

HIGHER-ORDER GRAPHON THEORY: FLUCTUATIONS, DEGENERACIES, AND INFERENCE

ANIRBAN CHATTERJEE, SOHAM DAN, AND BHASWAR B. BHATTACHARYA

ABSTRACT. Exchangeable random graphs, which include some of the most widely studied network models, have emerged as the mainstay of statistical network analysis in recent years. Graphons, which are the central objects in graph limit theory, provide a natural way to sample exchangeable random graphs. It is well known that network moments (motif/subgraph counts) identify a graphon (up to an isomorphism), hence, understanding the sampling distribution of subgraph counts in random graphs sampled from a graphon is pivotal for nonparametric network inference. Although there are a few results regarding the asymptotic normality of subgraph counts in graphon models, for many commonly appearing graphons this distribution is degenerate. This degeneracy phenomenon was overlooked until very recently and its consequences in network inference have remained unexplored. Towards this, we obtain the following results:

- We derive the joint asymptotic distribution of any finite collection of network moments in random graphs sampled from a graphon, that includes both the non-degenerate case (where the distribution is Gaussian) as well as the degenerate case (where the distribution has both Gaussian or non-Gaussian components). This provides the higher-order fluctuation theory for subgraph counts in the graphon model.
- We develop a novel multiplier bootstrap for graphons that consistently approximates the limiting distribution of the network moments (both in the Gaussian and non-Gaussian regimes). Using this and a procedure for testing degeneracy, we construct joint confidence sets for any finite collection of motif densities. This provides a general framework for statistical inference based on network moments in the graphon model.

Examples and simulations are provided to validate the general theory. To illustrate the broad scope of our results we also consider the problem of detecting global structure (that is, testing whether the graphon is a constant function) based on small subgraphs. We propose a consistent test for this problem, invoking celebrated results on quasi-random graphs, and derive its limiting distribution both under the null and the alternative.

1. INTRODUCTION

Networks provide a convenient way to represent complex relational data. The ubiquitous presence of network data in recent years has led to the development of several probabilistic models for random graphs that aim to capture various features of real-world networks. One of the most extensively studied models for network data are exchangeable random graphs [1, 10, 11, 21, 24, 38, 55], where the distribution of the network, given the location of the nodes, remains unchanged under permutations of the node labels. The celebrated Aldous-Hoover theorem [1, 38] shows that any exchangeable random graph of infinite size can be generated by first sampling independent node variables $\{U_i\}_{i \geq 1}$ uniformly on $[0, 1]$, and then connecting each pair of nodes (i, j) independently with probability $W(U_i, U_j)$, for some measurable function $W : [0, 1]^2 \rightarrow [0, 1]$ which is symmetric, that is, $W(x, y) = W(y, x)$, for all $x, y \in [0, 1]$. The function W is commonly referred to as a *graphon*. Graphons arise as limits of sequences of dense graphs and is the fundamental object in graph limit theory [14, 16, 55]. The theory of

2010 *Mathematics Subject Classification.* 05C80, 60F05, 05C60.

Key words and phrases. Inhomogeneous random graphs, network analysis, generalized U -statistics, subgraph counts.

graph limits has been extensively studied since its inception and is the backbone of several beautiful results in combinatorics, probability, statistics and related areas (see [55] for a book length treatment). As mentioned before, graphons provide a natural way for sampling finite exchangeable random graphs, a concept that has appeared independently in various contexts (see [11–13, 23, 35, 56] among others). We describe this formally in the following definition:

Definition 1.1 (Graphon random graph model). Given a graphon $W : [0, 1]^2 \rightarrow [0, 1]$, a W -random graph on the set of vertices $[n] := \{1, 2, \dots, n\}$, hereafter denoted by $G(n, W)$, is obtained by connecting the vertices i and j with probability $W(U_i, U_j)$ independently for all $1 \leq i < j \leq n$, where $\{U_i : 1 \leq i \leq n\}$ is an i.i.d. sequence of $U[0, 1]$ random variables. An alternative way to achieve this sampling is to generate i.i.d. sequences $\{U_i : 1 \leq i \leq n\}$ and $\{Y_{ij} : 1 \leq i < j \leq n\}$ of $U[0, 1]$ random variables and then assigning the edge (i, j) whenever $\{Y_{ij} \leq W(U_i, U_j)\}$, for $1 \leq i < j \leq n$.

The model in Definition 1.1 will be referred to as the W -random model or the graphon random graph model. This includes many well-known network models such as, the classical Erdős-Rényi random graph model (where $W = W_p \equiv p \in [0, 1]$ is the constant function), the stochastic block model [10, 37] (and its many variations), smooth graphons [30], random dot-product graphs [3, 73] (see also Lei [52]), and random geometric graphs [69], among others.

Network moments or motif counts are the frequencies of particular patterns (subgraphs) in a network, such as the number/density of edges, triangles, or stars in a network [2, 62, 78]. Motif counts encode structural information about the geometry of a network and are important summary statistics for potentially large networks. They are the building blocks of network models, such as Exponential Random Graph Models (ERGMs) [19, 40, 64, 65, 76, 83, 84], and many features of a network of practical interest can be derived from the motif counts, such as clustering coefficient [81], degree distribution [70], and transitivity [36] (see [74] for others). This has propelled the fast growing literature on counting and estimating network motifs under various sampling models (see [6, 7, 25, 32, 48, 59] and the references therein).

In the framework of the graphon model, network method of moments, introduced in the seminal papers [11, 15], is an important tool for inferring properties of the underlying graphon based on the motif counts of the observed network. This makes understanding the asymptotic properties of subgraph counts in W -random graphs a problem of central importance in network analysis. To this end, suppose G_n is the observed graph sampled from the W -random model $G(n, W)$. Then for a finite simple graph¹ (motif) $H = (V(H), E(H))$, with $V(H) = \{1, 2, \dots, |V(H)|\}$ such that $|V(H)| \geq 2$, the H -th empirical network moment is the number of copies of H in G_n . This will be denoted by $X(H, G_n)$, which can be expressed more formally as:

$$X(H, G_n) = \sum_{1 \leq i_1 < \dots < i_{|V(H)|} \leq n} \sum_{H' \in \mathcal{G}_H(\{i_1, \dots, i_{|V(H)|}\})} \prod_{(i_s, i_t) \in E(H')} \mathbf{1}\{Y_{i_s i_t} \leq W(U_{i_s}, U_{i_t})\}, \quad (1.1)$$

where, for any set $S \subseteq [n]$, $\mathcal{G}_H(S)$ denotes the collection of all subgraphs of the complete graph $K_{|S|}$ on the vertex set S which are isomorphic to H .² Note that

$$|\mathcal{G}_H(\{1, \dots, |V(H)|\})| = \frac{|V(H)|!}{|\text{Aut}(H)|},$$

where $\text{Aut}(H)$ is the set of all automorphisms of the graph H , that is, the collection of permutations of the vertex set $V(H)$ such that $(x, y) \in E(H)$ if and only if $(\sigma(x), \sigma(y)) \in E(H)$.

¹A graph is said to be *simple* if it has no self-loops and does not contain more than one edge between a pair of vertices.

²Note that we count unlabelled copies of H . Several other authors count labelled copies, which multiplies $X(H, G_n)$ by $|\text{Aut}(H)|$.

Therefore, by exchangeability,

$$\begin{aligned}\mathbb{E}[X(H, G_n)] &= \sum_{1 \leq i_1 < \dots < i_{|V(H)|} \leq n} \sum_{H' \in \mathcal{G}_H(\{i_1, \dots, i_{|V(H)|}\})} \prod_{(i_s, i_t) \in E(H')} \mathbb{P}(Y_{i_s i_t} \leq W(U_{i_s} U_{i_t})) \\ &= \frac{(n)_{|V(H)|}}{|\text{Aut}(H)|} t(H, W),\end{aligned}\tag{1.2}$$

where $(n)_{|V(H)|} := n(n-1) \cdots (n - |V(H)| + 1)$ and

$$t(H, W) = \int_{[0,1]^{|V(H)|}} \prod_{(a,b) \in E(H)} W(x_a, x_b) \prod_{a=1}^{|V(H)|} dx_a \tag{1.3}$$

is the *homomorphism density* of the graph H in the graphon W . The homomorphism density $t(H, W)$ can be interpreted as the probability that a W -random graph on $|V(H)|$ vertices contains the graph H , that is,

$$t(H, W) = \mathbb{P}[H \subseteq G(|V(H)|, W)].$$

One of the fundamental results in graph limit theory is that the homomorphism densities identify a graphon up to a measure-preserving transformation. The computation in (1.2) shows that

$$\hat{t}(H, G_n) := \frac{|\text{Aut}(H)|}{(n)_{|V(H)|}} X(H, G_n) \tag{1.4}$$

is an unbiased estimate of the homomorphism density $t(H, W)$. To assess the uncertainty and confidence of this estimate it is essential to understand the fluctuations (asymptotic distribution) of $\hat{t}(H, G_n)$ (equivalently that of $X(H, G_n)$). In fact, many inferential tasks in network analysis, such as estimating the clustering coefficient or testing for global structure, require understanding the joint distribution of multiple (more than 1) subgraph counts. This raises following natural questions:

- (Q1) Given a collection of r graphs $\mathcal{H} = \{H_1, H_2, \dots, H_r\}$, what is limiting joint distribution of $\mathbf{X}(\mathcal{H}, G_n) := (X(H_1, G_n), X(H_2, G_n), \dots, X(H_r, G_n))^\top$?
- (Q2) How can one construct asymptotically valid joint confidence sets for the homomorphism densities $\mathbf{t}(\mathcal{H}, W) = (t(H_1, W), t(H_2, W), \dots, t(H_r, W))^\top$ based on a single realization of the sampled graph G_n ?

Despite the growing interest in the random graphon model and the network method of moments, existing results provide only a limited understanding of these questions. In this paper we develop a framework for studying the asymptotic properties of network moments, which resolves the questions above in its full generality and closes several gaps in the existing literature. We summarize our results in the following sections.

1.1. Joint Distribution of Network Moments. The asymptotic distribution of subgraph counts in the Erdős Rényi model, where $W = W_p \equiv p$ is a constant function, has been classically studied, using various tools such as U -statistics [66, 67], method of moments [75], Stein's method [5], and martingales [42, 43] (see also [46, Chapter 6]). In particular, when $p \in (0, 1)$ is fixed and $G_n \sim G(n, W_p)$, $\mathbf{X}(\mathcal{H}, G_n)$ is known to be asymptotically jointly normal for any finite collection $\mathcal{H}$ of non-empty graphs (see [45, Section 9]). For general graphons W , the fluctuations of $X(H, G_n)$ (or that of the empirical homomorphism density $t(H, G_n)$ (see (2.2) for the definition) has received significant attention recently. This began with the work of Bickel et al. [11], where the asymptotic Gaussian distribution for subgraph counts was established, under certain sparsity assumptions. Later, using the framework of mod-Gaussian convergence, Féray, Méliot, and Nikeghbali [27] derived the asymptotic normality, moderate deviations, and local

limit theorems for the empirical homomorphism density. The joint Gaussian convergence of a finite collection of empirical homomorphism densities was established in Delmas et al. [22]. Recently, Zhang [86] derived rates of convergence to normality (Berry–Esseen type bounds), Zhang and Xia [85] obtained Edgeworth expansions, and Austern and Orbanz [4] studied connections to exchangeability, for $X(H, G_n)$ (or its related variations). Other related results include central limit theorems with rates of convergence for centered subgraph counts [47], analysis of localized subgraph counts [60], and motif counts in bipartite exchangeable networks [50].

One interesting feature that has escaped attention is that the limiting normal distribution of the subgraph counts obtained in the aforementioned works can be degenerate depending on the structure of the graphon W . For instance, in a planted bisection model [63] (a stochastic block model with two equal-sized communities and connection probabilities p and q within and between blocks, respectively), the limiting distribution of network moments such as edges and triangles are degenerate (see Case 4 in Example 3.1). This degeneracy phenomenon was noted in Féray et al. [27], and first systemically studied by Hladký et al. [34] when $H = K_R$ is the R -clique (the complete graph on R vertices), for some $R \geq 2$. This was extended to general subgraphs H by Bhattacharya et al. [8]. Here, it was shown that the usual Gaussian limit of $X(H, G_n)$ is degenerate when a certain regularity function, which encodes the homomorphism density of H incident to a given ‘vertex’ of W , is constant almost everywhere. In this case, the graphon H is said to be H -regular (see Definition 2.1) and the asymptotic distribution of $X(H, G_n)$ (with another normalization, differing by a factor $n^{\frac{1}{2}}$) can have two components: a Gaussian component and another independent (non-Gaussian) component which is a (possibly) infinite weighted sum of centered chi-squared random variables. This degeneracy phenomenon also appears in the subsequent work of Chatterjee and Huang [18] on the fluctuations of the largest eigenvalue. Very recently, Huang et al. [39] established an invariance principle for $X(H, G_n)$ that encompasses higher-order degeneracies.

