} > 0$ if $p_0 \in (0,1)$. Further, let X_n be a graph functional $X_n = X_n(G)$ computed from G. Suppose that, under the null hypothesis of a homogeneous ER graph, i.e. when $G \in \mathcal{G}_{ER} \subsetneq \mathcal{G}_{\text{HER}}$, the graph functional X_n has an expansion $X_n = \sum_{H\in\mathcal{H}} a_n(H)S_n(H)$ with $a_n(H) = a_n(H,p)$ that is dominated by a family of connected graphs $\mathcal{H}_0$. For all $H \in \mathcal{H}_0$, we suppose that $np_{\text{mean}}^{r(H)} \xrightarrow[n\to\infty]{} \infty$ with $r(H) = \max_{J\subseteq H} d(J) \ge 1$, where $d(J) = m(J)/n(J)$, the coefficients*

$$b_{mean}(H) = \sup_n \frac{n^{n(H)/2} p_{\text{mean}}^{m(H)/2} a_n(H)}{\sqrt{\mathbb{V}\text{ar}(X_n)}} \tag{16}$$

are finite and satisfy

$$\sum_{H\in\mathcal{H}_0} b_{mean}^2(H)|\text{aut}(H)| < \infty. \tag{17}$$

Further, let $X_n^* = X_n(G^*)$ denote the bootstrap version of the graph functional X_n of G^* that is generated according to the bootstrap scheme in Section 4.1. Then, we have $X_n^* = \sum_{H\in\mathcal{H}} a_n^*(H)S_n^*(H)$ with $a_n^*(H) = a_n(H,\hat{p})$ and

$$S_n^*(H) = \sum_{\varphi} \prod_{e\in E(\varphi(H))} (\mathbb{1}\{e \in E(G^*)\} - \mathbb{P}^*(e \in E(G^*))) = \sum_{\varphi} \prod_{e\in E(\varphi(H))} \left(\mathbb{1}\{e \in E(G^*)\} - \hat{p}\right),$$

and, as $n \to \infty$, it holds

$$R(X_n^*) := \frac{X_n^* - \mathbb{E}^*(X_n^*)}{\sqrt{\mathbb{V}\text{ar}^*(X_n^*)}} \xrightarrow{d} \mathcal{N}(0,1) \quad \text{in probability.}$$

Additionally, with $(c_n)_{n\in\mathbb{N}}$ and V^2 as defined in Theorem 1, we have $c_n^2 \mathbb{V}\text{ar}^*(X_n^*) \to V^2 > 0$ in probability.

Note that the asymptotic statement in Theorem 4 is "in distribution in probability" as it is usually used in the derivation of asymptotic bootstrap theory. We refer, for example, to classical literature such as Shao and Tu (2012). As a direct consequence, we get bootstrap consistency for the bootstrap procedure from Section 4.1 in the following sense.

Corollary 1 (Bootstrap consistency). *Under the assumptions of Theorems 1 and 4, we have*

$$\sup_{x\in\mathbb{R}} \left| \mathbb{P}_{H_0}\left(\frac{X_n - \mathbb{E}_{H_0}(X_n)}{\sqrt{\mathbb{V}\text{ar}_{H_0}(X_n)}} \le x \right) - \mathbb{P}^*\left(\frac{X_n^* - \mathbb{E}^*(X_n^*)}{\sqrt{\mathbb{V}\text{ar}^*(X_n^*)}} \le x \right) \right| \to 0 \quad \text{in probability,} \tag{18}$$

where $\mathbb{P}_{H_0}$ denotes the probability under the null hypothesis of a homogeneous ER graph and $\mathbb{P}^*$ denotes the bootstrap probability measure induced by the bootstrap scheme from Section 4 (conditional on G). Additionally, we have

$$c_n^2\left(\mathbb{V}\mathrm{ar}_{H_0}(X_n) - \mathbb{V}\mathrm{ar}^*(X_n^*)\right) \to 0 \quad \textit{in probability,} \tag{19}$$

leading to

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}_{H_0}\left(c_n(X_n - \mathbb{E}_{H_0}(X_n)) \le x\right) - \mathbb{P}^*(c_n(X_n^* - \mathbb{E}^*(X_n^*)) \le x)\right| \to 0 \text{ in probability.} \tag{20}$$

The bootstrap consistency results for the parametric bootstrap procedure in (18) and (20) enable the construction of goodness-of-fit tests based on graph functionals not only in an asymptotic, but also in a bootstrap manner. This has the great advantage that we do not rely anymore on finding the explicit centered subgraph representation of a graph functional as in Proposition 2 leading to a handily applicable class of goodness-of-fit procedures. In particular, if conditions (A) and (B) in (9) hold, an asymptotically valid bootstrap level-α test for testing the null hypothesis H_0 of an underlying homogeneous Erdös–Rényi graph against heterogeneous alternatives H_1 in (1) is given by the decision rule

$$\phi_\alpha^*(X_n) = \begin{cases} 1, & \text{if } \hat{R}(X_n) < z_{\alpha/2}^* \quad \text{or} \quad \hat{R}(X_n) > z_{1-\alpha/2}^*, \\ 0, & \text{else,} \end{cases}$$

where $z_{\alpha/2}^*$ and $z_{1-\alpha/2}^*$ are the $\alpha/2$ and the $(1-\alpha/2)$ quantiles of the empirical distribution of the resampled statistic $\hat{R}(X_n^*) = (X_n^* - \hat{\mathbb{E}}^*(X_n^*))/\sqrt{\widehat{\mathbb{V}\mathrm{ar}}^*(X_n^*)}$. Due to (19), an alternative asymptotically valid bootstrap level-α test is given by

$$\tilde{\phi}_\alpha^*(X_n) = \begin{cases} 1, & \text{if } X_n - \widehat{\mathbb{E}_{H_0}}(X_n) < \tilde{z}_{\alpha/2}^* \quad \text{or} \quad X_n - \widehat{\mathbb{E}_{H_0}}(X_n) > \tilde{z}_{1-\alpha/2}^* \\ 0, & \text{else,} \end{cases}$$

where $\tilde{z}_{\alpha/2}^*$ and $\tilde{z}_{1-\alpha/2}^*$ are the $\alpha/2$ and the $(1-\alpha/2)$ quantiles of the empirical distribution of the resampled statistic $X_n^* - \hat{\mathbb{E}}^*(X_n^*)$. Precisely, this general form now enables us to formulate suitable bootstrap versions of the asymptotic tests presented in Section 3.

Example 2 (Degree variance). A bootstrap version of the asymptotic level-α test using the degree variance V_n is given by

$$\phi_\alpha^*(V_n) = \begin{cases} 1, & \text{if } \frac{V_n - \widehat{\mathbb{E}_{H_0}}(V_n)}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(V_n)}} < z_{\alpha/2}^* \quad \text{or} \quad \frac{V_n - \widehat{\mathbb{E}_{H_0}}(V_n)}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(V_n)}} > z_{1-\alpha/2}^*, \\ 0, & \text{else,} \end{cases}$$

where $z_{\alpha/2}^*$ and $z_{1-\alpha/2}^*$ are the $\alpha/2$ and the $(1-\alpha/2)$ quantiles of the empirical distribution of the resampled statistic $\frac{V_n^* - \hat{\mathbb{E}}^*(V_n^*)}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}}^*(V_n^*)}}$. An alternative bootstrap test of asymptotic level-α is given by

$$\tilde{\phi}^*_\alpha(V_n) = \begin{cases} 1, & \text{if } V_n - \widehat{\mathbb{E}_{H_0}}(V_n) < z^*_{\alpha/2} \quad \text{or} \quad V_n - \widehat{\mathbb{E}_{H_0}}(V_n) > z^*_{1-\alpha/2} \\ 0, & \text{else,} \end{cases}$$

where $z^*_{\alpha/2}$ and $z^*_{1-\alpha/2}$ are the $\alpha/2$ and the $(1-\alpha/2)$ quantiles of the empirical distribution of the resampled statistic $V^*_n - \widehat{\mathbb{E}}^*(V^*_n)$.

Example 3 (Eigenvector centrality). As long as the conditions of Theorem 4 as well as (A) and (B) from Proposition 3 are satisfied, we may use any graph functional for constructing a goodness-of-fit test without deriving its centered subgraph count representation according to Proposition 2. This is particularly helpful for those graph functionals for which finding this representation may be tedious. This can be the case for quite complicated network metrics such as centrality metrics. These are popularly used in practice, but to the best of our knowledge there does not exist (yet) a well-established asymptotic theory. This is why the bootstrap procedure can be beneficial as it allows us to construct novel tests with statistics that have not been used before. A bootstrap test of asymptotic level-α for the average eigenvector centrality E_n is given by

$$\tilde{\phi}^*_\alpha(E_n) = \begin{cases} 1, & \text{if } E_n - \widehat{\mathbb{E}_{H_0}}(E_n) < z^*_{\alpha/2} \quad \text{or} \quad E_n - \widehat{\mathbb{E}_{H_0}}(E_n) > z^*_{1-\alpha/2} \\ 0, & \text{else,} \end{cases}$$

where $z^*_{\alpha/2}$ and $z^*_{1-\alpha/2}$ are the $\alpha/2$ and the $(1-\alpha/2)$ quantiles of the empirical distribution of the resampled statistic $E^*_n - \widehat{\mathbb{E}}^*(E^*_n)$. To the best of our knowledge, there does not exist a closed form of $\mathbb{E}_{H_0}(E_n)$ in the literature. Consequently, there is also no closed form of the plug-in version $\widehat{\mathbb{E}_{H_0}}(E_n)$. This is why we use a double bootstrap approach (see e.g., Efron, 1992) in order to estimate $\widehat{\mathbb{E}_{H_0}}(E_n)$.

5 | SIMULATION STUDY

To underline our findings and to gain further evidence for the proposed class of tests, we execute a simulation study for different parameter setups and various alternatives. We analyze the power of the tests and illustrate performance differences depending on the chosen graph functional.

5.1 | General setting

We investigate four different graph functionals for the construction of the test statistics. Namely, these are the degree variance statistic V_n, the average eigenvector centrality E_n, centered triangle counts $S_n(C_3)$ and centered two-star counts $S_n(P_3)$. We analyze their performance for the asymptotic version and for the bootstrap version of the testing procedure. Note that for E_n, we only consider the bootstrap version. To investigate the power of these test statistics, we need to generate networks from the alternative, that is, from the $\mathcal{G}_{\text{HER}}(n)$ model. In order to be able to detect the alternative, the generated networks have to deviate from the homogeneous null model by an increasing amount with increasing heterogeneity. To generate random graphs under the alternative, we set a mean connectivity p_{mean} such that the null model is a $\mathcal{G}(n, p_{\text{mean}})$ model. This is to ensure that the power of the tests is not due to differences in the mean connectivity (thus differing degrees and mean degrees), but actually due to the rising heterogeneity of the connection probabilities. Hence, we impose:

$$\binom{n}{2}^{-1} \sum_{1 \le i < j \le n} p_{ij} = p_{\text{mean}}.$$

This preserves the mean connectivity over all scenarios and settings. Although this limits the possible alternatives, it is not a big constraint. Having an observed network G, we can directly calculate its mean connectivity p_{mean}. In order to study the performances in various network sizes and density conditions, we use different values of $n = \{16, 32, 64, 128, 256, 512\}$ and $p_{\text{mean}} = p_{\text{mean}}(n) = \left\{ \frac{\log n}{n}, \frac{1}{\sqrt{n}}, \frac{\log n}{\sqrt{n}} \right\}$. Note that for these combinations of n and p_{mean}, we have $\frac{\log n}{n} < \frac{1}{\sqrt{n}} < \frac{\log n}{\sqrt{n}}$. For $n = \{256,512\}$, we will only report the performances of the asymptotic tests as the computation of the power curves of the bootstrap procedure is quite expensive. As we will see, the asymptotic results stand representative for the bootstrap results as they coincide for large n. As the $\mathcal{G}_{\text{HER}}(n)$ model class is very broad, we limit the simulation study to a few relevant scenarios for the alternative that serve as representatives for different possible ways the matrix $\boldsymbol{P}$ can be set up. These include the popular use case of SBMs and covariate models. For the SBMs, we use two different setups including an SBM with two blocks of equal block sizes as in our analysis of Section 3.4 and the extension to three blocks with random block sizes and varying intra-group probabilities. The exact setup of each alternative will be explained in the corresponding subsections below. We calculate the power of the tests by the application of 1000 replications for each setting. For the bootstrap version, we use $B = 700$ bootstrap replications. We set the test level as $\alpha = 0.05$ for all tests.