In this paper we generalize the above results, which only considers the marginal distribution of a single subgraph count, to joint distributions (recall (Q1)). Specifically, we derive the limiting joint distribution of $\mathbf{X}(\mathcal{H}, G_n) := (X(H_1, G_n), X(H_2, G_n), \dots, X(H_r, G_n))^\top$ (appropriately centered and scaled), when W is irregular with respect to $H_1, \dots, H_q$ for some $1 \leq q \leq r$, and regular with respect to $H_{q+1}, H_{q+2}, \dots, H_r$. This is significantly more delicate than marginal convergence, because of the non-Gaussian dependencies between and within the irregular and regular marginals. Towards this, using the asymptotic theory of generalized U -statistics developed by Janson and Nowicki [45] and the framework of multiple stochastic integrals we show the following (see Theorem 2.1 for the formal statement):

- The limiting distribution of $((X(H_i, G_n)))_{1 \leq i \leq q}$ (the irregular marginals) is a linear stochastic integral in terms of the regularity function.
- The limiting distribution of $((X(H_i, G_n)))_{q+1 \leq i \leq r}$ (the regular marginals) is the sum of two independent components; one of which is a multivariate Gaussian and the other is a bivariate stochastic integral in terms of the 2-point conditional kernel of H_i in W .

The stochastic integrals are with respect to the same underlying Brownian motion on $[0, 1]$, which captures the dependence between the different marginals. This result goes beyond the well-known sampling convergence (law of large numbers) for subgraph densities (see [55, Corollary 10.4]) and also the first-order Gaussian fluctuations. Hence, our results can be thought of as the higher-order fluctuation theory for subgraph counts in the random graphon model. The formal statement of the results are given in Section 2. In Section 3 we illustrate the general theory in some examples.

1.2. Joint Confidence Sets. To use the results described in the previous section for statistical inference (recall (Q2)), one needs to estimate the quantiles of the limiting distribution of the

subgraph counts (which depend on the unknown graphon W). This is particularly relevant because network moments commonly appear in inferential tasks such as goodness-of-fit and two-sample problems (see [17, 31, 51, 53, 61, 68, 77, 79, 82] among several others), which requires one to approximate the quantiles of the sampling distribution of the subgraph counts. Towards this different network bootstrap and subsampling methods have been proposed (see [9, 33, 53, 54, 58, 85] and the references therein). However, most of the existing results on bootstrap consistency are restricted to the regime where the subgraph count has a non-degenerate Gaussian distribution (and some of them also require the network to be sparse). The literature is surprisingly silent in the case where the Gaussian distribution is degenerate. The recent paper [77] appears to be the only one that directly address the degeneracy issue in the context of network two-sample testing. However, their result requires the network to be sparse (in addition to other technical conditions) and, hence, does not directly apply to the dense regime.

In this paper we develop a multiplier bootstrap method for approximating the limiting joint distribution of the network moments that remains valid even if the Gaussian distribution is degenerate. On a high level, this entails replacing the graphon W in the limiting distribution with its empirical counterpart (obtained from the observed graph G_n) and introducing random Gaussian multipliers (which are independent of G_n). For the irregular marginals (where the limiting distribution is Gaussian), the estimate takes the form a linear combination of Gaussians with weights given by an empirical estimate of the regularity function. On the other hand, for the regular marginals, the estimate is a quadratic form in Gaussians in terms of an empirical estimate of the 2-point conditional kernel (see (4.5) for the formal definition). We show that this estimate, interestingly, converges to the joint distribution of the network moments, conditional on the observed network G_n , with no additional assumptions on the graphon W (Theorem 4.1). We refer to this as the *graphon multiplier bootstrap*. Details are given in Section 4.

The graphon multiplier bootstrap, however, cannot be directly used for constructing confidence sets for the homomorphism densities, because we do not know which of the subgraphs in $\mathcal{H}$ are regular with respect to W . For this we develop a test for regularity based on a consistent estimate of the variance of the limiting Gaussian distribution (Proposition 5.1). Combining this with the graphon multiplier bootstrap we construct joint confidence sets for the homomorphism densities that are asymptotically valid for any finite collection of subgraphs (Theorem 5.1). To validate the theoretical results, we also study the finite sample performance of the proposed method in simulations. Details are given in Section 5.

1.3. Testing for Global Structure. The framework for analyzing the asymptotic properties network moments discussed above, readily applies to many problems in network inference. To illustrate, here we consider the problem of detecting global structure based on small subgraphs. Different variations of this problem have appeared in the literature. For instance, Gao and Lafferty [29] considered testing whether a degree-corrected block model has any structure, that is, whether it has a single community (which corresponds to no structure) versus it has more than 1 community (see also [28] for related results). In the graphon framework, detecting global structure corresponds to testing the null hypothesis:

$$H_0 : W = p \text{ almost everywhere for some } p \in (0, 1), \quad (1.5)$$

based on a single observed network G_n from the W -random model. For this problem, Fang and Röllin [26] proposed a universally consistent test based on the densities of the edge and the 4-cycle, invoking the celebrated result of Chung, Graham, and Wilson [20] about quasi-random graphs. In this paper, using the same quasi-randomness result, we propose a simpler test statistic which also gives a universally consistent test. Our proposal relies on the observation that H_0 in (1.5) holds if and only if $f(W) := t(K_2, W)^4 - t(C_4, W) = 0$, where K_2 denotes

the edge and C_4 denotes the 4-cycle. Consequently, a test which rejects for large values of $\hat{f}(G_n) := \hat{t}(K_2, G_n)^4 - \hat{t}(C_4, G_n)$ (recall (1.4)) will be universally consistent. In Section 6 we derive the limiting distribution of $\hat{f}(G_n)$ under both the null and the alternative, using the techniques employed in Section 1.1. This allows us to obtain a test with precise asymptotic level (unlike the test in [26] which is conservative) and also understand its fluctuations under the alternative.

2. ASYMPTOTIC JOINT DISTRIBUTION OF NETWORK MOMENTS

We begin by introducing the notion of regularity, the conditional 2-point kernel, and other related concepts in Section 2.1. In Section 2.2 we define the graph join operations. The asymptotic joint distribution of the subgraph counts are given in Section 2.3.

2.1. Conditional Homomorphism Density. Recall the definition of homomorphism density for a simple graph from (1.3). This extends easily to multigraphs as follows: The homomorphism density of a fixed multigraph $F = (V(F), E(F))$ (without loops) in a graphon W is defined as:

$$t(F, W) = \int_{[0,1]^{|V(F)|}} \prod_{(a,b) \in E(F)} W(x_a, x_b) \prod_{a=1}^{|V(F)|} dx_a. \quad (2.1)$$

Note that (2.1) is a natural continuum extension of the homomorphism density of a fixed graph $F = (V(F), E(F))$ into finite (unweighted) graph $G = (V(G), E(G))$ defined as:

$$t(F, G) := \frac{|\text{hom}(F, G)|}{|V(G)|^{|V(F)|}}, \quad (2.2)$$

where $|\text{hom}(F, G)|$ denotes the number of homomorphisms of F into G . In fact, $t(F, G)$ is the proportion of maps $\phi : V(F) \rightarrow V(G)$ which define a graph homomorphism. Defining the *empirical graphon* associated with the graph G as:

$$W^G(x, y) := \mathbf{1}\{(|V(G)|x, |V(G)|y) \in E(G)\}, \quad (2.3)$$

it can be easily verified that $t(F, G) = t(F, W^G)$. (In other words, to obtain the empirical graphon W^G from the graph G , partition $[0, 1]^2$ into $|V(G)|^2$ squares of side length $1/|V(G)|$, and let $W^G(x, y) = 1$ in the (i, j) -th square if $(i, j) \in E(G)$, and 0 otherwise.)

We now introduce the notion of conditional homomorphism densities and H -regularity of graphons.

Definition 2.1. (1-point conditional homomorphism density and H -regularity) Fix $a \in V(H)$ and $x \in [0, 1]$. Then *1-point conditional homomorphism density function* of H in a graphon W given the vertex a is defined as:

$$t_a(x, H, W) := \mathbb{E} \left[\prod_{(a,b) \in E(H)} W(U_a, U_b) \mid U_a = x \right].$$

In other words, $t_a(x, H, W)$ is the homomorphism density of H in the graphon W when the vertex $a \in V(H)$ is marked with the value $x \in [0, 1]$. A graphon W is said to be *H -regular* if

$$\bar{t}(x, H, W) := \frac{1}{|V(H)|} \sum_{a=1}^{|V(H)|} t_a(x, H, W) = t(H, W), \quad (2.4)$$

for almost every $x \in [0, 1]$. We say W is *H -irregular* if it is not H -regular.

To illustrate the notion of regularity, we consider the following 3 examples: (1) $H = K_2$ is the edge, (2) $H = K_3$ is the triangle, and (3) $H = K_{1,2}$ is the 2-star. These 3 choices of H will be the running examples throughout the paper.

- $H = K_2$ is the edge: In this case, for any $a \in \{1, 2\}$, by symmetry,

$$t_a(x, K_2, W) = \mathbb{E}[W(U_1, U_2) | U_a = x] = \int_0^1 W(x, y) dy := d_W(x),$$

is the *degree function* of W . Hence, a graphon W is K_2 -regular if and only the degree function d_W is constant almost everywhere, that is, W is *degree-regular*.

- $H = K_3$ is the triangle: Again, by symmetry, for all $1 \leq a \leq 3$,

$$t_a(x, K_3, W) = \int_0^1 \int_0^1 W(x, y) W(y, z) W(x, z) dy dz,$$

which is the homomorphism density of triangles incident at the point $x \in [0, 1]$.

- $H = K_{1,2}$ is 2-star: Suppose the vertices of $K_{1,2}$ are labeled $\{1, 2, 3\}$ with the central vertex labeled 1. Then we have the following:

$$- \text{ For } a = 1, t_1(x, K_{1,2}, W) = \int_0^1 \int_0^1 W(x, y) W(x, z) dy dz = d_W(x)^2.$$

$$- \text{ For } a \in \{2, 3\}, t_a(x, K_{1,2}, W) = \int_0^1 \int_0^1 W(x, y) W(y, z) dy dz.$$

Hence,

$$\bar{t}(x, K_{1,2}, W) = \frac{1}{3} \left(d_W(x)^2 + 2 \int_0^1 \int_0^1 W(x, y) W(y, z) dy dz \right).$$

Next, we define the 2-point conditional homomorphism density and the kernel derived from it. This kernel will arise in the non-Gaussian component of the limiting distribution of $X(H, G_n)$ in the regular regime.

Definition 2.2. (2-point conditional homomorphism density) Fix $a \neq b \in V(H)$ and $x \in [0, 1]$. Then the 2-point conditional homomorphism density function of H in a graphon W given the vertices a and b is defined as:

$$t_{a,b}(x, y, H, W) := \mathbb{E} \left[\prod_{(a,b) \in E(H)} W(U_a, U_b) \mid U_a = x, U_b = y \right].$$

Further, the 2-point conditional kernel of H is defined as:

$$W_H(x, y) = \frac{1}{2|\text{Aut}(H)|} \sum_{1 \leq a \neq b \leq |V(H)|} t_{a,b}(x, y, H, W). \quad (2.5)$$

For illustration, as before, we consider the following examples:

- $H = K_2$ is the edge: In this case, $t_{1,2}(x, y, K_2, G_n) = t_{2,1}(x, y, K_2, G_n) = W(x, y)$. Hence, $W_H(x, y) = \frac{W(x, y)}{2}$, that is, the 2-point conditional kernel is the scaled graphon W .
- $H = K_3$ is the triangle: By symmetry, $t_{a,b}(x, y, K_3, W) = t_{1,2}(x, y, K_3, W)$ for all $1 \leq a \neq b \leq 3$. Hence, the 2-point conditional kernel is given by,

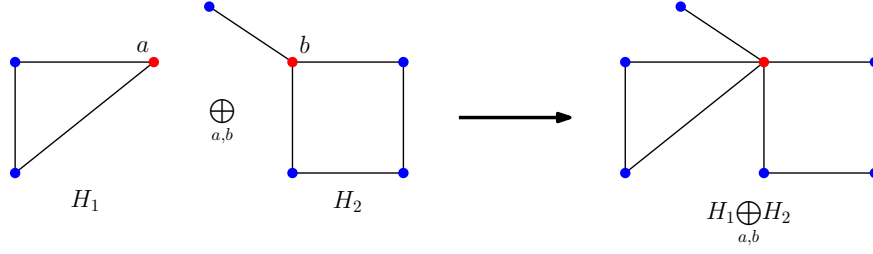
$$W_{K_3}(x, y) = \frac{1}{2} t_{1,2}(x, y, K_3, W) = \frac{1}{2} W(x, y) \int_0^1 W(x, z) W(z, y) dz,$$

since $|\text{Aut}(K_3)| = 3! = 6$.

- $H = K_{1,2}$ is 2-star: Suppose the vertices of $K_{1,2}$ are labeled $\{1, 2, 3\}$ with the central vertex labeled 1. Then we have the following:

$$- \text{ For } a = 1 \text{ and } b \in \{2, 3\},$$

$$t_{a,b}(x, y, K_{1,2}, W) = W(x, y) d_W(x) \text{ and } t_{b,a}(x, y, K_{1,2}, W) = W(x, y) d_W(y).$$

FIGURE 1. The (a, b) -vertex join of the graphs H_1 and H_2 .