5.2 | Performances

With the consideration of all described parameter variations, we investigate the power of 72 different scenarios. The main results and important insights are presented below.

5.2.1 | Stochastic block models

As mentioned, we construct two different versions of SBMs. The first one is characterized by two blocks of equal block sizes and is constructed as follows. We assign weights to the nodes with

$$\boldsymbol{w} = (w_1, \dots, w_n) = \Big(\underbrace{-\frac{1}{2}, \dots, -\frac{1}{2}}_{n/2 \text{ times}}, \underbrace{\frac{1}{2}, \dots, \frac{1}{2}}_{n/2 \text{ times}}\Big).$$

By premultiplying the weights with a factor $\lambda > 0$, we can achieve different degrees of heterogeneity. Concretely, the weights are transformed into connection probabilities p_{ij} using a logit link. We set

$$p_{ij}(a,\lambda)=\begin{cases}\frac{\exp(a+\lambda^2 w_i w_j)}{1+\exp(a+\lambda^2 w_i w_j)}, & \text{for } i\neq j\\ 0 & \text{else.}\end{cases}$$

The constant a is set to preserve the mean connectivity p_{mean} through the different heterogeneity levels. This is achieved by (numerically) minimizing the function

$$f(a)=\left|p_{\text{mean}}-\binom{n}{2}^{-1}\sum_{1\le i<j\le n}p_{ij}(a,\sigma^2)\right|.$$

In the end, the resulting probabilities have two values, one if $w_i = w_j$, and another one if $w_i \neq w_j$. Thus, the resulting models are two-community SBMs with two equally sized communities. We set $\lambda \in \{0, 0.5, 1, 1.5, \dots, 4\}$. Note that $\lambda = 0$ represents the null model where all edge probabilities are $p_{ij} = p_{\text{mean}}$, $i < j = 1, \dots, n$. Increasing values of λ yield models with rising heterogeneity: The intra-community connection probabilities increase, whereas the probabilities for inter-community edges decrease.

Performance results for $p_{\text{mean}} = \frac{1}{\sqrt{n}}$ are illustrated in Figure 3. Note that these results stand representative for all other values of p_{mean}, since the power of the tests is not directly influenced by the density of the underlying network. Overall, the results are in accordance with our theoretical analysis of Section 3.4. It is clearly visible that the tests based on C_3 have a superior performance in the underlying SBM cases as they are able to react way more sensitively than both other tests that are mainly based on P_3. While C_3 is able to detect differences even for low values of the heterogeneity parameter λ, the two other tests require a rather large deviation from the null model. In this context, the performances get more reliable for all tests with an increasing amount of nodes. For the case of $n = 16$, P_3 and V_n are even not at all sensitive for all investigated values of λ. On a further note, the similar results of these two tests in all situations underline our derivations that P_3 is the decisive component of the centered subgraph count representation of V_n. The performance of the eigenvector centrality test is similar to these of $S_n(P_3)$ and V_n.

For our second SBM setup, we construct an SBM with more flexible assumptions to its parameters. Compared to before, it consists of three blocks with random block sizes and varying intra-block probabilities. Formally, we set this up by assigning the weight vector $\boldsymbol{w}$ to the nodes with

$$\boldsymbol{w}'=(w_1',\dots,w_n')=\Big(\underbrace{-\frac{1}{2},\dots,-\frac{1}{2}}_{b_1\ \text{times}},\underbrace{0,\dots,0}_{b_2\ \text{times}},\underbrace{\frac{1}{4},\dots,\frac{1}{4}}_{b_3\ \text{times}}\Big).$$

In this context, b_k with $k = \{1, 2, 3\}$ denotes the block size of the corresponding block. Each b_k is randomly chosen with the constraint that $b_1 + b_2 + b_3 = n$ and $b_k \geq 2$ for each k. The probabilities p_{ij} are then determined analogously to the two-block SBM case explained above.

The resulting performances are depicted in Figure 4 for $p_{\text{mean}} = \frac{\log n}{n}$. Regarding the performance under the null, all tests stick to the desired level $\alpha = 0.05$. Apart from this, the results are, interestingly, quite different to before as both tests based on P_3 perform superior compared to the

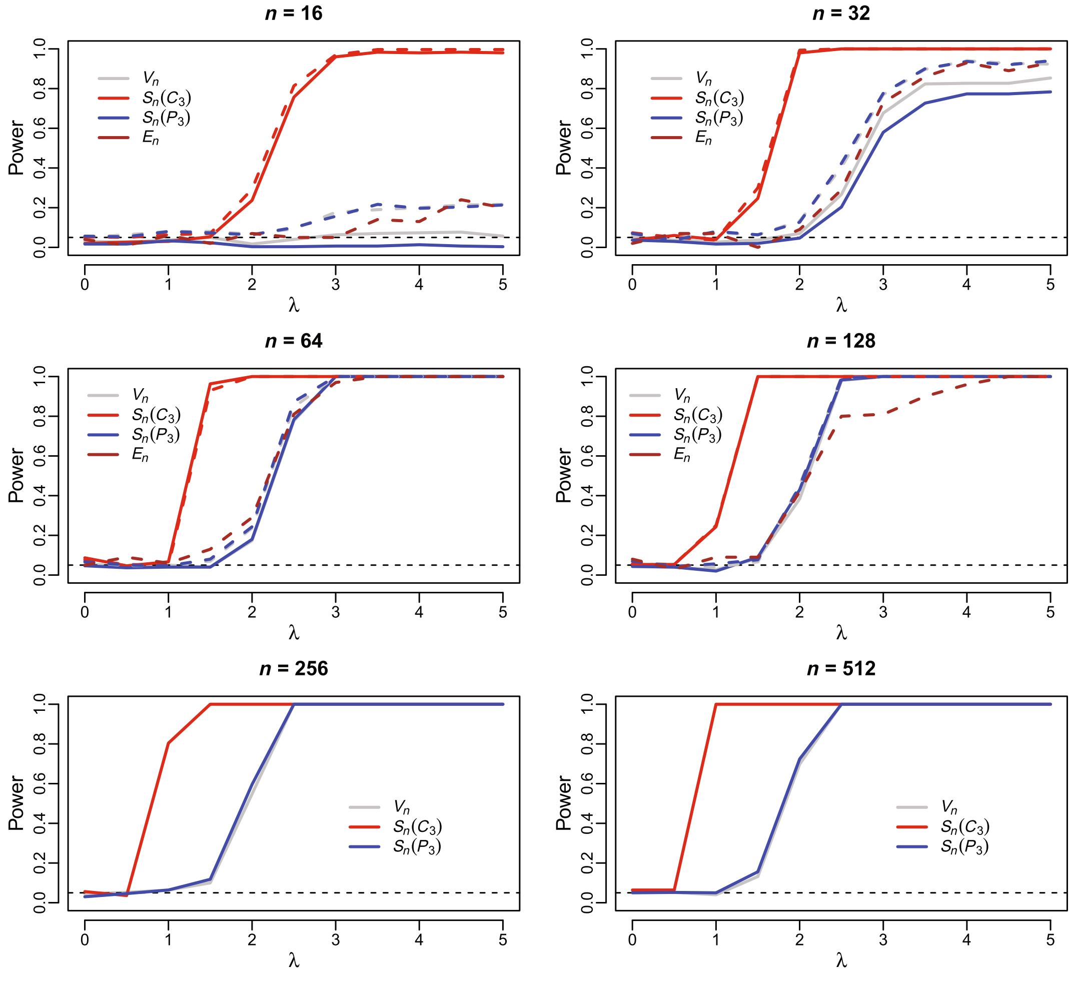

FIGURE 3 Power of the tests V_n, E_n, $S_n(C_3)$, and $S_n(P_3)$ for the alternative of an stochastic block model (SBM) with two blocks and $p_{\text{mean}} = \frac{1}{\sqrt{n}}$ for various values for n and λ. The asymptotic versions of the tests are represented by solid lines and the bootstrap version by dashed lines.

test C_3 and also to E_n. Apparently, the two-star structure is more sensitive to an increased number of blocks, varying block sizes and varying intra-block probabilities. The bootstrap versions of the tests confirm this behaviour by achieving very similar results to their asymptotic counterparts. The improved performance of P_3 and V_n seems quite plausible. As shown in Section 3.4, subgraph counts of P_3 are, in theory, sensitive to detecting SBMs with the assumptions of Theorem 3. Although they are less sensitive compared to C_3 for the two block case, they could benefit from splitting the graph into more blocks as this can be interpreted as a further increase in heterogeneity compared to the ER model. Furthermore, the varying block sizes and intra-block probabilities might influence this behaviour as well. Interestingly, the test based on C_3 performs superior again for denser setups which can be seen in Figure B1. In these denser networks, an increased λ enables the possibility of larger deviations between the intra-group probabilities and inter-group probabilities which could support the sensitivity of C_3.