– For the remaining vertex pairs $(2, 3)$ and $(3, 2)$,

$$t_{2,3}(x, y, K_{1,2}, W) = t_{3,2}(x, y, K_{1,2}, W) = \int_0^1 W(x, z)W(z, y)dz.$$

Hence, the 2-point conditional kernel is given by,

$$W_{K_{1,2}}(x, y) = \frac{1}{2} \left[\int_0^1 W(x, z)W(z, y)dz + W(x, y)(d_W(x) + d_W(y)) \right].$$

Remark 2.1. Note that a graphon W is H -regular (see Definition 2.1) if and only if the 2-point conditional kernel W_H is degree regular. This is because, for all $x \in [0, 1]$,

$$\int_0^1 W_H(x, y)dy = \frac{|V(H)| - 1}{2|\text{Aut}(H)|} \sum_{a=1}^{|V(H)|} t_a(x, H, W), \quad (2.6)$$

and the RHS of (2.6) is a constant if and only if W is H -regular. In fact, if W is H -regular, then $\frac{1}{|V(H)|} \sum_{a=1}^{|V(H)|} t_a(x, H, W) = t(H, W)$ almost everywhere. Hence, the degree function of W_H becomes

$$\int_0^1 W_H(x, y)dy = \frac{|V(H)|(|V(H)| - 1)}{2|\text{Aut}(H)|} \cdot t(H, W) := d_{W_H}, \quad (2.7)$$

for almost every $x \in [0, 1]$.

Note that $|W_H| \leq \frac{|V(H)|(|V(H)| - 1)}{2|\text{Aut}(H)|} := K$. Hence, W_H defines an operator $T_{W_H} : L^2[0, 1] \rightarrow L^2[0, K]$ as follows:

$$(T_{W_H}f)(x) = \int_0^1 W_H(x, y)f(y)dy, \quad (2.8)$$

for each $f \in L^2[0, 1]$. T_{W_H} is a symmetric Hilbert–Schmidt operator; thus it is compact and has a discrete spectrum, that is, it has a countable multiset of non-zero real eigenvalues, which we denote by $\text{Spec}(W_H)$, such that

$$\sum_{\lambda \in \text{Spec}(W_H)} \lambda^2 = \iint W_H(x, y)^2 dx dy < \infty.$$

Note that if W is H -regular, then d_{W_H} is an eigenvalue of the operator T_{W_H} (recall (2.8)) and $\phi \equiv 1$ is a corresponding eigenvector. In this case, we will use $\text{Spec}^-(W_H)$ to denote the collection $\text{Spec}(W_H)$ with the multiplicity of the eigenvalue d_{W_H} decreased by 1.

2.2. Graph Join Operations. The variance of the subgraph count $X(H, G_n)$ involves different graphs obtained by joining 2 isomorphic copies of H . To describe the asymptotic variance succinctly it is convenient to define some basic graph join operations (as in [8]). To this end, suppose $H = (V(H), E(H))$ is a graph with vertex set $V(H) = \{1, 2, \dots, |V(H)|\}$. Denote by $E^+(H)$ the ordered pairs of edges in H , that is, $E^+(H) = \{(a, b) : 1 \leq a \neq b \leq r, (a, b) \text{ or } (b, a) \in E(H)\}$.

Definition 2.3. Suppose $H_1 = (V(H_1), E(H_1))$ and $H_2 = (V(H_2), E(H_2))$ be two graphs with vertex sets $V(H_1) = \{1, 2, \dots, |V(H_1)|\}$ and $V(H_2) = \{1, 2, \dots, |V(H_2)|\}$ and edge sets $E(H_1)$ and $E(H_2)$, respectively.

- *Vertex Join:* For $a \in V(H_1)$ and $b \in V(H_2)$, the (a, b) -vertex join of H_1 and H_2 , denoted by

$$H_1 \bigoplus_{a,b} H_2,$$

is the graph obtained by identifying the a -th vertex of H_1 with the b -th vertex of H_2 (see Figure 1).

- *Weak Edge Join:* For $(a, b) \in E^+(H_1)$ and $(c, d) \in E^+(H_2)$, with $1 \leq a \neq b \leq r$ and $1 \leq c \neq d \leq r$, the $(a, b), (c, d)$ -weak edge join of H_1 and H_2 , denote by

$$H_1 \bigoplus_{(a,b),(c,d)} H_2,$$

is the graph obtained identifying the vertices a and c and the vertices b and d and keeping a single edge between the two identified vertices (see Figure 2).

- *Strong Edge Join:* For $(a, b) \in E^+(H_1)$ and $(c, d) \in E^+(H_2)$, with $1 \leq a \neq b \leq r$ and $1 \leq c \neq d \leq r$, the $(a, b), (c, d)$ -strong edge join of H_1 and H_2 ,

$$H_1 \bigoplus_{(a,b),(c,d)} H_2,$$

is the multi-graph obtained identifying the vertices a and c and the vertices b and d and keeping both the edges between the two identified vertices (see Figure 2).

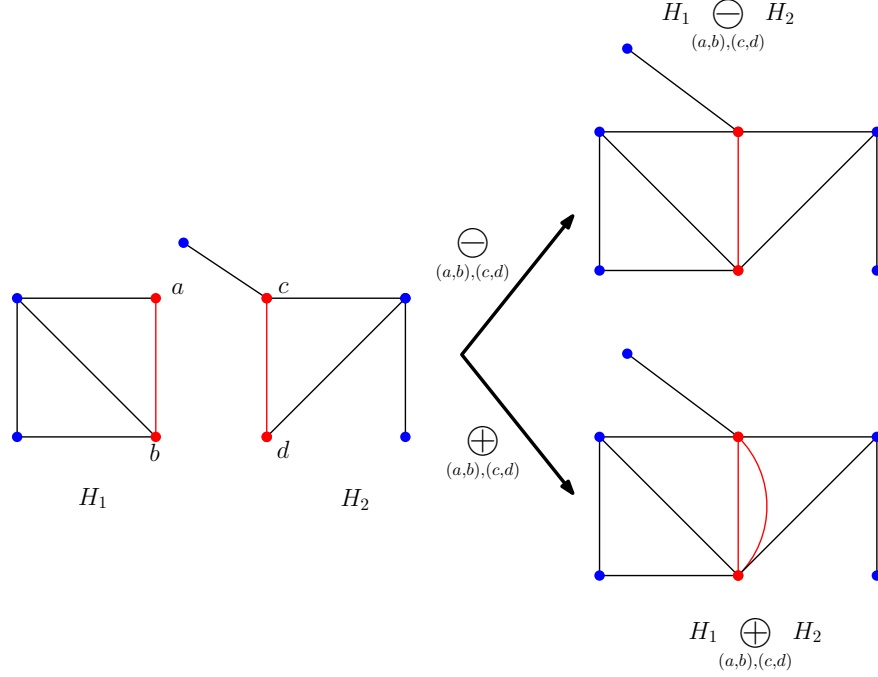
2.3. Joint Distribution of Subgraph Counts. Suppose $\mathcal{H} = \{H_1, \dots, H_r\}$ is a collection of finite simple graphs, where $H_i = (V(H_i), E(H_i))$ with vertices labeled $V(H_i) = \{1, 2, \dots, |V(H_i)|\}$ and $|V(H_i)| \geq 2$, for $1 \leq i \leq r$. To begin with, for any finite simple graph $H = (V(H), E(H))$, with $|V(H)| \geq 2$, define

$$Z(H, G_n) = \begin{cases} \frac{X(H, G_n) - \frac{(n)^{|V(H)|} t(H, W)}{|Aut(H)|}}{n^{|V(H)| - \frac{1}{2}}} & \text{if } W \text{ is } H\text{-irregular,} \\ \frac{X(H, G_n) - \frac{(n)^{|V(H)|} t(H, W)}{|Aut(H)|}}{n^{|V(H)| - 1}} & \text{if } W \text{ is } H\text{-regular.} \end{cases} \quad (2.9)$$

Our goal is to derive the limiting distribution of

$$\mathbf{Z}(\mathcal{H}, G_n) = (Z(H_1, G_n), Z(H_2, G_n), \dots, Z(H_r, G_n))^T. \quad (2.10)$$

For this we need to define the following covariance matrix:

FIGURE 2. The weak and strong edge joins of the graphs H_1 and H_2 .

Definition 2.4. Given a graphon W and finite collection of graphs $\{F_1, F_2, \dots, F_p\}$, such that W is regular with respect to $F_1, F_2, \dots, F_p$. Then define a $p \times p$ matrix $\Sigma := (\sigma_{ij})_{1 \leq i, j \leq p}$ as follows:

$$\sigma_{ij} = \frac{1}{2|\text{Aut}(F_i)||\text{Aut}(F_j)|} \sum_{\substack{(a,b) \in E^+(F_i) \\ (c,d) \in E^+(F_j)}} \left[t \left(F_i \ominus_{(a,b),(c,d)} F_j, W \right) - t \left(F_i \oplus_{(a,b),(c,d)} F_j, W \right) \right], \quad (2.11)$$

for all $1 \leq i, j \leq p$.

We are now ready to state our result about the limiting distribution of subgraph counts. To this end, denote by $\{B_t : t \in [0, 1]\}$ the standard Brownian motion on $[0, 1]$ and recall the framework of multiple Wiener-Itô stochastic integrals from Section G.

Theorem 2.1. Fix a graphon W and a finite collection of non-empty graphs $\mathcal{H} = \{H_1, H_2, \dots, H_r\}$, such that W is irregular with respect to $H_1, \dots, H_q$ for some $1 \leq q \leq r$ and regular with respect to $H_{q+1}, H_{q+2}, \dots, H_r$. Then

$$\mathbf{Z}(\mathcal{H}, G_n) \xrightarrow{D} \mathbf{Z}(\mathcal{H}, W) := (Z(H_1, W), Z(H_2, W), \dots, Z(H_r, W))^\top, \quad (2.12)$$

such that

- for $1 \leq i \leq q$,

$$Z(H_i, W) := \int_0^1 \left\{ \frac{1}{|\text{Aut}(H_i)|} \sum_{a=1}^{|V(H_i)|} t_a(x, H_i, W) - \frac{|V(H_i)|}{|\text{Aut}(H_i)|} t(H_i, W) \right\} dB_x,$$

- for $q + 1 \leq i \leq r$,

$$Z(H_i, W) := G_i + \int_0^1 \int_0^1 \left\{ W_{H_i}(x, y) - \frac{|V(H_i)|(|V(H_i) - 1|)}{2|\text{Aut}(H_i)|} t(H_i, W) \right\} dB_x dB_y,$$

where $\mathbf{G} = (G_{q+1}, \dots, G_r) \sim N_{r-q}(\mathbf{0}, \Sigma)$, with $\Sigma = ((\sigma_{ij}))_{q+1 \leq i, j \leq r}$ as in (2.11), is independent of $\{B_t\}_{t \in [0,1]}$.

The proof of Theorem 2.1 uses the asymptotic theory of generalized U -statistics developed in Janson and Nowicki [45]. This allows us to decompose $\mathbf{X}(\mathcal{H}, G_n)$ over sums of increasing complexity using a projection method (see also [44, Chapter 11]). The terms in the expansion are indexed by the vertices and edges subgraphs of the complete graph of increasing sizes, and the asymptotic behavior of $\mathbf{X}(\mathcal{H}, G_n)$ is determined by the joint distribution of non-zero terms indexed by the smallest size graphs. Then the machinery of multiple stochastic integral provides a convenient way to express the dependence among the irregular and regular marginals. The proof is given in Section A.

Theorem 2.1 recovers as special cases a number of existing results. For instance, when $\mathcal{H} = \{H\}$ is a singleton, we get the marginal distribution of $Z(H, G_n)$, which was proved for cliques in [34] and for general subgraphs in [8]. In this case the limiting distribution can be alternately expressed as in the following corollary, in terms of the graph join operations and the eigenvalues of the kernel W_H (recall the discussion following Remark 2.1). We show how to derive Corollary 2.1 from Theorem 2.1 in Section C.

Corollary 2.1 ([8, Theorem 2.9]). *Fix a graphon W and a non-empty graph $H = (V(H), E(H))$. Then as $n \rightarrow \infty$, the following hold:*

- If W is H -irregular,

$$Z(H, G_n) \xrightarrow{D} N(0, \tau_{H,W}^2), \quad (2.13)$$

where

$$\tau_{H,W}^2 = \frac{1}{|\text{Aut}(H)|^2} \left[\sum_{1 \leq a, b \leq |V(H)|} t \left(H \oplus_{a,b} H, W \right) - |V(H)|^2 t(H, W)^2 \right]. \quad (2.14)$$

- If W is H -regular, that is, $\tau_{H,W}^2 = 0$,

$$Z(H, G_n) \xrightarrow{D} \sigma_{H,W} \cdot Z + \sum_{\lambda \in \text{Spec}^-(W_H)} \lambda (Z_\lambda^2 - 1) \quad (2.15)$$

where $Z, \{Z_\lambda : \lambda \in \text{Spec}^-(W_H)\}$ are independent $N(0, 1)$,

$$\sigma_{H,W}^2 := \frac{1}{2|\text{Aut}(H)|^2} \sum_{(a,b),(c,d) \in E^+(H)} \left[t \left(H \ominus_{(a,b),(c,d)} H, W \right) - t \left(H \oplus_{(a,b),(c,d)} H, W \right) \right],$$

and $\text{Spec}^-(W_H)$ is the multiset $\text{Spec}(W_H)$ with multiplicity of the eigenvalue d_{W_H} decreased by 1.