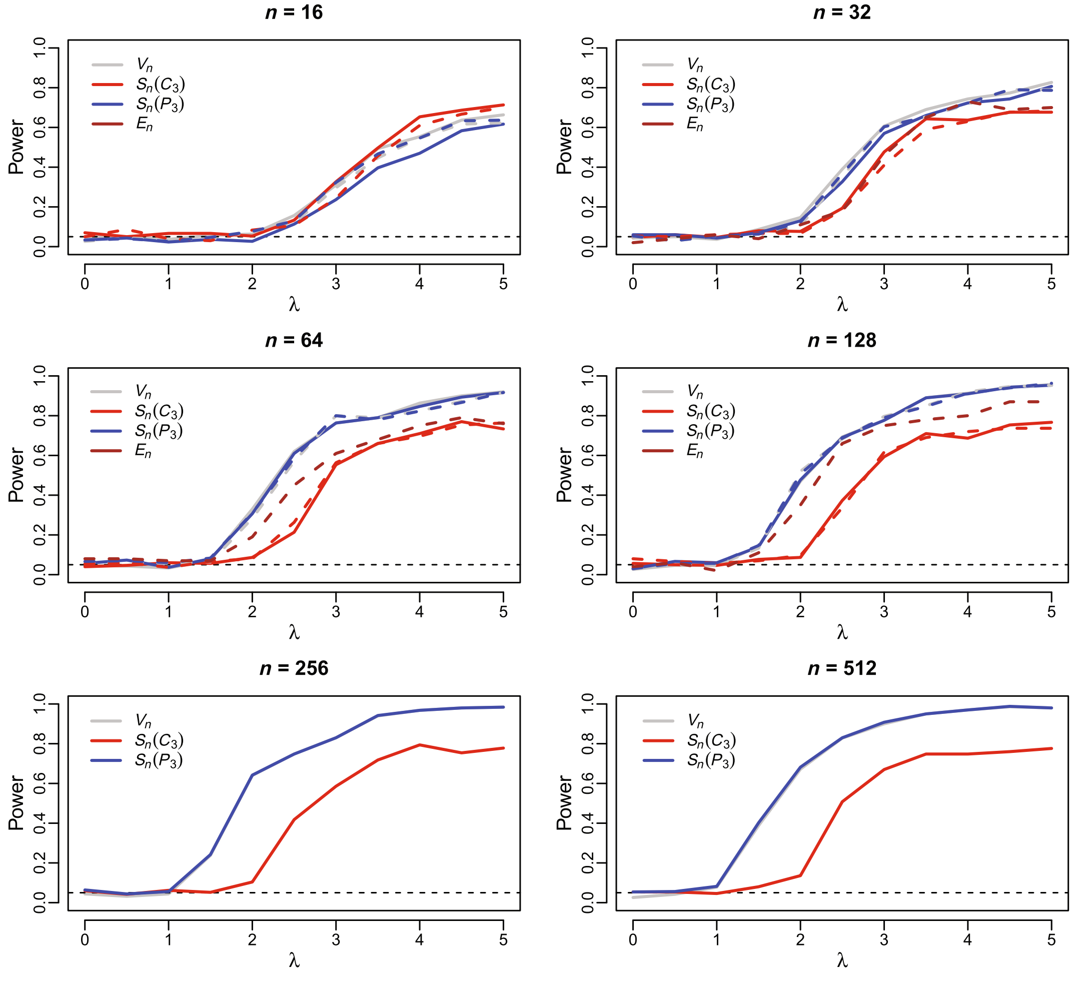

FIGURE 4 Power of the tests V_n, $S_n(C_3)$, E_n, and $S_n(P_3)$ for the alternative of an stochastic block model (SBM) with three blocks and $p_{\text{mean}} = \frac{\log n}{n}$. The asymptotic versions of the tests are represented by solid lines and the bootstrap version by dashed lines.

5.2.2 | Covariate models

As a further heterogeneous alternative, we study covariate models for which the connection probabilities are generated in a way such that vertices with similar properties are more likely to connect than others. This characteristic seems to be a realistic setting for modeling real world networks. Especially in social networks, we expect variables such as the age or the social status to affect whether people know each other or not. To achieve a setting like this, we associate each vertex $v_i \in \{v_1, \dots, v_n\}$ with a bivariate covariate

$$\begin{pmatrix} x_{i1} \\ x_{i2} \end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right),$$

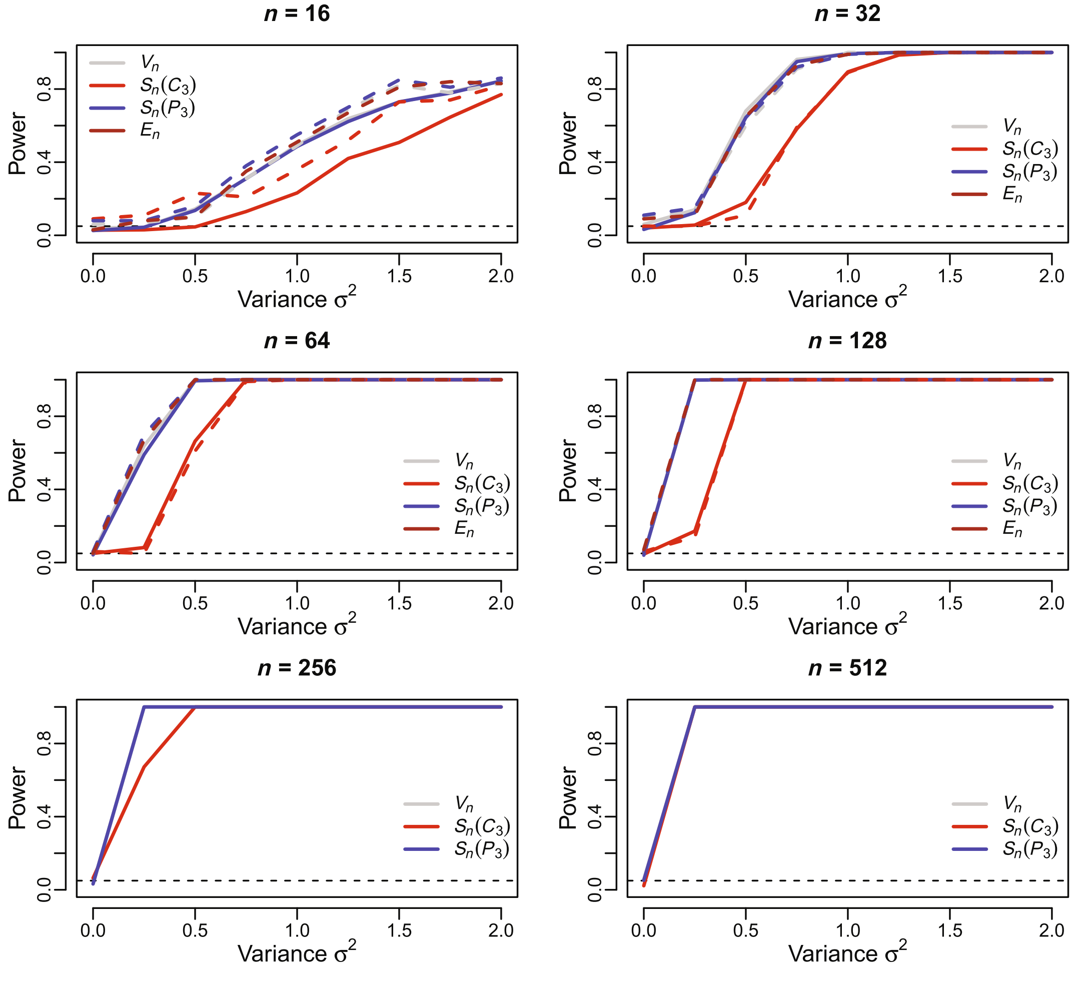

FIGURE 5 Power of the tests V_n, E_n, $S_n(C_3)$ and $S_n(P_3)$ for the alternative of a covariate model setup with $p_{\text{mean}} = \frac{1}{\sqrt{n}}$. The asymptotic versions of the tests are represented by solid lines and the bootstrap version by dashed lines.

drawn independently from a bivariate standard normal distribution. The covariates are then multiplied by a sequence of variances $\sigma^2 \in \{0, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2\}$. This yields versions of the same dataset with different degrees of heterogeneity. The connection probabilities (for given σ^2 and covariate set) are then calculated according to a logistic model:

$$p_{ij}(a, \sigma^2) = \begin{cases} \frac{\exp(a - \sigma^2 |x_{i1} - x_{j1}| - \sigma^2 |x_{i2} - x_{j2}|)}{1 + \exp(a - \sigma^2 |x_{i1} - x_{j1}| - \sigma^2 |x_{i2} - x_{j2}|)}, & \text{for } i \neq j, \\ 0, & \text{else.} \end{cases}$$

The constant a is again set to preserve the mean connectivity p_{mean} and can be determined as before. In this model, $\sigma^2 = 0$ represents the null hypothesis, where each vertex has the same covariate value and the resulting connection probabilities are all equal to p_{mean}. The bigger the variance σ^2, the more the model deviates from the null model due to the increasing heterogeneity of the covariates. In general, the vertices have a higher probability to connect if their covariates have similar values. Especially, vertices that have unusually large or small covariate values will only be able to draw few connections. The described procedure yields a series of connection probability matrices with increasing heterogeneity.

Results are illustrated in Figure 5 for $p_{\text{mean}} = \frac{1}{\sqrt{n}}$. First, the tests reliably stick to the desired level $\alpha = 0.05$. Regarding the performance under the alternative, the performance is overall quite good for all tests with advantages for E_n, V_n, and P_3 compared to C_3. For larger network sizes, all tests have no problems with detecting the deviation from the null model—even for the smallest investigated value of the heterogeneity parameter σ^2. Another result is that the bootstrap versions are able to outperform the asymptotic versions for small n and perform equally well for larger n. This behavior is also confirmed by the results of most of the other investigated parameter setups that are not reported in the paper. As expected, the derived parametric bootstrap procedure is a promising alternative to the asymptotic version of the proposed class of tests—particularly for small network sizes.

6 | CONCLUSION

Network data is quite complex and its handling is challenging from a statistical point of view. An important aspect is the complexity-reduction achieved by using a suitable network model. For a reliable analysis, however, the fitted model should be an adequate representation of the underlying data. In this context, we derived a class of goodness-of-fit tests for networks that serves as a unified approach and contains various formerly proposed tests as special cases. To do so, we used the broad class of graph functionals as test statistics and leveraged existing theory in order to derive novel asymptotic tests for evaluating if an underlying graph is generated by a homogeneous Erdös–Rényi model or some heterogeneous alternative. Moreover, we proposed a parametric bootstrap approach that particularly performs favorably in small network size situations compared to the asymptotic tests. We enriched our analysis by the application of our general procedure to three types of test statistics and derived power analysis results for the subgraphs of triangles and two-stars for the popular use case of certain SBMs. We illustrate our findings with an extensive simulation study in which we investigated multiple network parameter setups and also studied covariate models as a further heterogeneous alternative.

A possible topic for future research is to construct tests for which a heterogeneous model $\mathcal{G}_{\text{HER}}(n)$ represents the null model in the spirit of Ouadah et al. (2020), who exclusively considered the degree variance and without any resampling-based inference. A possible approach could be to try to extend the proof technique for the $\mathcal{G}_{\text{ER}}(n)$ case (Janson et al., 2011), that is based on a continuous time martingale theorem, to the heterogeneous case yielding a heterogeneous equivalent to the method of higher projections.

ACKNOWLEDGMENTS

This work was financially supported by the Mercator Research Center Ruhr (MERCUR) with project number PR-2019-0019 and by the European Union with the project GADMO. The authors are grateful to two reviewers and an associate editor for their helpful and constructive comments. Open Access funding enabled and organized by Projekt DEAL.

ORCID

Barbara Brune https://orcid.org/0000-0002-3154-8445
Jonathan Flossdorf https://orcid.org/0000-0003-2153-0281
Carsten Jentsch https://orcid.org/0000-0001-7824-1697

REFERENCES

Bassett, D. S., & Sporns, O. (2017). Network neuroscience. *Nature Neuroscience*, *20*, 353–364.

Bhattacharyya, S., & Bickel, P. J. (2015). Subsampling bootstrap of count features of networks. *The Annals of Statistics*, *43*, 2384–2411.

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

Bubeck, S., Ding, J., Eldan, R., & Rácz, M. Z. (2016). Testing for high-dimensional geometry in random graphs. *Random Structures & Algorithms*, *49*, 503–532.

Carrington, P. J., Scott, J., & Wasserman, S. (2005). *Models and methods in social network analysis* (Vol. *28*). Cambridge university press.

Dan, S., & Bhattacharya, B. B. (2020). *Goodness-of-fit tests for inhomogeneous random graphs*. In *International conference on machine learning* (pp. 2335–2344). PMLR.

Efron, B. (1992). *Bootstrap methods: Another look at the jackknife*. In *Breakthroughs in statistics: Methodology and distribution* (pp. 569–593). Springer.

Flossdorf, J., Fried, R., & Jentsch, C. (2023). Online monitoring of dynamic networks using flexible multivariate control charts. *Social Network Analysis and Mining*, *13*, 1–18.

Flossdorf, J., & Jentsch, C. (2021). Change detection in dynamic networks using network characteristics. *IEEE Transactions on Signal and Information Processing over Networks*, *7*, 451–464.

Gao, C., & Lafferty, J. (2017). Testing network structure using relations between small subgraph probabilities. *ArXiv Preprint arXiv:1704.06742*.

Gao, C., Lu, Y., & Zhou, H. H. (2015). Rate-optimal graphon estimation. *The Annals of Statistics*, *43*, 2624–2652.

Gilbert, E. N. (1959). Random graphs. *The Annals of Mathematical Statistics*, *30*, 1141–1144.

Hoeffding, W. (1992). *A class of statistics with asymptotically normal distribution*. In *Breakthroughs in statistics* (pp. 308–334). Springer.

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

Hu, J., Zhang, J., Qin, H., Yan, T., & Zhu, J. (2021). Using maximum entry-wise deviation to test the goodness of fit for stochastic block models. *Journal of the American Statistical Association*, *116*, 1373–1382.

Janson, S., Luczak, T., & Rucinski, A. (2011). *Random graphs*. John Wiley & Sons.

Jin, J., Ke, Z., & Luo, S. (2018). *Network global testing by counting graphlets*. In *International conference on machine learning* (pp. 2333–2341). PMLR.

Jin, J., Ke, Z. T., & Luo, S. (2021). Optimal adaptivity of signed-polygon statistics for network testing. *The Annals of Statistics*, *49*, 3408–3433.

Lee, A. J. (2019). *U-statistics: Theory and practice*. Routledge.

Lee, D.-H., & Dong, M. (2009). Dynamic network design for reverse logistics operations under uncertainty. *Transportation Research Part E: Logistics and Transportation Review*, *45*, 61–71.

Lei, J. (2016). A goodness-of-fit test for stochastic block models. *The Annals of Statistics*, *44*, 401–424.

Lospinoso, J., & Snijders, T. A. (2019). Goodness of fit for stochastic actor-oriented models. *Methodological Innovations*, *12*, 2059799119884282.

Lovász, L. (2012). *Large networks and graph limits* (Vol. *60*). American Mathematical Soc.

Maugis, P., Priebe, C. E., Olhede, S. C., & Wolfe, P. J. (2017). Statistical inference for network samples using subgraph counts. *ArXiv Preprint arXiv:1701.00505*.

Maugis, P.-A., Olhede, S., Priebe, C., & Wolfe, P. (2020). Testing for equivalence of network distribution using subgraph counts. *Journal of Computational and Graphical Statistics*, *29*, 455–465.

Nowicki, K. (1989). Asymptotic normality of graph statistics. *Journal of Statistical Planning and Inference*, *21*, 209–222.

Nowicki, K., & Wierman, J. C. (1988). Subgraph counts in random graphs using incomplete u-statistics methods. *Discrete Mathematics*, *72*, 299–310.

Ofori-Boateng, D., Gel, Y. R., & Cribben, I. (2021). Nonparametric anomaly detection on time series of graphs. *Journal of Computational and Graphical Statistics*, *30*, 756–767.

Ospina-Forero, L., Deane, C. M., & Reinert, G. (2019). Assessment of model fit via network comparison methods based on subgraph counts. *Journal of Complex Networks*, *7*, 226–253.

Ouadah, S., Robin, S., & Latouche, P. (2020). Degree-based goodness-of-fit tests for heterogeneous random graph models: Independent and exchangeable cases. *Scandinavian Journal of Statistics*, *47*, 156–181.

Prill, R. J., Iglesias, P. A., & Levchenko, A. (2005). Dynamic properties of network motifs contribute to biological network organization. *PLoS Biology*, *3*, e343.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social Networks*, *29*, 173–191.

Sarkar, P., & Moore, A. W. (2005). Dynamic social network analysis using latent space models. *Acm Sigkdd Explorations Newsletter*, *7*, 31–40.

Shao, J., & Tu, D. (2012). *The jackknife and bootstrap*. Springer Science & Business Media.

Snijders, T. A. (1981). The degree variance: An index of graph heterogeneity. *Social Networks*, *3*, 163–174.

Snijders, T. A. (1996). Stochastic actor-oriented models for network change. *Journal of Mathematical Sociology*, *21*, 149–172.

Xu, W., & Reinert, G. (2021). *A stein goodness-of-test for exponential random graph models*. In *International conference on artificial intelligence and statistics* (pp. 415–423). PMLR.

Yuan, M., Liu, R., Feng, Y., & Shang, Z. (2022). Testing community structure for hypergraphs. *The Annals of Statistics*, *50*, 147–169.

How to cite this article: Brune, B., Flossdorf, J., & Jentsch, C. (2025). Goodness-of-fit testing based on graph functionals for homogeneous Erdös–Rényi graphs. *Scandinavian Journal of Statistics*, *52*(1), 332–380. https://doi.org/10.1111/sjos.12750

APPENDIX A. PROOFS

The following appendix contains all proofs of our novel theorems/propositions/statements in the paper.

Proof of Proposition 3. We have

$$\frac{X_n - \widehat{\mathbb{E}_{H_0}}(X_n)}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(X_n)}} = \frac{\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(X_n)}}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(X_n)}} \left(\frac{X_n - \mathbb{E}_{H_0}(X_n)}{\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(X_n)}} + \frac{\mathbb{E}_{H_0}(X_n) - \widehat{\mathbb{E}_{H_0}}(X_n)}{\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(X_n)}} \right).$$

Then, for the first factor on the last right-hand side, we have $\mathbb{V}\mathrm{ar}_{H_0}(X_n)/\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(X_n) \xrightarrow{P} 1$ due to condition (A) in (9). Further, according to condition (*B*) in (9), the second term in parentheses, that is, $(\mathbb{E}_{H_0}(X_n) - \widehat{\mathbb{E}_{H_0}}(X_n))/(\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(X_n)})$ vanishes as $n \to \infty$. Altogether, as the first term in parentheses, i.e. $(X_n - \mathbb{E}_{H_0}(X_n))/\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(X_n)}$ converges in distribution to $\mathcal{N}(0, 1)$ according to Theorem 1, by Slutsky's Lemma, the claimed result follows. ∎

Proof of Proposition 4. To derive the rate of $\hat{p} - p$, from $\mathbb{E}(\hat{p}) = p$ and

$$\mathbb{V}\mathrm{ar}(\hat{p}) = \binom{n}{2}^{-2} \sum_{1 \le i < j \le n} \mathbb{V}\mathrm{ar}(A_{ij}) = \binom{n}{2}^{-2} \sum_{1 \le i < j \le n} p(1-p),$$

under the null of a homogeneous ER graph, we get

$$\hat{p} - p = O_P\left(\frac{\sqrt{p}}{n}\right).$$

Consequently, we also have

$$\begin{aligned}
\hat{p}(1-\hat{p}) - p(1-p) &= \hat{p}(1-\hat{p}) - \hat{p}(1-p) + \hat{p}(1-p) - p(1-p) = \hat{p}(p-\hat{p}) + (\hat{p}-p)(1-p) \\
&= (\hat{p} - p + p)(p - \hat{p}) + (\hat{p}-p)(1-p) = p(p-\hat{p}) - (\hat{p}-p)^2 + (\hat{p}-p)(1-p) \\
&= \mathcal{O}_{\mathbb{P}}\left(\frac{p^{3/2}}{n}\right) + \mathcal{O}_{\mathbb{P}}\left(\frac{p}{n^2}\right) + \mathcal{O}_{\mathbb{P}}\left(\frac{\sqrt{p}}{n}\right) = \mathcal{O}_{\mathbb{P}}\left(\frac{\sqrt{p}}{n}\right).
\end{aligned}$$

We then consider the following expansion:

$$\frac{V_n - \widehat{\mathbb{E}_{H_0}}(V_n)}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(V_n)}} = \frac{\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(V_n)}}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(V_n)}}\left(\frac{V_n - \mathbb{E}_{H_0}(V_n)}{\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(V_n)}} - \frac{\widehat{\mathbb{E}_{H_0}}(V_n) - \mathbb{E}_{H_0}(V_n)}{\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(V_n)}}\right).$$

For the difference of the expectations in the numerator of the second term in brackets, we get

$$\widehat{\mathbb{E}_{H_0}}(V_n) - \mathbb{E}_{H_0}(V_n) = \frac{(n-1)(n-2)}{n}\left(\hat{p}(1-\hat{p}) - p(1-p)\right) = \mathcal{O}(n) \cdot \mathcal{O}_{\mathbb{P}}\left(\frac{\sqrt{p}}{n}\right) = \mathcal{O}_{\mathbb{P}}(\sqrt{p}).$$

As the variance $\mathbb{V}\mathrm{ar}_{H_0}(V_n)$ has order np^2, and $np \to \infty$ by assumption, we obtain

$$\frac{\widehat{\mathbb{E}_{H_0}}(V_n) - \mathbb{E}_{H_0}(V_n)}{\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(V_n)}} = \mathcal{O}_{\mathbb{P}}\left(\frac{\sqrt{p}}{\sqrt{np^2}}\right) = \mathcal{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{np}}\right) \xrightarrow{P} 0, \text{ as well as } \frac{\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(V_n)}}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(V_n)}} \xrightarrow{P} 1,$$

by similar arguments using $\hat{p} - p = O_P(\frac{\sqrt{p}}{n})$. Combining these two convergence results and applying Slutzky's lemma, we get that $\frac{V_n - \widehat{\mathbb{E}_{H_0}}(V_n)}{\sqrt{\widehat{\mathbb{V}\mathrm{ar}_{H_0}}(V_n)}}$ has the same asymptotic distribution as $\frac{V_n - \mathbb{E}_{H_0}(V_n)}{\sqrt{\mathbb{V}\mathrm{ar}_{H_0}(V_n)}}$ stated in (13). ■

Proof of Proposition 5. We have

$$\begin{aligned}
&\frac{\hat{S}_n(H)}{\sqrt{\mathrm{aut}(H)(n)_{n(H)}(\hat{p}(1-\hat{p}))^{m(H)}}} \\
&\quad = \sqrt{\frac{\mathrm{aut}(H)(n)_{n(H)}(p(1-p))^{m(H)}}{\mathrm{aut}(H)(n)_{n(H)}(\hat{p}(1-\hat{p}))^{m(H)}}}\left(\frac{S_n(H)}{\sqrt{\mathrm{aut}(H)(n)_{n(H)}(p(1-p))^{m(H)}}}\right.
\end{aligned}$$

$$+ \frac{\widehat{S}_n(H) - S_n(H)}{\sqrt{\mathrm{aut}(H)(n)_{n(H)}(p(1-p))^{m(H)}}}\Bigg)$$
$$= \sqrt{\frac{(p(1-p))^{m(H)}}{(\widehat{p}(1-\widehat{p}))^{m(H)}}}\left(\frac{S_n(H)}{\sqrt{\mathrm{aut}(H)(n)_{n(H)}(p(1-p))^{m(H)}}} + \frac{\widehat{S}_n(H) - S_n(H)}{\sqrt{\mathrm{aut}(H)(n)_{n(H)}(p(1-p))^{m(H)}}}\right).$$