Remark 2.2. An interesting question that arises from Corollary 2.1, is whether the distribution in (2.15) is always non-degenerate? This is known to be true when H is the clique [34] and if H is the 2-star or the 4-cycle [8]. However, there are non-trivial cases where the limit in (2.15) is degenerate (see [8, Example 4.6]). Instances where one (but not both) of the two components of the distribution in (2.15) is degenerate also has interesting combinatorial properties (see [8, Section 4]). Additional degeneracies appear in the multivariate case. For instance, the matrix Σ

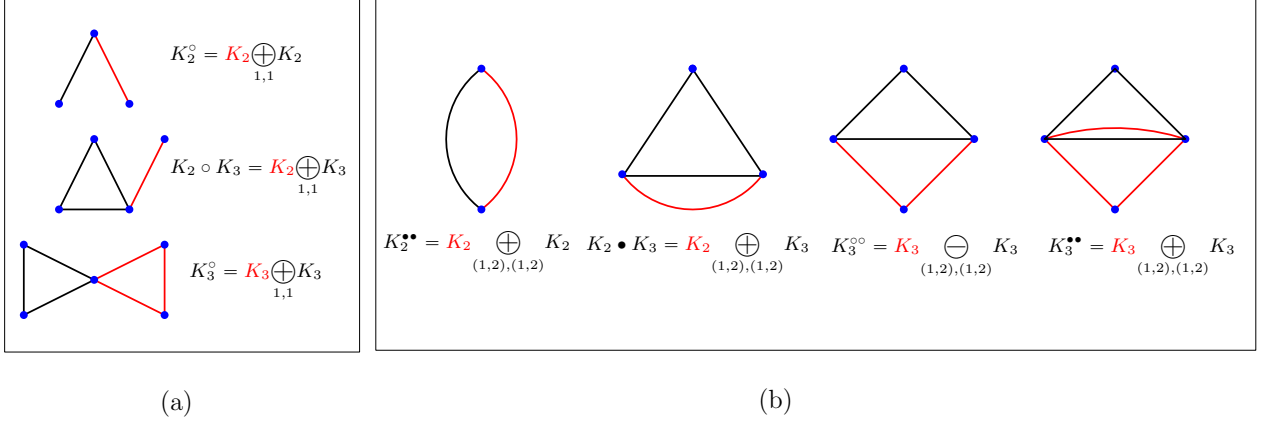


FIGURE 3. Graphs obtained from (a) vertex join operations and (b) edge join operations, between a copy of K_2 and a copy K_3 .

in Theorem 2.1 can be singular. This is the case, for example, in the Erdős-Rényi model where the matrix Σ has rank 1 for any finite collections of graphs (see Example 3.2).

Another case which has appeared in prior work is when all the graph in $\mathcal{H}$ are irregular with respect to W (see [27] for the univariate case and [22] for the multivariate case). In this case, since a linear stochastic integral has a Gaussian distribution, the limiting distribution of $\mathbf{Z}(\mathcal{H}, G_n)$ is multivariate Gaussian (see Theorem 1.5 in [44]). The covariance matrix of this Gaussian distribution can be expressed in terms of the graph join operations as follows:

Corollary 2.2 ([22, Corollary 7.6]). *Fix a graphon W and a finite collection of non-empty graphs $\mathcal{H} = \{H_1, \dots, H_r\}$, such that W is H_i -irregular for all $1 \leq i \leq r$. Then*

$$\mathbf{Z}(\mathcal{H}, G_n) \xrightarrow{D} N_r(\mathbf{0}, \Gamma),$$

where $\Gamma = ((\tau_{ij}))_{1 \leq i, j \leq r}$, with

$$\tau_{ij} = \frac{1}{|\text{Aut}(H_i)| |\text{Aut}(H_j)|} \left[\sum_{a=1}^{|V(H_i)|} \sum_{b=1}^{|V(H_j)|} t \left(H_i \oplus_{a,b} H_j, W \right) - |V(H_i)| |V(H_j)| t(H_i, W) t(H_j, W) \right].$$

(Note that $\tau_{ii} = \tau_{H_i, W}^2$, for $\tau_{H_i, W}^2$ as defined in (2.14) with H replaced by H_i .)

3. EXAMPLES

In this section we compute the limiting distribution of $\mathbf{Z}(\mathcal{H}, G_n)$ in a few examples. We begin with the joint distribution of the counts of edges and triangles.

Example 3.1. (Edges and triangles) Fix a graphon W and suppose $\mathcal{H} = \{K_2, K_3\}$ be the edge and the triangle. There are 4-cases depending on whether or not W is K_2 or K_3 -regular.

Case 1: W is irregular with respect to K_2 and K_3 : In this case, Corollary 2.2 applies. To this end, as shown in Figure 3 (a), denote by K_2° , K_3° , and $K_2 \circ K_3$ the graphs obtained by the vertex joins of 2 copies of K_2 , 2 copies of K_3 , and one copy K_2 and one copy of K_3 , respectively. Then by Corollary 2.2,

$$\begin{pmatrix} Z(K_2, G_n) \\ Z(K_3, G_n) \end{pmatrix} \xrightarrow{D} N_2 \left(\mathbf{0}, \begin{pmatrix} \tau_{11} & \tau_{12} \\ \tau_{21} & \tau_{22} \end{pmatrix} \right),$$

where $\tau_{11} := t(K_2^\circ, W) - t(K_2, W)^2$, $\tau_{22} := \frac{1}{4}[t(K_3^\circ, W) - t(K_3, W)^2]$, and

$$\tau_{12} = \tau_{21} := \frac{1}{2}[t(K_2 \circ K_3, W) - t(K_2, W)t(K_3, W)].$$

For a specific example of a graphon which is irregular with respect to K_2 and K_3 , consider

$$\tilde{W}_1(x, y) := \frac{1}{2}(x + y), \quad (3.1)$$

for $x, y \in [0, 1]$. In this case, $t_a(x, K_2, \tilde{W}_1) = d_{\tilde{W}_1}(x) = \frac{1}{2}(x + \frac{1}{2})$, for $a \in \{1, 2\}$, and $t_b(x, K_3, \tilde{W}_1) = \frac{1}{8}(x^2 + \frac{7x}{6} + \frac{1}{3})$, for $b \in \{1, 2, 3\}$, are both non-constant functions, hence, $\tilde{W}_1$ is K_2 and K_3 -irregular.

Case 2: W is regular with respect to K_2 and irregular with respect to K_3 : In this case, Theorem 2.1 shows that,

$$\begin{pmatrix} Z(K_2, G_n) \\ Z(K_3, G_n) \end{pmatrix} \xrightarrow{D} \begin{pmatrix} G + \frac{1}{2} \int_0^1 \int_0^1 (W(x, y) - t(K_2, W)) dB_x dB_y \\ \frac{1}{2} \int_0^1 \left(\int_0^1 \int_0^1 W(x, y) W(y, z) W(z, x) dy dz - t(K_3, W) \right) dB_x \end{pmatrix},$$

where $G \sim N(0, \sigma^2)$ is independent of the Brownian motion $\{B_t\}_{t \in [0, 1]}$ and

$$\sigma^2 = \frac{1}{2} \{t(K_2, W) - t(K_2^{\bullet\bullet}, W)\}.$$

Here, $K_2^{\bullet\bullet}$ is the graph obtained by the strong edge join of 2 copies of K_2 , as shown in Figure 3(b). For a concrete example of a graphon which is K_2 -regular and K_3 -irregular, consider the graphon $\tilde{W}_2$ shown in Figure 4(a). This can be expressed more formally as:

$$\tilde{W}_2(x, y) = \begin{cases} 1 & \text{if } (x, y) \in [0, \frac{1}{3}] \times [\frac{2}{3}, 1] \cup [\frac{2}{3}, 1] \times [0, \frac{1}{3}], \\ 1 & \text{if } (x, y) \in [\frac{1}{3}, \frac{2}{3}] \times [\frac{1}{3}, \frac{2}{3}], \\ 0 & \text{otherwise.} \end{cases} \quad (3.2)$$

The ‘graph’ representation of this graphon is shown in Figure 4(b), which corresponds to a clique and a disjoint complete bipartite graph of equal block sizes. In this case, the degree function $d_{\tilde{W}_2}(x) = \frac{1}{3}$, for all $x \in [0, 1]$, hence, $\tilde{W}_2$ is K_2 -regular. Further, for $b \in \{1, 2, 3\}$,

$$t_b(x, K_3, \tilde{W}_2) = \begin{cases} 0 & \text{if } (x, y) \in [0, \frac{1}{3}] \cup [\frac{2}{3}, 1], \\ \frac{1}{9} & \text{if } (x, y) \in [\frac{1}{3}, \frac{2}{3}], \end{cases}$$

which means $\tilde{W}_2$ is K_3 -irregular.

Case 3: W is irregular with respect to K_2 and regular with respect to K_3 : In this case, from Theorem 2.1 we have,

$$\begin{pmatrix} Z(K_2, G_n) \\ Z(K_3, G_n) \end{pmatrix} \xrightarrow{D} \begin{pmatrix} \int_0^1 \left(\int_0^1 W(x, y) dy - t(K_2, W) \right) dB_x \\ G + \frac{1}{2} \int_0^1 \int_0^1 \left(\int_0^1 W(x, y) W(y, z) W(z, x) dz - t(K_3, W) \right) dB_x dB_y \end{pmatrix},$$

where $G \sim N(0, \sigma^2)$ is independent of the Brownian motion $\{B_t\}_{t \in [0, 1]}$ and

$$\sigma^2 = \frac{1}{2} \{t(K_3^{\circ\circ}, W) - t(K_3^{\bullet\bullet}, W)\}.$$

Here, $K_3^{\circ\circ}$ and $K_3^{\bullet\bullet}$ are the graphs obtained by the weak and strong edge joins of 2 copies of K_3 , as shown in Figure 3(b), respectively. For an example of a graphon which is K_2 -irregular and K_3 -regular, consider the graphon $\tilde{W}_3$ in Figure 5(a). This is a 6×6 block graphon taking values 1, $\frac{1}{2}$, and 0 in the gray, green, and white blocks, respectively. The ‘graph’ representation of this graphon is shown in Figure 5(b), which corresponds

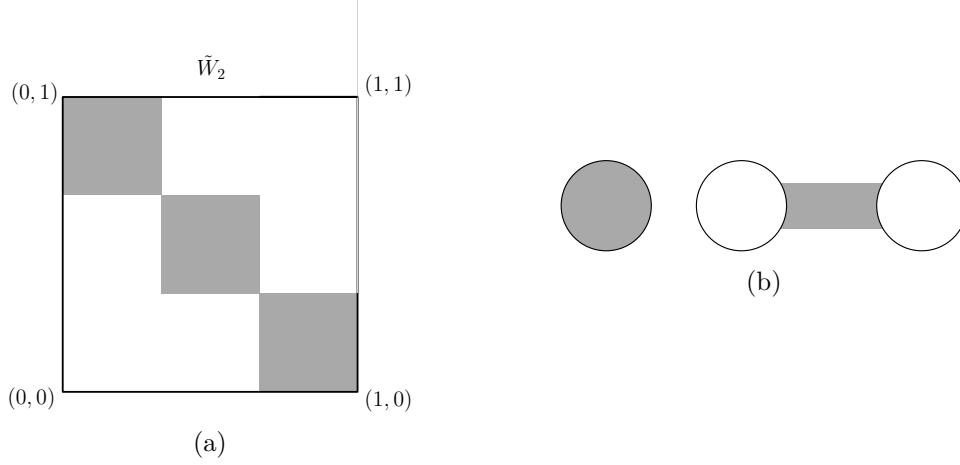


FIGURE 4. (a) A K_2 -regular and K_3 -irregular graphon $\tilde{W}_2$ and (b) its ‘graph’ representation.

to 2 disjoint complete tri-partite graphs with equal block sizes and a random bipartite graph with edge probability $\frac{1}{2}$ between 2 blocks of the tri-partite graphs. The bipartite connections change the degrees of the corresponding vertices, but do not change their 1-point triangle densities, hence, $\tilde{W}_3$ is K_2 -irregular but K_3 -regular.

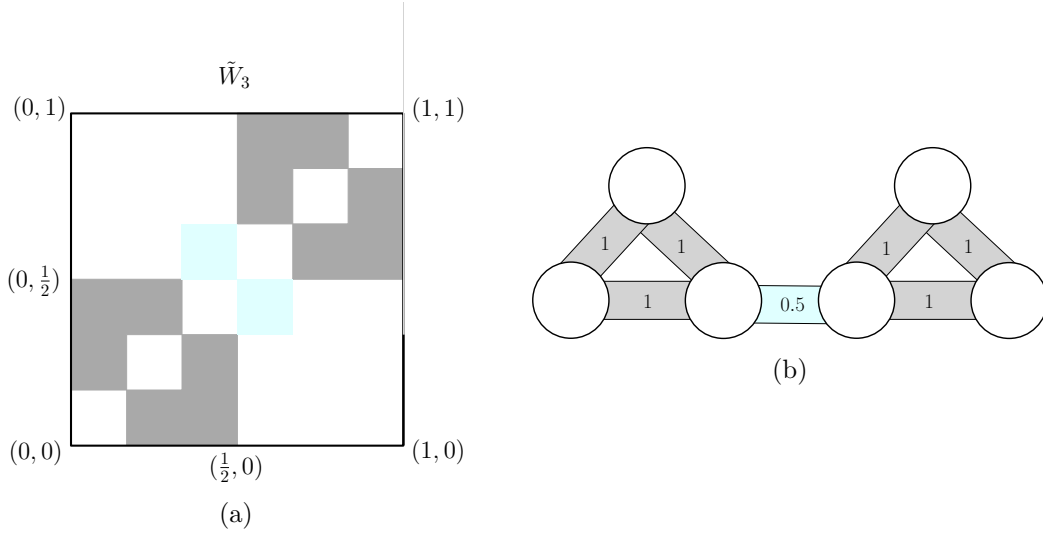


FIGURE 5. (a) A K_2 -irregular and K_3 -regular graphon $\tilde{W}_3$ and (b) its ‘graph’ representation.

Case 4: W is regular with respect to K_2 and K_3 : Once again, an application of Theorem 2.1 gives,

$$\begin{pmatrix} Z(K_2, G_n) \\ Z(K_3, G_n) \end{pmatrix} \xrightarrow{D} \begin{pmatrix} G_1 + \frac{1}{2} \int_0^1 \int_0^1 (W(x, y) - t(K_2, W)) dB_x dB_y \\ G_2 + \frac{1}{2} \int_0^1 \int_0^1 \left(\int_0^1 W(x, y) W(y, z) W(z, x) dz - t(K_3, W) \right) dB_x dB_y \end{pmatrix}.$$

Here, $\{B_t\}_{t \in [0,1]}$ is the standard Brownian motion and independently

$$\begin{pmatrix} G_1 \\ G_2 \end{pmatrix} \sim N_2 \left(\mathbf{0}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \right),$$

where $\sigma_{11} := \frac{1}{2} \{t(K_2, W) - t(K_2^{\bullet\bullet}, W)\}$, $\sigma_{22} := \frac{1}{2} \{t(K_3^{\circ\circ}, W) - t(K_3^{\bullet\bullet}, W)\}$, and

$$\sigma_{12} = \sigma_{21} := \frac{1}{2} \{t(K_2 \circ K_3) - t(K_2 \bullet K_3)\},$$

with $K_2^{\bullet\bullet}$, $K_3^{\circ\circ}$, $K_3^{\bullet\bullet}$, $K_2 \circ K_3$, and $K_2 \bullet K_3$ as shown in Figure 3(b). A simple example of a graphon which is K_2 and K_3 regular is the constant graphon $W_p \equiv p$, which, incidentally, is H -regular, for all finite graphs H . More generally, consider the R -block graphon, for some $R \geq 1$, with equal block sizes, taking values $a \in [0, 1]$ in the diagonal blocks and $b \in [0, 1]$ in the off-diagonal blocks. This graphon is also K_2 and K_3 regular.

Next, we consider the case when $W_p \equiv p$ is the constant function $p \in (0, 1)$, that is, $G_n \sim G(n, p)$ is the Erdős-Rényi random graph. In this case, it is well-known that the joint asymptotic distribution of the subgraph counts is a multivariate Gaussian (see [45, Section 9]). In the following example we show how to obtain this classical result from our general theorem.

Example 3.2. (Erdős-Rényi random graph) Suppose $W = W_p \equiv p$, that is, $G_n \sim G(n, p)$ is an Erdős-Rényi random graph with edge probability p . In this case, for any collection of finite subgraphs $\mathcal{H} = \{H_1, H_2, \dots, H_r\}$, the limiting joint distribution of $Z(\mathcal{H}, G_n)$ is known to be a multivariate Gaussian. Moreover, the covariance matrix of the Gaussian has rank 1 [45, Section 9]. Here, we show how to derive this result from Theorem 2.1. Note that W_p is regular with respect to H_i , for all $1 \leq i \leq r$ (recall (2.4)). Also,

$$W_{H_i}(x, y) = \frac{|V(H_i)|(|V(H_i)| - 1)}{2|\text{Aut}(H_i)|} p^{|E(H_i)|} = \frac{|V(H_i)|(|V(H_i)| - 1)}{2|\text{Aut}(H_i)|} t(H_i, W),$$

for $1 \leq i \leq r$. Hence, the bivariate stochastic integral in Theorem 2.1 vanishes, and the limiting distribution is given by,

$$\mathbf{Z}(\mathcal{H}, G_n) \xrightarrow{D} N_r(\mathbf{0}, \Sigma), \quad (3.3)$$

where $\Sigma = (\sigma_{ij})_{1 \leq i, j \leq r}$ with

$$\sigma_{ij} = \frac{2|E(H_i)||E(H_j)|}{|\text{Aut}(H_i)||\text{Aut}(H_j)|} p^{|E(H_i)| + |E(H_j)| - 1} (1 - p). \quad (3.4)$$

Now, for every $2 \leq i \leq r$ and $1 \leq j \leq r$ observe that

$$\frac{\sigma_{1j}}{\sigma_{ij}} = \frac{|E(H_1)|}{|E(H_i)|} \frac{|\text{Aut}(H_i)|}{|\text{Aut}(H_1)|} p^{|E(H_1)| - |E(H_i)|}.$$

Hence, the i -th column of Σ is a multiple of the first column of Σ , for $2 \leq i \leq r$, that is, the matrix Σ has rank 1.

In the next example we discuss the global clustering coefficient, which can be expressed in terms of the counts of 2-stars and triangles.

Example 3.3. (Global clustering coefficient/transitivity) The *global clustering coefficient* of a graph G is defined as (see [57]):

$$\eta(G) := \frac{3 \times \text{number of triangles in } G}{\text{the number of 2-stars in } G} = \frac{3X(K_3, G)}{X(K_{1,2}, G)} = \frac{\hat{t}(K_3, G)}{\hat{t}(K_{1,2}, G)}, \quad (3.5)$$

where $\hat{t}(\cdot, G)$ is defined in (1.4). This is a measure of clustering in the graph G and is also known as the *transitivity ratio* (see [80, Page 243]). Extending (3.5), one can define the *global clustering coefficient* of a graphon W as follows:

$$\eta(W) := \frac{t(K_3, W)}{t(K_{1,2}, W)} = \mathbb{P}(\text{the nodes } (1, 2, 3) \text{ are connected} \mid (1, 2) \text{ and } (1, 3) \text{ are connected}),$$

assuming $t(K_{1,2}, W) > 0$. Clearly, when $G_n \sim G(n, W)$, then $\eta(G_n)$ is a consistent estimate of $\eta(W)$. Using the asymptotic joint distribution of $(X(K_{1,2}, G_n), X(K_3, G_n))$ from Theorem 2.1 and the delta method, we can derive the asymptotic distribution of $\eta(G_n)$. The limit depends on whether or not the graphon W is $K_{1,2}$ and K_3 regular, hence, 4 cases can arise, similar to Example 3.1. We can also quantify the uncertainty of $\eta(G_n)$ in estimating $\eta(W)$, using the results on joint confidence sets in Section 4.

4. GRAPHON MULTIPLIER BOOTSTRAP

Note that the asymptotic distribution of the subgraph counts obtained in Theorem 2.1 depends on the graphon W . Hence, to use this result for statistical inference of the homomorphism densities, one needs to estimate quantiles of the asymptotic distribution. When the limit is Gaussian, that is, W is H -irregular, this entails estimating the asymptotic variance consistently. However, if the limit is non-Gaussian, which is the case when W is H -regular, this is more delicate. This becomes even more challenging in the multivariate regime, when there is a combination of irregular and regular components.

In this section, we introduce the *graphon multiplier bootstrap*, a method for estimating the quantiles of the limiting distribution $\mathbf{Z}(\mathcal{H}, G_n)$ (recall (2.9)), based on the observed network G_n itself and additional external randomness. To begin with, denote by $\mathbf{A}_{G_n} = ((w_{st}))_{s,t=1}^n$ the adjacency matrix of G_n and, as before, let W^{G_n} be the empirical graphon corresponding to G_n (recall (2.3)). Then the empirical homomorphism density of a graph $H = (V(H), E(H))$ in G_n can be expressed as (recall (2.1)):

$$t(H, W^{G_n}) = \frac{1}{n^{|V(H)|}} \sum_{\mathbf{s} \in [n]^{|V(H)|}} \prod_{(i,j) \in E(H)} w_{s_i s_j}. \quad (4.1)$$

Moreover, the number of copies of H in the observed in G_n as defined in (1.1) can be alternatively expressed as:

$$X(H, G_n) := \frac{1}{|\text{Aut}(H)|} \sum_{\mathbf{s} \in ([n]^{|V(H)|})} \prod_{(i,j) \in E(H)} w_{s_i s_j}, \quad (4.2)$$

where $([n]^{|V(H)|})$ is the set of all $|V(H)|$ -tuples $\mathbf{s} = (s_1, \dots, s_{|V(H)|}) \in [n]^{|V(H)|}$ with distinct indices.³ Note that the cardinality of $(n)^{|V(H)|}$ is $\frac{n!}{(n-|V(H)|)!} = (n)^{|V(H)|}$. To obtain the bootstrap estimate of the asymptotic distribution of $X(H, G_n)$ we need to define the empirical counterparts of the 1-point and 2-point conditional homomorphism densities (recall (2.1)). (Hereafter, for simplicity, we will assume H has no isolated vertex.)

Definition 4.1. (Empirical 1-point subgraph density)

Fix $a \in V(H)$ and $v \in V(G_n)$. Denote by $X_a(v, H, G_n)$ the number of injective homomorphism $\phi : V(H) \rightarrow V(G_n)$ such that $\phi(a) = v$. More formally,

$$X_a(v, H, G_n) = \sum_{\mathbf{s}_{\{a\}^c}} \prod_{y \in N_H(a)} w_{v s_y}(G_n) \prod_{(x,y) \in E(H \setminus \{a\})} w_{s_x s_y}(G_n),$$

³For a set S , the set S^N denotes the N -fold cartesian product $S \times S \times \dots \times S$.

where the sum is over tuples $\mathbf{s}_{\{a\}^c} := (s_x)_{x \in V(H) \setminus \{a\}} \in ([n] \setminus \{v\})_{|V(H)|-1}$ and $N_H(a)$ denotes the neighbors of a in the graph H . Then the empirical 1-point subgraph density function is defined as:

$$\hat{t}(v, H, G_n) := \frac{1}{|\text{Aut}(H)|} \sum_{a=1}^{|V(H)|} \frac{X_a(v, H, G_n)}{n^{|V(H)|-1}}. \quad (4.3)$$

Note that (4.3) counts (up to constant factors depending on the automorphisms of H) the fraction of copies of H in G_n passing through the vertex $v \in V(G_n)$. To illustrate we consider the following examples:

- $H = K_2$ is the edge: In this case, $\hat{t}(v, H, G_n) = \frac{d_v}{n}$, where d_v is the degree of the vertex v in G_n .
- $H = K_3$ is the triangle: Suppose the vertices of K_3 are labeled $\{1, 2, 3\}$. By symmetry, for all $1 \leq a \leq 3$,

$$X_a(v, K_3, G_n) = \sum_{1 \leq s_1 \neq s_2 \leq n} w_{vs_1} w_{vs_2} w_{s_1 s_2},$$

which is twice the number of triangles in G_n with v as one of the vertex. Therefore,

$$\hat{t}(v, K_3, G_n) = \frac{1}{2n^2} \sum_{1 \leq s_1 \neq s_2 \leq n} w_{vs_1} w_{vs_2} w_{s_1 s_2}.$$

- $H = K_{1,2}$ is the 2-star: Suppose the vertices of $K_{1,2}$ are labeled $\{1, 2, 3\}$ with the central vertex labeled 1. Then we have the following:
 - For $a = 1$, $X_1(v, K_{1,2}, G_n) = \sum_{1 \leq s_1 \neq s_2 \leq n} w_{vs_1} w_{vs_2}$, is twice the number of 2-stars in G_n with v as the central vertex.
 - For $a \in \{2, 3\}$, $X_a(v, K_{1,2}, G_n) = \sum_{s_1=1}^n w_{vs_1} (d_{s_1} - 1)$, is the number of 2-star in G_n where v is a leaf vertex.

Hence,

$$\hat{t}(v, K_{1,2}, G_n) = \frac{1}{2n^2} \left(\sum_{1 \leq s_1 \neq s_2 \leq n} w_{vs_1} w_{vs_2} + 2 \sum_{s_1=1}^n w_{vs_1} (d_{s_1} - 1) \right).$$

Next, we define the 2-point subgraph density of G_n , which is the empirical analogue of 2-point conditional kernel (2.5).