Hence, to prove the claimed result, by using the continous mapping theorem and by Slutsky's Lemma, it remains to show

$$\text{(a)} \quad \frac{\widehat{S}_n(H) - S_n(H)}{\sqrt{\mathrm{aut}(H)(n)_{n(H)}(p(1-p))^{m(H)}}} \xrightarrow{P} 0, \quad \text{and} \quad \text{(b)} \quad \frac{(p(1-p))^{m(H)}}{(\widehat{p}(1-\widehat{p}))^{m(H)}} \xrightarrow{P} 1.$$

For showing part (b), it is equivalent to show that the reciprocal, that is,

$$\frac{(\widehat{p}(1-\widehat{p}))^{m(H)}}{(p(1-p))^{m(H)}} = \left(\frac{\widehat{p}}{p}\right)^{m(H)}\left(\frac{1-\widehat{p}}{1-p}\right)^{m(H)},$$

converges to 1. As $\widehat{p} - p = O_P(\frac{\sqrt{p}}{n})$, we have $\widehat{p}/p - 1 = O_P(\frac{1}{\sqrt{p}n}) = o_P(1)$ as $n \to \infty$ due to $np^{2/3} \to \infty$ by assumption. Consequently, by continuous mapping, we have also $(p/\widehat{p})^{m(H)} - 1^{m(H)} = (p/\widehat{p})^{m(H)} - 1 = o_P(1)$. Similarly, we have

$$\frac{1-\widehat{p}}{1-p} = \frac{1-p}{1-p} + \frac{p-\widehat{p}}{1-p} = 1 + \frac{p-\widehat{p}}{1-p} = 1 + O_P\left(\frac{\sqrt{p}}{n}\right) = 1 + o_P(1),$$

because $p = p(n) \to p_0 \in [0, 1)$. Hence, by continuous mapping, we have $((1-p)/(1-\widehat{p}))^{m(H)} - 1^{m(H)} = ((1-p)/(1-\widehat{p}))^{m(H)} - 1 = o_P(1)$ as well. Altogether, we proved part (b). For part (a), under the null of an ER graph, it is natural to decompose $\widehat{S}_n(H)$ and to re-write it in terms of centered subgraph counts. For this purpose, let G be a graph on n vertices, and H be another graph with $n(H) \leq n$. Then, similar to Ex. 6.50 in Janson et al. (2011), let $J \subseteq H$ be a graph that is obtained from H by dropping some of the edges present in H. Hence, we have $n(J) = n(H)$, but $m(J) \leq m(H)$. We consider the $(n)_{n(J)}$ different injective mappings $\varphi_{n(J)}$ from vertices of J into $\{1, \dots, n\}$. Each mapping $\varphi_{n(J)}$ maps J onto a copy $\varphi_{n(j)}(J)$ of J in a complete graph with n vertices. Then, for the feasible centered subgraph count $\widehat{S}_n(H)$, we get

$$\begin{aligned}
\widehat{S}_n(H) &= \sum_{\varphi} \prod_{e \in E(\varphi(H))} (\mathbb{1}\{e \in E(G)\} - \widehat{p}) = \sum_{\varphi_{n(H)}} \prod_{e \in E(\varphi_{n(H)}(H))} (\mathbb{1}\{e \in E(G)\} - p + p - \widehat{p}) \\
&= \sum_{\substack{J \subseteq H \\ J \text{ induced}}} \left[\sum_{\varphi_{n(J)}} \prod_{e \in E_{n(J)}(\varphi_{n(J)}(H))} (\mathbb{1}\{e \in E(G)\} - p)\right](p - \widehat{p})^{m(H)-m(J)} \\
&= \sum_{\substack{J \subseteq H \\ J \text{ induced}}} \left(\prod_{h=n(H)-k(J)}^{n(H)-1} (n-h)\right) S_n(\widetilde{J})(p - \widehat{p})^{m(H)-m(J)}, \qquad \text{(A1)}
\end{aligned}$$

where $k(J)$ denotes the number of isolated vertices in J, $\widetilde{J}$ denotes the graph obtained from J after removing all isolated vertices, $\prod_{h=n(H)}^{n(H)-1}(n-h) := 1$ and the sum

$\sum_{J\subseteq H, J \text{ induced}}$ ranges over all substructures that can be obtained from H by dropping j, $j = 0, 1, \ldots, m(H)$ edges of the edges present in H. Separating the case, where no edge is dropped at all, we get

$$\widehat{S}_n(H) = S_n(H) + \sum_{\substack{J\subsetneq H \\ J \text{ induced}}} \left(\prod_{h=n(H)-k(J)}^{n(H)-1} (n-h) \right) S_n(\widetilde{J})(p-\widehat{p})^{m(H)-m(J)},$$

where $m(H) - m(J) \geq 1$ by construction. For the sake of readability, we suppress the expression "J is induced" in the index of the sum in the remainder of this proof. As we assumed $H \neq \emptyset$, $H \neq P_2$, using Proposition 1 as well as $S_n(\emptyset) := 1$, we get

$$\begin{aligned}
&\frac{\widehat{S}_n(H) - S_n(H)}{\sqrt{\mathrm{aut}(H)(n)_{n(H)}(p(1-p))^{m(H)}}} \\
&= \frac{\sum_{J\subsetneq H} \left(\prod_{h=n(H)-k(J)}^{n(H)-1} (n-h) \right) S_n(\widetilde{J}) \left(p-\widehat{p}\right)^{m(H)-m(J)}}{\sqrt{\mathrm{aut}(H)(n)_{n(H)}(p(1-p))^{m(H)}}} \\
&= \frac{\sum_{J\subsetneq H_{m(J)>0}} \left(\prod_{h=n(H)-k(J)}^{n(H)-1} (n-h) \right) S_n(\widetilde{J}) \left(p-\widehat{p}\right)^{m(H)-m(J)}}{\sqrt{\mathrm{aut}(H)(n)_{n(H)}(p(1-p))^{m(H)}}} \\
&\quad + \frac{\sum_{J\subsetneq H_{m(J)=0}} \left(\prod_{h=n(H)-k(J)}^{n(H)-1} (n-h) \right) S_n(\widetilde{J}) \left(p-\widehat{p}\right)^{m(H)-m(J)}}{\sqrt{\mathrm{aut}(H)(n)_{n(H)}(p(1-p))^{m(H)}}} \\
&= I + II,
\end{aligned}$$

with an obvious notation for I and II. For the term I, using Proposition 1, which holds for non-empty graphs, we get

$$I = O_P\left(\frac{\sum_{J\subsetneq H_{m(J)>0}} n^{k(J)} n^{n(\widetilde{J})/2} p^{m(\widetilde{J})/2} \left(\frac{\sqrt{p}}{n} \right)^{m(H)-m(J)}}{\sqrt{n^{n(H)}(p(1-p))^{m(H)}}} \right).$$

Further, as $m(\widetilde{J}) = m(J)$ and $n(\widetilde{J}) = n(H) - k(J)$ holds by construction, the last right-hand side becomes

$$\begin{aligned}
&O_P\left(\frac{\sum_{J\subsetneq H_{m(J)>0}} n^{k(J)} n^{\frac{n(H)-k(J)}{2}} p^{m(J)/2} \left(\frac{\sqrt{p}}{n} \right)^{m(H)-m(J)}}{\sqrt{n^{n(H)} p^{m(H)}}} \right) \\
&= O_P\left(\frac{\sum_{J\subsetneq H_{m(J)>0}} n^{\frac{n(H)+k(J)}{2}} p^{\frac{m(J)}{2}} p^{\frac{m(H)-m(J)}{2}} n^{-m(H)+m(J)}}{\sqrt{n^{n(H)} p^{m(H)}}} \right) \\
&= O_P\left(\frac{\sum_{J\subsetneq H_{m(J)>0}} n^{\frac{n(H)}{2}} n^{\frac{k(J)}{2}} p^{\frac{m(H)}{2}} n^{-m(H)+m(J)}}{\sqrt{n^{n(H)} p^{m(H)}}} \right) = O_P\left(\sum_{\substack{J\subsetneq H \\ m(J)>0}} n^{-m(H)+m(J)} n^{\frac{k(J)}{2}} \right).
\end{aligned}$$

Now, as $m(J) > 0$ and H is connected, we have $k(J) \leq m(H) - m(J)$. Hence, as for all J with $J \subsetneq H$, we have $m(H) - m(J) \geq 1$ and as the sum $\sum_{J \subsetneq H}$ is finite, because $n(H)$ is finite, the last right-hand side is bounded by

$$O_P\left(\sum_{\substack{J \subsetneq H \\ m(J)>0}} n^{-m(H)+m(J)} \sqrt{n^{m(H)-m(j)}} \right) = O_P\left(\sum_{\substack{J \subsetneq H \\ m(J)>0}} \sqrt{\left(\frac{1}{n}\right)^{m(H)-m(J)}} \right) \leq O_P\left(\frac{1}{\sqrt{n}} \right) = o_P(1).$$

For the term II, due to $m(J) = 0$ implying $J = \emptyset$ and using $S_n(\emptyset) := 1$, we get

$$II = O_P\left(\frac{n^{n(H)} \left(\frac{\sqrt{p}}{n} \right)^{m(H)}}{\sqrt{n^{n(H)} (p(1-p))^{m(H)}}} \right) = O_P\left(n^{n(H)/2 - m(H)} \right) = o_P(1),$$

as $H \neq P_2$ and H connected leading to $n(H) \geq 3$ and $m(H) \geq n(H) - 1$. ■

Proof of Proposition 6. Following the steps in (A1), we get

$$\begin{aligned}
T_n(H) &= \frac{1}{\mathrm{aut}(H)} \sum_{\varphi} \prod_{e \in E(\varphi(H))} (\mathbb{1}\{e \in E(G)\}) = \frac{1}{\mathrm{aut}(H)} \sum_{\varphi_{n(H)}} \prod_{e \in E(\varphi_{n(H)}(H))} (\mathbb{1}\{e \in E(G)\} - p + p) \\
&= \frac{1}{\mathrm{aut}(H)} \sum_{\substack{J \subseteq H \\ J \text{ induced}}} \left[\sum_{\varphi_{n(J)}} \prod_{e \in E_{n(J)}(\varphi(H))} (\mathbb{1}\{e \in E(G)\} - p) \right] p^{m(H)-m(J)} \\
&= \frac{1}{\mathrm{aut}(H)} \sum_{\substack{J \subseteq H \\ J \text{ induced}}} \left(\prod_{h=n(H)-k(J)}^{n(H)-1} (n-h) \right) S_n(\widetilde{J}) p^{m(H)-m(J)}.
\end{aligned}$$

■

Proof of Theorem 2. Starting with the expected value of $S_n(C_3)$, as all edges form independently, we get

$$\begin{aligned}
\mathbb{E}(S_n(C_3)) &= \sum_{\varphi} \mathbb{E}\left(\prod_{e \in E(\varphi(C_3))} (\mathbb{1}\{e \in E(G)\} - p) \right) = \sum_{\varphi} \prod_{e \in E(\varphi(C_3))} \mathbb{E}\left((\mathbb{1}\{e \in E(G)\} - p) \right) \\
&= \sum_{\varphi} \prod_{e \in E(\varphi(C_3))} (\mathbb{P}\{e \in E(G)\} - p).
\end{aligned}$$