Definition 4.2. (Empirical 2-point subgraph density) Fix $a \neq b \in V(H)$ and $u, v \in V(G_n)$. Denote by $X_{a,b}(u, v, H, G_n)$ the number of injective homomorphism $\phi : V(H) \rightarrow V(G_n)$ such that $\phi(a) = u$ and $\phi(b) = v$. More formally,

$$X_{a,b}(u, v, H, G_n) = w_{uv}^+ \sum_{\mathbf{s}_{\{a,b\}^c}} \prod_{y \in N_H(a) \setminus \{b\}} w_{us_y}(G_n) \prod_{y \in N_H(b) \setminus \{a\}} w_{vs_y}(G_n) \prod_{(x,y) \in E(H \setminus \{a,b\})} w_{s_x s_y}(G_n),$$

where the sum is over tuples $\mathbf{s}_{\{a,b\}^c} := (s_x)_{x \in V(H) \setminus \{a,b\}} \in ([n] \setminus \{u, v\})_{|V(H)|-2}$ and $w_{uv}^+ = w_{uv}$ if $(a, b) \in E(H)$ and $w_{uv} = 1$ otherwise. The 2-point subgraph density is then defined as:

$$\hat{W}_H^{G_n}(u, v) = \frac{1}{2|\text{Aut}(H)|} \sum_{1 \leq a \neq b \leq |V(H)|} \frac{X_{a,b}(u, v, H, G_n)}{n^{|V(H)|-2}}. \quad (4.4)$$

By convention we define $\hat{W}_H^{G_n}(u, u) = 0$ for all $1 \leq u \leq n$.

Note that $((\hat{W}_H^{G_n}(u, v))_{1 \leq u, v \leq n})$ is a $n \times n$ matrix which counts (up to constant factors depending on the automorphisms of H) the fraction of copies of H in G_n passing through the vertices $u, v \in V(G_n)$. To illustrate we consider the following examples:

- $H = K_2$ is the edge: In this case, $\hat{W}_H^{G_n}(u, v) = \frac{1}{2}w_{uv}$, is the scaled adjacency matrix of G_n .
- $H = K_3$ is the triangle: Suppose the vertices of K_3 are labeled $\{1, 2, 3\}$. By symmetry, for all $1 \leq a, b \leq 3$,

$$X_{a,b}(u, v, K_3, G_n) = \sum_{s_1=1}^n w_{uv} w_{us_1} w_{vs_1},$$

which is the number of triangles in G_n with u and v as vertices. Therefore,

$$\hat{W}_{K_3}^{G_n}(u, v) = \frac{1}{2n} \sum_{1 \leq s_1 \leq n} w_{uv} w_{us_1} w_{vs_1}.$$

- $H = K_{1,2}$ is the 2-star: Suppose the vertices of $K_{1,2}$ are labeled $\{1, 2, 3\}$ with the central vertex labeled 1. Then we have the following for $1 \leq u \neq v \leq n$:
 - For $a = 1$ and $b \in \{2, 3\}$, $X_{1,b}(u, v, K_{1,2}, G_n) = w_{uv} \sum_{s_1=1}^n w_{us_1} - w_{uv} = w_{uv}(d_u - 1)$, is the number of 2-stars in G_n with u as the central vertex and v as the leaf vertex. Similarly, $X_{1,b}(v, u, K_{1,2}, G_n) = w_{vu} \sum_{s_1=1}^n w_{vs_1} - w_{vu} = w_{vu}(d_v - 1)$, is the number of 2-stars in G_n with v as the central vertex and u as the leaf vertex. Also, note that $X_{b,1}(u, v, K_{1,2}, G_n) = X_{1,b}(v, u, K_{1,2}, G_n)$.
 - For $a, b \in \{2, 3\}$, $X_{a,b}(u, v, K_{1,2}, G_n) = \sum_{s_1=1}^n w_{us_1} w_{vs_1}$, is the number of 2-star in G_n with u, v as leaf vertices.

Hence,

$$\hat{W}_{K_{1,2}}^{G_n}(u, v) = \frac{1}{2n} \left[w_{uv}(d_u + d_v - 2) + \sum_{s_1=1}^n w_{us_1} w_{vs_1} \right].$$

With the above definitions we can now describe multiplier bootstrap estimates of the limiting distribution $\mathbf{Z}(\mathcal{H}, W)$. For this, recall that $\mathcal{H} = \{H_1, H_2, \dots, H_r\}$ is such that W is irregular with respect to $H_1, H_2, \dots, H_q$ and W is regular with respect to $H_{q+1}, \dots, H_r$. Suppose $Z_1, Z_2, \dots, Z_n$ are i.i.d. $N(0, 1)$ independent of the graph G_n . Then define

$$\hat{\mathbf{Z}}(H_i, G_n) = \begin{cases} \frac{1}{\sqrt{n}} \sum_{v=1}^n (\hat{t}(v, H_i, G_n) - \bar{t}(H_i, G_n)) Z_v & \text{if } 1 \leq i \leq q, \\ \frac{1}{n} \sum_{1 \leq u, v \leq n} (\hat{W}_{H_i}^{G_n}(u, v) - \bar{W}_{H_i}^{G_n}) (Z_u Z_v - \delta_{u,v}) & \text{if } q+1 \leq i \leq r, \end{cases} \quad (4.5)$$

where $\delta_{u,v} = \mathbf{1}\{u = v\}$, $\bar{t}(H_i, G_n) = \frac{1}{n} \sum_{v=1}^n \hat{t}(v, H_i, G_n)$, and

$$\bar{W}_{H_i}^{G_n} = \frac{1}{n^2} \sum_{1 \leq u, v \leq n} \hat{W}_{H_i}^{G_n}(u, v). \quad (4.6)$$

Denote

$$\hat{\mathbf{Z}}(\mathcal{H}, G_n) = (\hat{\mathbf{Z}}(H_1, G_n), \hat{\mathbf{Z}}(H_2, G_n), \dots, \hat{\mathbf{Z}}(H_r, G_n))^\top. \quad (4.7)$$

Note that $\hat{\mathbf{Z}}(\mathcal{H}, G_n)$ depends only on the observed graph G_n and the Gaussian multipliers $Z_1, Z_2, \dots, Z_n$, but not on the graphon W . In the following theorem we show that $\hat{\mathbf{Z}}(\mathcal{H}, G_n)$, conditional on the graph G_n , converges to $\mathbf{Z}(\mathcal{H}, W)$ as in Theorem 2.1.

Theorem 4.1. *Fix a graphon W and a finite collection of non-empty graphs $\mathcal{H} = \{H_1, H_2, \dots, H_r\}$ such that W is irregular with respect to $H_1, \dots, H_q$, for some $1 \leq q \leq r$, and regular with respect to $H_{q+1}, H_{q+2}, \dots, H_r$. Suppose G_n is a realization from $G(n, W)$ and $\hat{\mathbf{Z}}_n(\mathcal{H}, G_n)$ be as defined in (4.7). Then, almost surely as $n \rightarrow \infty$,*

$$\hat{\mathbf{Z}}_n(\mathcal{H}, G_n) | G_n \xrightarrow{D} \mathbf{Z}(\mathcal{H}, W), \quad (4.8)$$

where $\mathbf{Z}(\mathcal{H}, W)$ is as in (2.12).

The proof of Theorem 4.1 is given in Section D. It shows the asymptotic distribution of $\hat{\mathbf{Z}}_n(\mathcal{H}, G_n)|G_n$ is the same as that of the subgraph counts $\mathbf{Z}(H, G_n)$ (recall (2.9)). Hence, we can use the distribution $\hat{\mathbf{Z}}_n(\mathcal{H}, G_n)|G_n$, which depends only on the observed graph G_n , to approximate the quantiles of the limiting distribution $\mathbf{Z}(\mathcal{H}, W)$. This allow us to construct joint confidence sets for the homomorphism densities as described in Section 5.

Remark 4.1. Recently, Lin et al. [54] proposed a bootstrap method for approximating the sampling distribution of a network moment in the sparse regime (where the networks have $o(n^2)$ edges), which bears some similarity to our approach. Specifically, the authors use a multiplier bootstrap to estimate the terms in the Hoeffding decomposition of a network moment and also approximates the local subgraph counts based on sampling for fast computation. However, as in most prior work, the bootstrap consistency essentially requires the network moment to be have a non-degenerate Gaussian limit. Moreover, the result only applies in the sparse regime and for the marginal distribution a single network moment that is either acyclic or a cycle.

5. JOINT CONFIDENCE SETS

Suppose $\mathcal{H} = \{H_1, H_2, \dots, H_r\}$ is a collection of non-empty graphs, with $H_i = (V(H_i), E(H_i))$ and $|V(H_i)| \geq 2$, for $1 \leq i \leq r$. In this section, we construct a joint confidence set for the collection of homomorphism densities

$$\mathbf{t}(\mathcal{H}, W) = (t(H_1, W), t(H_2, W), \dots, t(H_r, W))^\top,$$

given a sample G_n from $G(n, W)$. Note that, although Theorem 4.1 provides a way to estimate the quantiles of the limiting distribution $\mathbf{Z}(\mathcal{H}, W)$, this result cannot be directly applied for constructing a confidence set, because it is a-priori unknown whether or not W is H_i -regular for some $1 \leq i \leq r$. For this, we propose a *testimation* strategy for constructing joint confidence sets, which first tests for H_i -regularity based on the observed graph G_n , for $1 \leq i \leq r$, and then uses Theorem 4.1 to estimate the appropriate quantiles. The rest of this section is organized as follows: In Section 5.1 we discuss the test for regularity. Using this and the graphon multiplier bootstrap from the previous section we provide an algorithm for constructing confidence sets in Section 5.2. We illustrate the performance of the algorithm in simulations in Section 5.3.

5.1. Testing for Regularity. Given a graphon W and finite simple graph $H = (V(H), E(H))$, with $|V(H)| \geq 2$, the regularity testing problem for the pair (H, W) can be formulated as follows:

$$H_0 : W \text{ is } H\text{-regular} \quad \text{versus} \quad H_1 : W \text{ is not } H\text{-regular}. \quad (5.1)$$

Recall that W is H -regular if and only if the asymptotic variance $\tau_{H,W} = 0$ (recall (2.14)). For notational convinience define,

$$R(H, W) = |\text{Aut}(H)|^2 \tau_{H,W}^2 = \sum_{1 \leq a, b \leq |V(H)|} t \left(H \oplus_{a,b} H, W \right) - |V(H)|^2 t(H, W)^2. \quad (5.2)$$

Clearly, W is H -regular if and only if $R(H, W) = 0$. Note that, since the vertex joins to 2 simple graphs is another simple graph, $R(H, W)$ is a function of homomorphism densities of simple graphs. Hence, $R(H, W)$ can be consistently estimated from G_n based on the following simple estimate:

$$R(H, G_n) = \sum_{1 \leq a, b \leq |V(H)|} \hat{t} \left(H \oplus_{a,b} H, G_n \right) - |V(H)|^2 \hat{t}(H, G_n)^2. \quad (5.3)$$

The following lemma shows that $R(H, G_n)$ converges to zero at rate faster than $\sqrt{n}$ when W is H -regular.

Proposition 5.1. *Suppose $R(H, G_n)$ be as defined in (5.3). Then the following hold:*

- (1) *When W is H -regular, $\sqrt{n}R(H, G_n) \xrightarrow{P} 0$.*
- (2) *When W is not H -regular, $\sqrt{n}R(H, G_n) \xrightarrow{P} \infty$.*

The proof of Proposition 5.1 is given in Section E.1. Now, consider the test function

$$\phi(H, G_n) = \mathbf{1} \{ \sqrt{n}R(H, G_n) > 1 \}. \quad (5.4)$$

Proposition 5.1 implies that under H_0 as in (5.1), $\mathbb{P}(\phi(H, G_n)) \rightarrow 0$, and under H_1 , $\mathbb{P}(\phi(H, G_n)) \rightarrow 1$. Hence, the test (5.4) is consistent for the regularity testing problem (5.1).

5.2. Constructing Confidence Sets. Using the test for regularity, we can now describe our algorithm for constructing a joint confidence set for $\mathbf{t}(\mathcal{H}, W)$ as follows:

- For each $1 \leq i \leq r$, consider the hypothesis testing problem:

$$H_{0,i} : W \text{ is } H_i\text{-regular} \quad \text{versus} \quad H_{1,i} : W \text{ is not } H_i\text{-regular}.$$

Let $S(\mathcal{H}, G_n) := \{1 \leq i \leq r : \sqrt{n}R(H_i, G_n) > 1\}$, be the set of indices where the hypothesis of H_i -regularity is rejected.

- Define

$$\mathbf{Q}(\mathcal{H}, G_n) = (Q(H_1, G_n), \dots, Q(H_r, G_n)), \quad (5.5)$$

where (recall (4.5))

$$Q(H_i, G_n) = \begin{cases} \frac{1}{\sqrt{n}} \sum_{v=1}^n (\hat{t}(v, H_i, G_n) - \bar{t}(H_i, G_n)) Z_v & \text{if } i \in S(\mathcal{H}, G_n), \\ \frac{1}{n} \sum_{1 \leq u, v \leq n} (\hat{W}_{H_i}^{G_n}(u, v) - \bar{W}_{H_i}^{G_n}) (Z_u Z_v - \delta_{u,v}) & \text{if } i \notin S(\mathcal{H}, G_n), \end{cases}$$

with $Z_1, Z_2, \dots, Z_n$ are i.i.d. $N(0, 1)$ independent of the graph G_n . Denote by $\hat{q}_{1-\alpha, \mathcal{H}, G_n}$ the $(1 - \alpha)$ -th quantile of distribution of $\|\mathbf{Q}(\mathcal{H}, G_n)\|_2 | G_n$. (Note that the distribution of $\mathbf{Q}(\mathcal{H}, G_n)$ given G_n does not depend on the graphon W , it only depends on the randomness of the Gaussian multipliers $\{Z_u\}_{1 \leq u \leq n}$. Hence, in practice, $\hat{q}_{1-\alpha, \mathcal{H}, G_n}$ will be computed from the empirical quantiles of $\|\mathbf{Q}(\mathcal{H}, G_n)\|_2 | G_n$ obtained by repeatedly sampling the Gaussian multipliers.)