Obviously, for the $\mathcal{G}_{\mathrm{ER}}(n)$ model, we get $\mathbb{E}_{\mathrm{ER}}(S_n(C_3)) = \sum_{\varphi} \prod_{e \in E(\varphi(C_3))} (p - p) = 0$ by construction, where $\mathbb{P}\{e \in E(G)\} = p$. In contrast, for the SBM case, we have

$$\begin{aligned}
&\mathbb{E}_{\mathrm{SBM}}(S_n(C_3)) \\
&\quad = \sum_{\varphi} \prod_{e \in E(\varphi(C_3))} (\mathbb{P}\{e \in E(G)\} - p) \\
&\quad = n\left(\frac{n}{2} - 1\right)\left(\frac{n}{2} - 2\right)(p_{\mathrm{intra}} - p)^3 + n\left(\frac{n}{2} - 1\right)\left(\frac{n}{2}\right)(p_{\mathrm{intra}} - p)(p_{\mathrm{inter}} - p)^2
\end{aligned}$$

$$
\begin{aligned}
&+ n\left(\frac{n}{2}\right)\left(\frac{n}{2}-1\right)(p_{\text{inter}}-p)(p_{\text{intra}}-p)(p_{\text{inter}}-p) + n\left(\frac{n}{2}\right)\left(\frac{n}{2}-1\right)(p_{\text{inter}}-p)^2(p_{\text{intra}}-p)\\
&= n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)(p_{\text{intra}}-p)^3 + 3n\left(\frac{n}{2}\right)\left(\frac{n}{2}-1\right)(p_{\text{intra}}-p)(p_{\text{inter}}-p)^2.
\end{aligned}
$$

Next, we express $p = p_{\text{mean}}$ in terms of p_{intra} and p_{inter}, which results in $p = \frac{n-2}{2n-2}p_{\text{intra}} + \frac{n}{2n-2}p_{\text{inter}}$. Therefore, the last right-hand side becomes

$$
\begin{aligned}
&n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\left(p_{\text{intra}} - \left(\frac{n-2}{2n-2}p_{\text{intra}} + \frac{n}{2n-2}p_{\text{inter}}\right)\right)^3\\
&\quad + 3n\left(\frac{n}{2}\right)\left(\frac{n}{2}-1\right)\left(p_{\text{intra}} - \left(\frac{n-2}{2n-2}p_{\text{intra}} + \frac{n}{2n-2}p_{\text{inter}}\right)\right)\\
&\quad \times \left(p_{\text{inter}} - \left(\frac{n-2}{2n-2}p_{\text{intra}} + \frac{n}{2n-2}p_{\text{inter}}\right)\right)^2\\
&= n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\left(\frac{n}{2n-2}p_{\text{intra}} - \frac{n}{2n-2}p_{\text{inter}}\right)^3\\
&\quad + 3n\left(\frac{n}{2}\right)\left(\frac{n}{2}-1\right)\left(\frac{n}{2n-2}p_{\text{intra}} - \frac{n}{2n-2}p_{\text{inter}}\right)\Big)\left(\frac{n-2}{2n-2}p_{\text{inter}} - \frac{n-2}{2n-2}p_{\text{intra}}\right)^2,
\end{aligned}
$$

which is asymptotically equivalent to

$$
\begin{aligned}
&n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\frac{1}{8}(p_{\text{intra}}-p_{\text{inter}})^3 + 3n\left(\frac{n}{2}\right)\left(\frac{n}{2}-1\right)\frac{1}{8}(p_{\text{intra}}-p_{\text{inter}})(p_{\text{inter}}-p_{\text{intra}})^2\\
&= n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\frac{1}{8}\epsilon^3 + 3n\left(\frac{n}{2}\right)\left(\frac{n}{2}-1\right)\frac{1}{8}\epsilon(-\epsilon)^2\\
&= \frac{1}{8}n\left(\frac{n}{2}-1\right)\left[\frac{n}{2}-2+\frac{3}{2}n\right]\epsilon^3 = \frac{n\left(\frac{n}{2}-1\right)(2n-2)}{8}\epsilon^3 = \frac{n(n-1)(n-2)}{8}\epsilon^3 = O(n^3\epsilon^3).
\end{aligned}
$$

With the help of Proposition 1, we can determine the variance for the ER-model as

$$
\mathbb{V}\text{ar}_{\text{ER}}(S_n(C_3)) = 6n(n-1)(n-2)p^3(1-p)^3 = 36\binom{n}{3}p^3(1-p)^3 = O(n^3p^3).
$$

Altogether, asymptotically, we get

$$
\frac{\mathbb{E}_{\text{SBM}}(S_n(C_3)) - \mathbb{E}_{\text{ER}}(S_n(C_3))}{\sqrt{\mathbb{V}\text{ar}_{\text{ER}}(S_n(C_3))}} = O\left(\frac{n^3\epsilon^3}{\sqrt{n^3p^3}}\right).
$$

Considering the case of fixed alternatives with $\epsilon \sim p$, we get the following rates: if $p \to p_0 \in (0,1)$, we get an $O(n^{3/2})$. If $p \to 0$, we get an $O((np)^{3/2})$. Hence, if $np \to \infty$, we get divergence to $+\infty$ in all cases. For the case of local alternatives, where $\epsilon = \epsilon_n$ is allowed to decrease to 0 at a faster rate than $p = p_n$, we get the following: if $p \to p_0 \in (0,1)$, we get an $O(n^{3/2}\epsilon^3)$. If $p \to 0$, we get an $O(n^{3/2}\epsilon^3 p^{-3/2})$.

Further, continuing with the expectation of $T_n(C_3)$, we have

$$
\mathbb{E}(T_n(C_3)) = \frac{1}{\text{aut}(C_3)}\sum_{\varphi}\prod_{e\in E(\varphi(C_3))}\mathbb{P}\{e \in E(G)\}.
$$

For the $\mathcal{G}_{\mathrm{ER}}(n)$ model, this obviously simplifies to

$$\mathbb{E}_{\mathrm{ER}}(T_n(C_3)) = \frac{1}{6}n(n-1)(n-2)p^3 = \binom{n}{3}p^3.$$

For the SBM case, we have

$$\begin{aligned}
&\mathbb{E}_{\mathrm{SBM}}(T_n(C_3))\\
&= \frac{1}{6}\left[n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)p_{\mathrm{intra}}^3 + 3n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)p_{\mathrm{intra}}p_{\mathrm{inter}}^2\right]\\
&= \frac{1}{6}\left[n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)(p_{\mathrm{intra}}-p+p)^3 + 3n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)(p_{\mathrm{intra}}-p+p)(p_{\mathrm{inter}}-p+p)^2\right]\\
&= \frac{1}{6}\left\{n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\left[(p_{\mathrm{intra}}-p)^3 + 3(p_{\mathrm{intra}}-p)^2p + 3(p_{\mathrm{intra}}-p)p^2 + p^3\right]\right.\\
&\quad + 3n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\Big((p_{\mathrm{intra}}-p)\left[(p_{\mathrm{inter}}-p)^2 + 2(p_{\mathrm{inter}}-p)p + p^2\right]\\
&\quad \left.+\ p\left[(p_{\mathrm{inter}}-p)^2 + 2(p_{\mathrm{inter}}-p)p + p^2\right]\Big)\right\}.
\end{aligned}$$

Analogously to before, we express p in terms of p_{intra} and p_{inter}, which results in $p = \frac{n-2}{2n-2}p_{\mathrm{intra}} + \frac{n}{2n-2}p_{\mathrm{inter}}$. This is asymptotically equal to $p = \frac{1}{2}p_{\mathrm{intra}} + \frac{1}{2}p_{\mathrm{inter}}$. Hence, we have

$$\begin{aligned}
&\frac{1}{6}\left\{n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\left[\frac{1}{8}\epsilon^3 + \frac{3}{4}\epsilon^2p + \frac{3}{2}\epsilon p^2 + p^3\right]\right.\\
&\qquad \left.+\ 3n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\left(\frac{1}{2}\epsilon\left[\frac{1}{4}\epsilon^2 + (-\epsilon)p + p^2\right] + p\left[\frac{1}{4}(-\epsilon)^2 + 2(-\epsilon)p + p^2\right]\right)\right\}\\
&\quad = \frac{1}{6}\left\{n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\left[\frac{1}{8}\epsilon^3 + \frac{3}{4}\epsilon^2p + \frac{3}{2}\epsilon p^2 + p^3\right]\right.\\
&\qquad \left.+\ 3n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\left[\frac{1}{8}\epsilon^3 - \frac{1}{2}\epsilon^2p + \frac{1}{2}\epsilon p^2 + \frac{1}{4}\epsilon^2p - \epsilon p^2 + p^3\right]\right\}\\
&\quad = \frac{1}{6}\left\{\left[n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right) + 3n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\right]\frac{1}{8}\epsilon^3\right.\\
&\qquad + \left[n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\frac{3}{4} - \frac{3}{4}n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\right]\epsilon^2p\\
&\qquad + \left[n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\frac{3}{2} - \frac{3}{2}n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\right]\epsilon p^2\\
&\qquad \left.+ \left[n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right) + 3n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\right]p^3\right\}\\
&\quad = \frac{1}{6}\left\{n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2+\frac{3}{2}n\right)\frac{1}{8}\epsilon^3 + n\left(\frac{n}{2}-1\right)\frac{3}{4}\left(\frac{n}{2}-2-\frac{n}{2}\right)\epsilon^2p\right.\\
&\qquad \left.+\ n\left(\frac{n}{2}-1\right)\frac{3}{2}\left(\frac{n}{2}-2-\frac{n}{2}\right)\epsilon p^2 + n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2+\frac{3}{2}n\right)p^3\right\}\\
&\quad = \frac{1}{6}\left[\frac{n(n-1)(n-2)}{8}\epsilon^3 - \frac{3n(n-2)}{4}\epsilon^2p - \frac{3n(n-2)}{2}\epsilon p^2 + n(n-1)(n-2)p^3\right].
\end{aligned}$$

Hence, it yields

$$\begin{aligned}
&\mathbb{E}_{\mathrm{SBM}}(T_n(C_3)) - \mathbb{E}_{\mathrm{ER}}(T_n(C_3)) \\
&\quad = \frac{1}{6}\left[\frac{n(n-1)(n-2)}{8}\epsilon^3 - \frac{3n(n-2)}{4}\epsilon^2 p - \frac{3n(n-2)}{2}\epsilon p^2 + n(n-1)(n-2)p^3\right] - \binom{n}{3}p^3 \\
&\quad = \frac{1}{6}\left[\frac{n(n-1)(n-2)}{8}\epsilon^3 - \frac{3n(n-2)}{4}\epsilon^2 p - \frac{3n(n-2)}{2}\epsilon p^2\right] \\
&\quad = O(n^3\epsilon^3) + O(n^2\epsilon^2 p) + O(n^2\epsilon p^2).
\end{aligned}$$

With (15) from Proposition 6, we get

$$\mathbb{V}\mathrm{ar}_{\mathrm{ER}}(T_n(C_3)) = \binom{n}{3}p^3(1-p)^3 + 3\binom{n}{3}p^4(1-p)^2 + 3(n-2)\binom{n}{3}p^5(1-p).$$

Altogether, asymptotically, we get

$$\frac{\mathbb{E}_{\mathrm{SBM}}(T_n(C_3)) - \mathbb{E}_{\mathrm{ER}}(T_n(C_3))}{\sqrt{\mathbb{V}\mathrm{ar}_{\mathrm{ER}}(T_n(C_3))}} = O\left(\frac{n^3\epsilon^3 + n^2\epsilon^2 p + n^2\epsilon p^2}{\sqrt{n^3p^3 + n^3p^4 + n^4p^5}}\right).$$

Considering the case of fixed alternatives with $\epsilon \sim p$, we get the following rates: if $p \to p_0 \in (0,1)$, we get an $O(n)$. If $p \to 0$ and $np^2 \to \infty$, we get an $O(n\sqrt{p})$. If $p \to 0$ and $np^2 \to K \geq 0$, we get an $O((np)^{3/2})$. Hence, if $np \to \infty$, we get divergence to $+\infty$ in all three cases. For the case of local alternatives, where $\epsilon = \epsilon_n$ is allowed to decrease to 0 at a faster rate than $p = p_n$, we get the following: if $n\epsilon \to \infty$ and $np^2 \to \infty$, we get an $O(n\epsilon^3 p^{-5/2})$, if $n\epsilon \to K_\epsilon \geq 0$ and $np^2 \to \infty$, we get an $O(\epsilon p^{-5/2})$. If $n\epsilon \to \infty$ and $np^2 \to K \geq 0$, we get an $O(n^{3/2}\epsilon^3 p^{-3/2})$, if $n\epsilon \to K_\epsilon \geq 0$ and $np^2 \to K \geq 0$, we get an $O(n^{1/2}\epsilon p^{-3/2})$. ∎

Proof of Theorem 3. Analogously to the proof of Theorem 2, we have

$$\mathbb{E}(S_n(P_3)) = \sum_{\varphi} \prod_{e \in E(\varphi(P_3))} (\mathbb{P}\{e \in E(G)\} - p)$$

Again, it is $\mathbb{E}_{\mathrm{ER}}(S_n(P_3)) = \sum_{\varphi} \prod_{e \in E(\varphi(P_3))} (p - p) = 0$ for the $\mathcal{G}_{\mathrm{ER}}(n)$ model. For the SBM case, we have

$$\begin{aligned}
&\mathbb{E}_{\mathrm{SBM}}(S_n(P_3)) \\
&\quad = \sum_{\varphi} \prod_{e \in E(\varphi(P_3))} (\mathbb{P}\{e \in E(G)\} - p) \\
&\quad = n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)(p_{\mathrm{intra}} - p)^2 + n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)(p_{\mathrm{inter}} - p)^2 \\
&\qquad + n\left(\frac{n}{2}\right)\left(\frac{n}{2}-1\right)(p_{\mathrm{inter}} - p)(p_{\mathrm{intra}} - p) + n\left(\frac{n}{2}\right)\left(\frac{n}{2}-1\right)(p_{\mathrm{intra}} - p)(p_{\mathrm{inter}} - p) \\
&\quad = n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)(p_{\mathrm{intra}} - p)^2 + n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)(p_{\mathrm{inter}} - p)^2 \\
&\qquad + 2n\left(\frac{n}{2}\right)\left(\frac{n}{2}-1\right)(p_{\mathrm{intra}} - p)(p_{\mathrm{inter}} - p).
\end{aligned}$$

We again express p of the ER model in terms of p_{intra} and p_{inter}, which results in $p = \frac{n-2}{2n-2}p_{\text{intra}} + \frac{n}{2n-2}p_{\text{inter}}$. The last right-hand side therefore is

$$
\begin{aligned}
& n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\left(p_{\text{intra}} - \left(\frac{n-2}{2n-2}p_{\text{intra}} + \frac{n}{2n-2}p_{\text{inter}}\right)\right)^2 \\
&\quad + n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\left(p_{\text{inter}} - \left(\frac{n-2}{2n-2}p_{\text{intra}} + \frac{n}{2n-2}p_{\text{inter}}\right)\right)^2 \\
&\quad + 2n\left(\frac{n}{2}\right)\left(\frac{n}{2}-1\right) \\
&\quad \left(p_{\text{intra}} - \left(\frac{n-2}{2n-2}p_{\text{intra}} + \frac{n}{2n-2}p_{\text{inter}}\right)\right)\left(p_{\text{inter}} - \left(\frac{n-2}{2n-2}p_{\text{intra}} + \frac{n}{2n-2}p_{\text{inter}}\right)\right) \\
&= n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\left(\frac{n}{2n-2}p_{\text{intra}} - \frac{n}{2n-2}p_{\text{inter}}\right)^2 \\
&\quad + n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\left(\frac{n-2}{2n-2}p_{\text{inter}} - \frac{n-2}{2n-2}p_{\text{intra}}\right)^2 \\
&\quad + 2n\left(\frac{n}{2}\right)\left(\frac{n}{2}-1\right) \\
&\quad \left(\frac{n}{2n-2}p_{\text{intra}} - \frac{n}{2n-2}p_{\text{inter}}\right)\left(\frac{n-2}{2n-2}p_{\text{inter}} - \frac{n-2}{2n-2}p_{\text{intra}}\right),
\end{aligned}
$$

which is asymptotically equivalent to

$$
\begin{aligned}
& n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\frac{1}{4}(p_{\text{intra}} - p_{\text{inter}})^2 + n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\frac{1}{4}(p_{\text{inter}} - p_{\text{intra}})^2 \\
&\quad + 2n\left(\frac{n}{2}\right)\left(\frac{n}{2}-1\right)\frac{1}{2}(p_{\text{intra}} - p_{\text{inter}})\frac{1}{2}(p_{\text{inter}} - p_{\text{intra}}) \\
&= n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\frac{1}{4}\epsilon^2 + n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\frac{1}{4}(-\epsilon)^2 + 2n\left(\frac{n}{2}\right)\left(\frac{n}{2}-1\right)\frac{1}{2}\epsilon\frac{1}{2}(-\epsilon) \\
&= \frac{n(\frac{n}{2}-1)}{4}\left[\frac{n}{2} - 2 + \frac{n}{2} - 2\frac{n}{2}\right]\epsilon^2 \\
&= \frac{n(\frac{n}{2}-1)}{4}(-2)\epsilon^2 = -\frac{n(n-2)}{4}\epsilon^2 = O(n^2\epsilon^2).
\end{aligned}
$$

With the help of Proposition 1, we can determine the variance for the ER-model as

$$
\mathbb{V}\text{ar}_{\text{ER}}(S_n(P_3)) = 2n(n-1)(n-2)p^2(1-p)^2 = 12\binom{n}{3}p^2(1-p)^2 = O(n^3p^2)
$$

Altogether, asymptotically, we get

$$
\frac{\mathbb{E}_{\text{SBM}}(S_n(P_3)) - \mathbb{E}_{\text{ER}}(S_n(P_3))}{\sqrt{\mathbb{V}\text{ar}_{\text{ER}}(S_n(P_3))}} = O\left(\frac{n^2\epsilon^2}{\sqrt{n^3p^2}}\right)
$$

Considering the case of fixed alternatives with $\epsilon \sim p$, we get the following rates: if $p \to p_0 \in (0,1)$, we get an $O(n^{1/2})$. If $p \to 0$, we get an $O(n^{1/2}p)$. Hence, if $np^2 \to \infty$, we get divergence to $+\infty$ in all cases. For the case of local alternatives, where $\epsilon = \epsilon_n$ is allowed to decrease to 0 at a faster rate than $p = p_n$, we get the following: if $p \to p_0 \in (0,1)$, we get an $O(n^{1/2}\epsilon^2)$. If $p \to 0$, we get an $O(n^{1/2}\epsilon^2p^{-1})$.

Further, let us now move on to the raw subgraph counts. It yields

$$\mathbb{E}(T_n(P_3)) = \frac{1}{\mathrm{aut}(P_3)} \sum_{\varphi} \prod_{e \in E(\varphi(C_3))} \mathbb{P}\{e \in E(G)\}.$$

For the $\mathcal{G}_{\mathrm{ER}}(n)$ model, this obviously simplifies to

$$\mathbb{E}_{\mathrm{ER}}(T_n(P_3)) = \frac{1}{2} n(n-1)(n-2)p^2 = 3\binom{n}{3}p^2.$$

For the SBM case, we have

$$\begin{aligned}
&\mathbb{E}_{\mathrm{SBM}}(T_n(P_3)) \\
&\quad = \frac{1}{2}\Big[n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)p_{\mathrm{intra}}^2 + n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)p_{\mathrm{inter}}^2 \\
&\qquad + 2n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)p_{\mathrm{intra}}p_{\mathrm{inter}}\Big] \\
&\quad = \frac{1}{2}\Big[n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)(p_{\mathrm{intra}}-p+p)^2 + n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)(p_{\mathrm{inter}}-p+p)^2 \\
&\qquad + 2n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)(p_{\mathrm{intra}}-p+p)(p_{\mathrm{inter}}-p+p)\Big] \\
&\quad = \frac{1}{2}\Big\{n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\left[(p_{\mathrm{intra}}-p)^2 + 2(p_{\mathrm{intra}}-p)p + p^2\right] \\
&\qquad + n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\left[(p_{\mathrm{inter}}-p)^2 + 2(p_{\mathrm{inter}}-p)p + p^2\right] \\
&\qquad + 2n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\left[(p_{\mathrm{inter}}-p)(p_{\mathrm{intra}}-p) + (p_{\mathrm{intra}}-p)p + (p_{\mathrm{inter}}-p)p + p^2\right]\Big\}.
\end{aligned}$$

Again, we express some p in terms of p_{intra} and p_{inter}, which results in $p = \frac{n-2}{2n-2}p_{\mathrm{intra}} + \frac{n}{2n-2}p_{\mathrm{inter}}$. This is asymptotically equivalent to $p = \frac{1}{2}p_{\mathrm{intra}} + \frac{1}{2}p_{\mathrm{inter}}$. Hence, we have

$$\begin{aligned}
&\frac{1}{2}\Big\{n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\left[\frac{1}{4}\epsilon^2 + \epsilon p + p^2\right] + n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\left[-\frac{1}{4}\epsilon^2 - \epsilon p + p^2\right] \\
&\qquad + 2n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\left[-\frac{1}{4}\epsilon^2 + p^2\right]\Big\} \\
&\quad = \frac{1}{2}\Big\{\Big[n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right) - n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\Big]\frac{1}{4}\epsilon^2 + \Big[n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right) \\
&\qquad - n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\Big]\epsilon p + \Big[n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right) + 3n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)\Big]p^2\Big\} \\
&\quad = \frac{1}{2}\Big\{n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2+\frac{n}{2}-\frac{2n}{2}\right)\frac{1}{4}\epsilon^2 \\
&\qquad + n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2-\frac{n}{2}\right)\epsilon p + n\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2+\frac{3}{2}n\right)p^2\Big\} \\
&\quad = \frac{1}{2}\left[-\frac{n(n-2)}{4}\epsilon^2 - n(n-2)\epsilon p + n(n-1)(n-2)p^2\right].
\end{aligned}$$

Hence, it yields

$$\begin{aligned}\mathbb{E}_{\mathrm{SBM}}(T_n(P_3)) - \mathbb{E}_{\mathrm{ER}}(T_n(P_3)) &= \frac{1}{2}\left[-\frac{n(n-2)}{4}\epsilon^2 - n(n-2)\epsilon p + n(n-1)(n-2)p^2\right] - 3\binom{n}{3}p^2 \\ &= \frac{1}{2}\left[-\frac{n(n-2)}{4}\epsilon^2 - n(n-2)\epsilon p\right] \\ &= O(n^2\epsilon^2) + O(n^2\epsilon p).\end{aligned}$$

Together with (15) from Proposition 6, we get

$$\begin{aligned}\mathbb{V}\mathrm{ar}_{\mathrm{ER}}(T_n(P_3)) &= \mathbb{V}\mathrm{ar}\left(\frac{1}{2}\left[S_n(P_3) + (n-2)S_n(P_2)p + n(n-1)(n-2)S_n(\emptyset)p^2\right]\right) \\ &= \frac{1}{2}n(n-1)(n-2)p^2(1-p)^2 + \frac{1}{2}(n-2)^2 n(n-1)p^3(1-p) \\ &= 3\binom{n}{3}p^2(1-p)^2 + 3(n-2)\binom{n}{3}p^3(1-p) = O(n^3p^2) + O(n^4p^3).\end{aligned}$$