- Define

$$\tilde{\mathbf{Z}}(\mathcal{H}, G_n) = (\tilde{Z}(H_1, G_n), \tilde{Z}(H_2, G_n), \dots, \tilde{Z}(H_r, G_n))^\top, \quad (5.6)$$

with

$$\tilde{Z}(H_i, G_n) = \begin{cases} \frac{X(H_i, G_n) - \frac{(n)^{|V(H_i)|} t(H_i, W)}{|Aut(H_i)|}}{n^{|V(H_i)| - \frac{1}{2}}} & \text{if } i \in S(\mathcal{H}, G_n), \\ \frac{X(H_i, G_n) - \frac{(n)^{|V(H_i)|} t(H_i, W)}{|Aut(H_i)|}}{n^{|V(H_i)| - 1}} & \text{if } i \notin S(\mathcal{H}, G_n). \end{cases} \quad (5.7)$$

Report the confidence set

$$\mathcal{C}(\mathcal{H}, G_n) = \{\mathbf{t}(\mathcal{H}, W) : \|\tilde{\mathbf{Z}}(\mathcal{H}, G_n)\|_2 \leq \hat{q}_{1-\alpha, \mathcal{H}, G_n}\}, \quad (5.8)$$

where $\mathbf{t}(\mathcal{H}, W) = (t(H_1, W), t(H_2, W), \dots, t(H_r, W))^\top$.

The following theorem shows that the set $C(\mathcal{H}, G_n)$ is a confidence set for the vector of homomorphism densities $\mathbf{t}(\mathcal{H}, W)$ with asymptotically α coverage probability.

Theorem 5.1. *Let $C(\mathcal{H}, G_n)$ be as defined in (5.8). Then $\lim_{n \rightarrow \infty} \mathbb{P}(C(\mathcal{H}, G_n)) = 1 - \alpha$.*

The proof of Theorem 5.1 is given in Section E.2. The proof involves showing that $\mathbf{Q}(\mathcal{H}, G_n)|G_n$ and $\tilde{\mathbf{Z}}(\mathcal{H}, G_n)$ (recall (5.5) and (5.6), respectively), both converge to the distribution of $\mathbf{Z}(\mathcal{H}, W)$ asymptotically.

Remark 5.1. (Marginal Confidence Intervals) The algorithm for constructing joint confidence sets described above takes a simpler form when $\mathcal{H} = \{H\}$ is a singleton. In other words, suppose, we want to construct a (marginal) confidence interval for $t(H, W)$. Then, recalling Corollary 2.1, we proceed as follows:

- If $\sqrt{n}R(H, G_n) > 1$ (that is, H_0 in (5.1) is rejected), then define

$$L(H, G_n) = \left[\hat{t}(H, G_n) - z_{\alpha/2} \frac{|\text{Aut}(H)| \hat{\tau}_{H, G_n}}{\sqrt{n}}, \hat{t}(H, G_n) + z_{\alpha/2} \frac{|\text{Aut}(H)| \hat{\tau}_{H, G_n}}{\sqrt{n}} \right],$$

where $\hat{t}(H, G_n) = \frac{|\text{Aut}(H)|}{(n)^{|V(H)|}} X(H, G_n)$ (as defined in (1.4)),

$$\hat{\tau}_{H, G_n}^2 = \frac{1}{n} \sum_{v=1}^n (\hat{t}(v, H, G_n) - \bar{t}(H, G_n))^2,$$

and z_α is the $(1 - \alpha)$ -th quantile of standard Gaussian distribution.

- If $\sqrt{n}R(H, G_n) \leq 1$ (that is, H_0 in (5.1) is accepted), then define

$$L(H, G_n) = \left[\hat{t}(H, G_n) - \hat{q}_{\alpha/2, H, G_n} \frac{|\text{Aut}(H)|}{n}, \hat{t}(H, G_n) - \hat{q}_{1-\alpha/2, H, G_n} \frac{|\text{Aut}(H)|}{n} \right].$$

Here, $\hat{q}_{1-\alpha, H, G_n}$ is the α -th quantile of the random variable $\frac{1}{n} \sum_{i=1}^n \lambda_i(H, G_n) (Z_i^2 - 1) |G_n$, where $\{\lambda_i(H, G_n)\}_{1 \leq i \leq n}$ are the eigenvalues of the matrix $((\hat{W}_H^{G_n}(u, v) - \bar{W}_H^{G_n}))_{1 \leq u, v \leq n}$ (recall (4.4) and (4.6)) and $\{Z_i\}_{1 \leq i \leq n}$ are i.i.d. $N(0, 1)$.

From Corollary 2.1 and the proof of Theorem 5.1, it easily follows that $\lim_{n \rightarrow \infty} \mathbb{P}(L(H, G_n)) = 1 - \alpha$, that is, $L(H, G_n)$ is an asymptotically valid confidence interval for $t(H, W)$.

5.3. Simulations. In this section we evaluate the performance of the algorithm for constructing joint confidence sets in simulations.

5.3.1. Confidence Interval for the Edge Density. For the confidence interval of the edge density $t(K_2, W)$ we consider the following 2 choices of W :

- $W = W_-(x, y) := xy$, for $x, y \in [0, 1]$. This graphon is K_2 -irregular (the degree function $d_{W_-}(x) = \frac{x}{2}$) and $t(K_2, W_-) = \frac{1}{4}$.
- Next, we consider the K_2 -regular graphon

$$W = W_+(x, y) := \begin{cases} \frac{1}{2} & \text{if } (x, y) \in [0, \frac{1}{2}] \times [\frac{1}{2}, 1] \cup [\frac{1}{2}, 1] \times [0, \frac{1}{2}], \\ 0 & \text{otherwise.} \end{cases}$$

Note that this graphon corresponds to the random bipartite graph with equal block sizes and edge probability $\frac{1}{2}$. Note that $t(K_2, W_+) = \frac{1}{4}$.

Using the method described in Remark 5.1 we construct 100 instances of the 95% confidence interval for $t(K_2, W)$, when $W = W_-$ (Figure 6(a)) and $W = W_+$ (Figure 6(b)). Each interval is computed based on a graph of size $n = 400$ sampled from the model $G(n, W)$, for $W \in \{W_-, W_+\}$, and the quantiles are estimated using 1000 resamples from the conditional distribution. The black horizontal line represents the population edge density $t(K_2, W) = \frac{1}{4}$ (in both cases) and

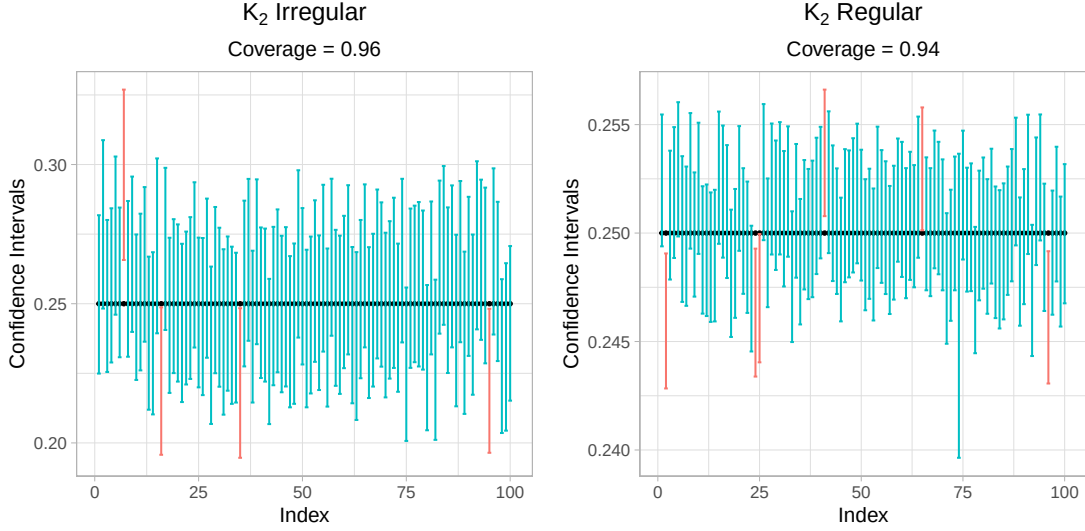


FIGURE 6. 100 instances of the 95% confidence interval for edge density $t(K_2, W)$ with (a) $W = W_-$ and (b) $W = W_+$.

the intervals not containing $\frac{1}{4}$ are shown in red. Observe that in both cases the fraction of intervals containing $\frac{1}{4}$ (the empirical coverage probability) is very close to 0.95, as predicted by the asymptotic theory.

5.3.2. Joint Confidence Sets for Edge and Triangle Densities. We now use our algorithm to construct the joint confidence set for the edge and the triangle densities $(t(K_2, W), t(K_3, W))$. Here, 4 possible cases can arise depending on whether or not W is K_2 or K_3 -regular (recall Example 3.1). For each of the 4 graphons considered in Example 3.1, we show below the heatmap of 100 instances of the 95% confidence ellipsoid (recall (5.8)). In all the simulations, the confidence sets are computed based on graphs of size $n = 400$ and the quantiles are estimated using 1000 resamples from the conditional distribution. The empirical coverage is given by the fraction of confidence ellipsoids that contain the true homomorphism densities (which is marked by the black point).

- Figure 7(a) shows the joint confidence sets for $(t(K_2, W), t(K_3, W))$ when $W = \tilde{W}_1(x, y) = \frac{1}{2}(x + y)$ (recall (3.1)). This graphon is both K_2 and K_3 -irregular. Also, for this graphon $(t(K_2, \tilde{W}_1), t(K_3, \tilde{W}_1)) = (\frac{1}{2}, \frac{5}{32})$. In this case, the empirical coverage is 94%.
- Figure 7(b) shows the joint confidence sets for $(t(K_2, W), t(K_3, W))$ when $W = \tilde{W}_2(x, y)$ is the graphon in (3.2). This graphon is K_2 -regular and K_3 -irregular. Furthermore, $(t(K_2, \tilde{W}_2), t(K_3, \tilde{W}_2)) = (\frac{1}{3}, \frac{1}{27})$. In this case, the empirical coverage is 96%.
- Figure 8(a) shows the joint confidence sets for $(t(K_2, W), t(K_3, W))$ when $W = \tilde{W}_3(x, y)$ is the graphon shown in Figure 5. This graphon is K_2 -irregular and K_3 -regular. Furthermore, a direct computation shows that $(t(K_2, \tilde{W}_3), t(K_3, \tilde{W}_3)) = (\frac{13}{36}, \frac{1}{18})$. In this case, the empirical coverage is 95%.
- Figure 8(b) shows the joint confidence sets for $(t(K_2, W), t(K_3, W))$ when $W = W_{\frac{1}{2}} \equiv \frac{1}{2}$ is the constant function $\frac{1}{2}$ (which corresponds to the Erdős-Rényi graph $G(n, \frac{1}{2})$). This graphon is both K_2 and K_3 -regular. Also, $(t(K_2, W), t(K_3, W)) = (\frac{1}{2}, \frac{1}{8})$. In this case, the empirical coverage is 94%.

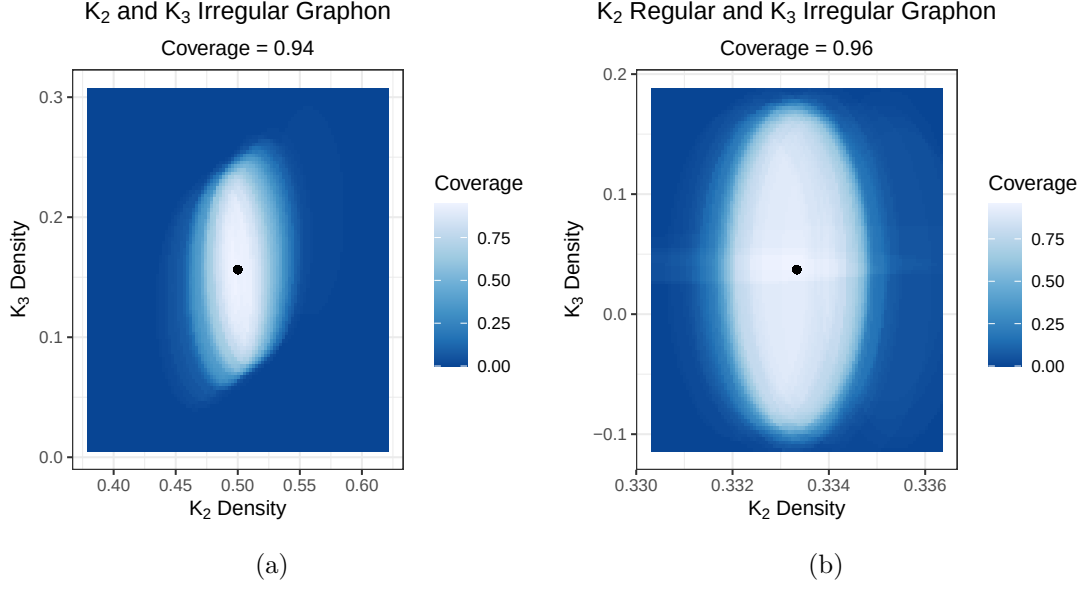


FIGURE 7. 100 instances of the 95% confidence sets for $(t(K_2, W), t(K_3, W))$ with (a) $W = \tilde{W}_1$ and (b) $W = \tilde{W}_2$.