Altogether, asymptotically, we get

$$\frac{\mathbb{E}_{\mathrm{SBM}}(T_n(P_3)) - \mathbb{E}_{\mathrm{ER}}(T_n(P_3))}{\sqrt{\mathbb{V}\mathrm{ar}_{\mathrm{ER}}(T_n(P_3))}} = O\left(\frac{n^2\epsilon^2 + n^2\epsilon p}{\sqrt{n^3p^2 + n^4p^3}}\right).$$

Considering the case of fixed alternatives with $\epsilon \sim p$, we get the following rates: if $p \to p_0 \in (0,1)$, we get an $O(1)$. If $p \to 0$ and $np \to \infty$, we get an $O(\sqrt{p})$. If $p \to 0$ and $np \to K \geq 0$, we get an $O(\sqrt{np})$. Hence, we do not get divergence to $+\infty$ in all cases. For the case of local alternatives, where $\epsilon = \epsilon_n$ is allowed to decrease to 0 at a faster rate than $p = p_n$, we get the following: if $p \to p_0 \in (0,1)$, we get $O(\epsilon)$, if $p \to 0$ and $np \to \infty$, we get $O(\epsilon p^{-1/2})$, and if $p \to 0$ and $np \to K \geq 0$, we get $O(n^{1/2}\epsilon)$. ∎

Proof of Theorem 4. By assumption, $G \in \mathcal{G}_{\mathrm{HER}}$, that is, G is generated by a heterogeneous ER-graph $\mathcal{G}(n, \boldsymbol{P})$ with some $\boldsymbol{P} = (p_{ij})$, $p_{ij} = p_{ji}$, $p_{ij} \in [0,1)$ for all i,j such that its mean connectivity $p_{\mathrm{mean}} = \binom{n}{2}^{-1}\sum_{1\leq i<j\leq n} p_{n,ij}$ satisfies $p_{\mathrm{mean}} = p_{\mathrm{mean}}(n) \xrightarrow[n\to\infty]{} p_0 \in [0,1]$. Then, conditional on $\boldsymbol{A} = (A_{ij})_{1\leq i,j\leq n}$, G^* is generated by a homogeneous ER-graph $\mathcal{G}(n, \hat{p})$, where $\hat{p} = \binom{n}{2}^{-1}\sum_{1\leq i<j\leq n} A_{ij}$. Furthermore, we have that, under the null, $X_n = \sum_{H\in\mathcal{H}} a_n(H)S_n(H)$, which is dominated by some family of connected graphs $\mathcal{H}_0$ such that for all $H \in \mathcal{H}_0$, it holds $np_{\mathrm{mean}}^{r(H)} \xrightarrow[n\to\infty]{} \infty$. Hence, under both the null and the alternative, we immediately get $X_n^* = \sum_{H\in\mathcal{H}_0} a_n^*(H)S_n^*(H)$, which is dominated by the same family of connected graphs $\mathcal{H}_0$, and where $a_n^*(H) = a_n(H, \hat{p})$ for $a_n(H) = a_n(H, p)$.

Hence, in view of Theorem 1, we have to check, whether for all $H \in \mathcal{H}_0$, we have that $n\hat{p}^{r(H)} \xrightarrow[n\to\infty]{} \infty$ in probability, the coefficients

$$\hat{b}(H) = \sup_n \frac{n^{n(H)/2}\hat{p}^{m(H)/2}a_n^*(H)}{\sqrt{\mathbb{V}\mathrm{ar}^*(X_n^*)}}, \tag{A2}$$

are bounded in probability and satisfy

$$\sum_{H \in \mathcal{H}_0} \hat{b}^2(H)|\mathrm{aut}(H)| = O_P(1), \tag{A3}$$

respectively. For this purpose, let us consider $\hat{p} = \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} A_{ij}$ in more detail. To be more precise, as p_{ij}, and consequently also the distribution of A_{ij}, are allowed to depend on n, we will write $A_{n,ij}$ and $p_{n,ij}$ in the following. For the expectation, we get

$$\mathbb{E}(\hat{p}) = \mathbb{E}\left(\binom{n}{2}^{-1} \sum_{1 \le i < j \le n} A_{n,ij}\right) = \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} \mathbb{E}(A_{n,ij}) = \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} p_{n,ij} = p_{\mathrm{mean}}.$$

For the variance, as the edges are formed independently according to Bernoulli distributions with connection probabilities $p_{n,ij}$, we have

$$\mathbb{V}\mathrm{ar}(\hat{p}) = \binom{n}{2}^{-2} \sum_{1 \le i < j \le n} \mathbb{V}\mathrm{ar}(A_{n,ij}) = \binom{n}{2}^{-2} \sum_{1 \le i < j \le n} p_{n,ij}(1 - p_{n,ij}).$$

Hence, this leads to

$$\hat{p} - p_{\mathrm{mean}} = O_P\left(\sqrt{\binom{n}{2}^{-2} \sum_{1 \le i < j \le n} p_{n,ij}(1 - p_{n,ij})}\right) \le O_P\left(\frac{\min\{p_{\mathrm{mean}}, 1 - p_{\mathrm{mean}}\}}{n}\right) \le O_P\left(\frac{1}{n}\right), \tag{A4}$$

due to $\binom{n}{2} = n^2$ and $1 - p_{n,ij} \le 1$ and $\sup_{i,j} p_{ij} = o(1)$ if $p_0 = 0$ and $\sup_{n;i,j} p_{ij} < 1$ and $\inf_{n;i,j} p_{ij} > 0$ if $p_0 \in (0, 1)$. Then, a Taylor series argument leads to

$$\begin{aligned} np^{r(H)} &= n\left(p^{r(H)} + r(H)\tilde{p}^{r(H)-1}(\hat{p} - p)\right) = np^{r(H)} + O_P\left(nr(h)\tilde{p}^{r(H)-1}\frac{\sqrt{p}}{n}\right) \\ &= np^{r(H)} + O_P\left(n\left(\frac{\sqrt{p}}{n}\right)^{r(H)-1}\frac{\sqrt{p}}{n}\right) = np^{r(H)} + O_P\left(n\left(\frac{\sqrt{p}}{n}\right)^{r(H)}\right) \\ &\to \infty, \end{aligned}$$

in probability, because $r(h) \ge 1$. To show that (A2) is bounded in probability, note that $X_n^* = \sum_{H \in \mathcal{H}_0} a_n^*(H) S_n^*(H)$ with $a_n^*(H) = a_n(H, \hat{p})$ and

$$\mathbb{V}\mathrm{ar}^*(X_n^*) = \sum_{H \in \mathcal{H}_0} a_n^{*2}(H) \mathbb{V}\mathrm{ar}(S_n^*(H)),$$

where, according to Proposition 1,

$$\mathbb{V}\mathrm{ar}^*(S_n^*(H)) = |\mathrm{aut}(H)|(n)_{n(H)}(\hat{p}(1 - \hat{p}))^{m(H)}.$$

By plugging-in, we get

$$\hat{b}(H) = \sup_n \frac{n^{n(H)/2}\hat{p}^{m(H)/2}a_n(H)}{\sqrt{\mathbb{V}\mathrm{ar}^*(X_n^*)}} = \sup_n \frac{n^{n(H)/2}\hat{p}^{m(H)/2}a_n(H)}{\sqrt{|\mathrm{aut}(H)|(n)_{n(H)}(\hat{p}(1-\hat{p}))^{m(H)}}}$$

which remains bounded in probability by similar arguments as above due to (16) and making use of (A4). Finally, using the same arguments, (A3) follows from (17). ■

APPENDIX B. ADDITIONAL TABLES AND FIGURES

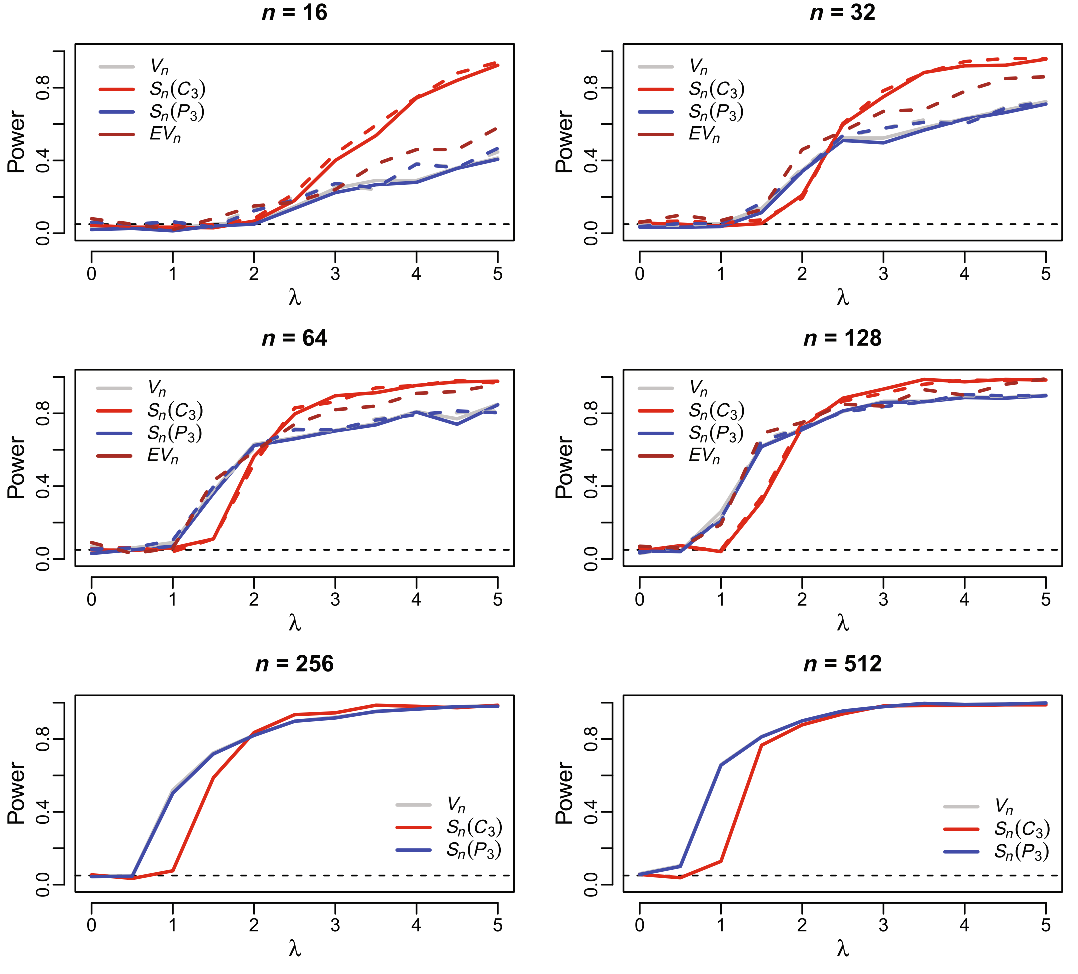

FIGURE B1 Power of the tests V_n, E_n, $S_n(C_3)$ and $S_n(P_3)$ for the alternative of a SBM setup with three blocks with $p_{\text{mean}} = \frac{\log n}{\sqrt{n}}$. The asymptotic versions of the tests are represented by solid lines and the bootstrap version by dashed lines.