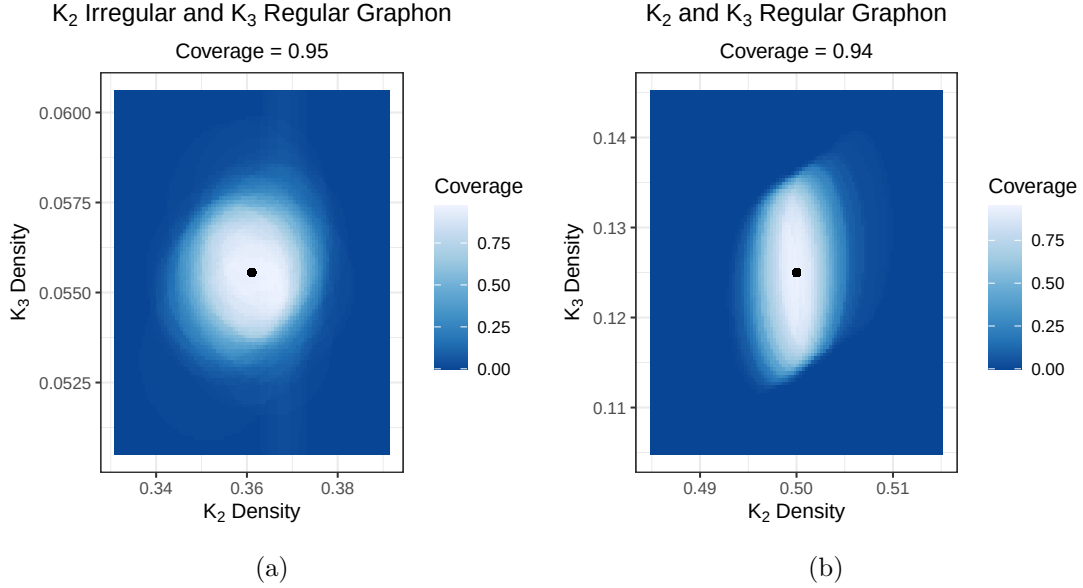


FIGURE 8. 100 instances of the 95% confidence sets for $(t(K_2, W), t(K_3, W))$ with (a) $W = \tilde{W}_3$ the graphon in Figure 5 and (b) $W = W_{\frac{1}{2}} \equiv \frac{1}{2}$ the constant graphon $\frac{1}{2}$.

The results above show that the proposed method achieves the desired coverage in different simulation settings. It is worth recalling our method does not have any prior knowledge about whether or not W is K_2 or K_3 -regular. We first test for the presence of regularity and construct the confidence sets depending on the outcome of the tests as described in Section 5.2.

6. TESTING FOR GLOBAL STRUCTURE

Testing global network properties based on counts of subgraphs is a central theme in many statistical network analysis problems. A basic problem in this direction is to test whether the network is generated completely at random or whether it has some additional structure. In the context of stochastic models this entails testing whether or not the network has any community structure [28, 29]. For graphon models, global structure testing can be formulated as the following hypothesis (recall (1.5)):

$$H_0 : W = p \text{ almost everywhere for some } p \in (0, 1) \text{ versus } H_1 : W \text{ is non-constant.} \quad (6.1)$$

To find a consistent test for this hypothesis, we need to find a functional $f : \mathcal{W} \rightarrow \mathbb{R}$ which has the property that $f(W) = 0$ if and only if W is a constant function almost everywhere. A classical result of Chung, Graham, and Wilson about quasi-random graphs [20] implies that the function $f(W) = t(K_2, W)^4 - t(C_4, W)$ satisfies this property (see [55, Claim 11.53]). Hence, one can construct a consistent test for (6.1) by estimating this functional based on the observed graph G_n . To this end, define,

$$\hat{f}(G_n) = \hat{t}(K_2, G_n)^4 - \hat{t}(C_4, G_n), \quad (6.2)$$

where, for any graph H , $\hat{t}(H, G_n) = \frac{|\text{Aut}(H)|}{(n)_{|V(H)|}} X(H, G_n)$ (as defined in (1.4)).

In the following theorem we derive the asymptotic distribution of $\hat{f}(G_n)$ under the null hypothesis.

Proposition 6.1. *Under H_0 as in (1.5),*

$$n^{\frac{3}{2}} \hat{f}(G_n) \xrightarrow{D} N(0, \vartheta^2). \quad (6.3)$$

where $\vartheta^2 := 32t(K_2, W)^6(1 - t(K_2, W))^2$.

The proof is given in Section F.1. As in the proof of Theorem 2.1, it uses the method of orthogonal projections. One interesting feature of the statistic $\hat{f}(G_n)$ is that it has fluctuations of order $O(n^{-\frac{3}{2}})$ under H_0 , even though we know from Example 3.2 that both $\hat{t}(K_2, G_n)$ and $\hat{t}(C_4, G_n)$ have fluctuations of $O(1/n)$. This means $\hat{f}(G_n)$ cancels the contributions from $\hat{t}(K_2, G_n)$ and $\hat{t}(C_4, G_n)$ in the $O(1/n)$ scale and the leading asymptotic contribution of $\hat{f}(G_n)$ is determined from the third-order projection. The same scaling appears in the Erdős-Zuckerberg (EZ) statistic considered in [29], for testing the presence of community structure in degree-corrected block models.

To apply Proposition 6.1 to test the hypothesis (6.1), we need to consistently estimate the asymptotic variance in (6.3). Towards this, note that, since $\hat{t}(K_2, G_n) \xrightarrow{P} t(K_2, W)$ (follows from (E.1) and Corollary 10.4 from [55]), by Slutsky's lemma:

$$T_n := n^{\frac{3}{2}} \frac{\hat{f}(G_n)}{4\sqrt{2} \hat{t}(K_2, G_n)^3(1 - \hat{t}(K_2, G_n))} \xrightarrow{D} N(0, 1), \quad (6.4)$$

under H_0 . Hence, the test which rejects when $|T_n| > z_{\alpha/2}$ is asymptotically level α . In the following proposition we show that this test is consistent, that is, it can detect any non-constant graphon with probability going to 1 (see Section F.2 for the proof).

Proposition 6.2. *For any graphon W such that $|t(K_2, W)^4 - t(C_4, W)| > 0$ we have,*

$$\mathbb{P}(|T_n| > z_{\alpha/2}) \rightarrow 1.$$

Proposition 6.2 provides a test for (6.1) that has precise asymptotic level and is consistent in detecting all non-constant graphons. In comparison, the asymptotic null distribution of the test statistic in Fang and Röllin [26] is unknown and the resulting test is conservative (see [26, Remark 3.3]). The framework of orthogonal projections and the results obtained in Section 2 allow us derive the asymptotic distribution of $\hat{f}(G_n)$ both under the null (as in Proposition 6.1) and the alternative (see Proposition F.1 in Section F.3). This will allow us to approximate the asymptotic power of the test based on T_n (recall (6.4)), and also obtain a confidence interval for $f(W)$ using the method in Section 5.

Acknowledgements. B. B. Bhattacharya was supported by NSF CAREER grant DMS 2046393, NSF grant DMS 2113771, and a Sloan Research Fellowship.

REFERENCES

- [1] D. J. Aldous. Representations for partially exchangeable arrays of random variables. *Journal of Multivariate Analysis*, 11(4):581–598, 1981.
- [2] U. Alon. Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, 8(6):450–461, 2007.
- [3] A. Athreya, D. E. Fishkind, M. Tang, C. E. Priebe, Y. Park, J. T. Vogelstein, K. Levin, V. Lyzinski, Y. Qin, and D. L. Sussman. Statistical inference on random dot product graphs: a survey. *Journal of Machine Learning Research*, 18(226):1–92, 2018.
- [4] M. Austern and P. Orbanz. Limit theorems for distributions invariant under groups of transformations. *The Annals of Statistics*, 50(4):1960–1991, 2022.
- [5] A. D. Barbour, M. Karoński, and A. Ruciński. A central limit theorem for decomposable random variables with applications to random graphs. *Journal of Combinatorial Theory, Series B*, 47(2):125–145, 1989.
- [6] A. Bhattacharya, A. Bishnu, A. Ghosh, and G. Mishra. On triangle estimation using tripartite independent set queries. *Theory of Computing Systems*, 65(8):1165–1192, 2021.
- [7] B. B. Bhattacharya, S. Das, and S. Mukherjee. Motif estimation via subgraph sampling: The fourth-moment phenomenon. *The Annals of Statistics*, 50(2):987–1011, 2022.
- [8] B. B. Bhattacharya, A. Chatterjee, and S. Janson. Fluctuations of subgraph counts in graphon based random graphs. *Combinatorics, Probability and Computing*, 32(3):428–464, 2023.
- [9] S. Bhattacharyya and P. J. Bickel. Subsampling bootstrap of count features of networks. *The Annals of Statistics*, 43(6):2384 – 2411, 2015.
- [10] P. J. Bickel and A. Chen. A nonparametric view of network models and newman–girvan and other modularities. *Proceedings of the National Academy of Sciences*, 106(50):21068–21073, 2009.
- [11] P. J. Bickel, A. Chen, and E. Levina. The method of moments and degree distributions for network models. *The Annals of Statistics*, 39(5):2280–2301, 2011.
- [12] M. Boguná and R. Pastor-Satorras. Class of correlated random networks with hidden variables. *Physical Review E*, 68(3):036112, 2003.
- [13] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures & Algorithms*, 31(1):3–122, 2007.
- [14] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztegombi. Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing. *Advances in Mathematics*, 219(6):1801–1851, 2008.
- [15] C. Borgs, J. Chayes, and L. Lovász. Moments of two-variable functions and the uniqueness of graph limits. *Geometric and Functional Analysis*, 19:1597–1619, 2010.
- [16] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztegombi. Convergent sequences of dense graphs II. multiway cuts and statistical physics. *Annals of Mathematics*, pages 151–219, 2012.
- [17] G. Bravo-Hernsdorff, L. M. Gunderson, P.-A. Maugis, and C. E. Priebe. Quantifying network similarity using graph cumulants. *Journal of Machine Learning Research*, 24(187):1–27, 2023.
- [18] A. Chatterjee and J. Huang. Fluctuation of the largest eigenvalue of a kernel matrix with application in graphon-based random graphs. *arXiv:2401.01866*, 2024.

- [19] S. Chatterjee and P. Diaconis. Estimating and understanding exponential random graph models. *The Annals of Statistics*, 41(5):2428–2461, 2013.
- [20] F. R. K. Chung, R. L. Graham, and R. M. Wilson. Quasi-random graphs. *Combinatorica*, 9:345–362, 1989.
- [21] H. Crane. *Probabilistic Foundations of Statistical Network Analysis*. CRC Press, 2018.
- [22] J.-F. Delmas, J.-S. Dhersin, and M. Sciaudeau. Asymptotic for the cumulative distribution function of the degrees and homomorphism densities for random graphs sampled from a graphon. *Random Structures & Algorithms*, 58(1):94–149, 2021.
- [23] P. Diaconis and D. Freedman. On the statistics of vision: the Julesz conjecture. *Journal of Mathematical Psychology*, 24(2):112–138, 1981.
- [24] P. Diaconis and S. Janson. Graph limits and exchangeable random graphs. *Rendiconti di Matematica e delle sue Applicazioni. Serie VII*, pages 33–61, 2008.
- [25] T. Eden, A. Levi, D. Ron, and C. Seshadhri. Approximately counting triangles in sublinear time. *SIAM Journal on Computing*, 46(5):1603–1646, 2017.
- [26] X. Fang and A. Röhl. Rates of convergence for multivariate normal approximation with applications to dense graphs and doubly indexed permutation statistics. *Bernoulli*, 21(4):2157 – 2189, 2015.
- [27] V. Féray, P.-L. Méliot, and A. Nikeghbali. Graphons, permutons and the Thoma simplex: three mod-Gaussian moduli spaces. *Proceedings of the London Mathematical Society*, 121(4):876–926, 2020.
- [28] C. Gao and J. Lafferty. Testing network structure using relations between small subgraph probabilities. *arXiv:1704.06742*, 2017.
- [29] C. Gao and J. Lafferty. Testing for global network structure using small subgraph statistics. *arXiv:1710.00862*, 2017.
- [30] C. Gao, Y. Lu, and H. H. Zhou. Rate-optimal graphon estimation. *The Annals of Statistics*, 43(6):2624–2652, 2015.
- [31] D. Ghoshdastidar, M. Gutzeit, A. Carpentier, and U. von Luxburg. Two-sample tests for large random graphs using network statistics. In *Conference on Learning Theory*, pages 954–977. PMLR, 2017.
- [32] M. Gonen, D. Ron, and Y. Shavitt. Counting stars and other small subgraphs in sublinear-time. *SIAM Journal on Discrete Mathematics*, 25(3):1365–1411, 2011.
- [33] A. Green and C. R. Shalizi. Bootstrapping exchangeable random graphs. *Electronic Journal of Statistics*, 16(1):1058–1095, 2022.
- [34] J. Hladký, C. Pelekis, and M. Šileikis. A limit theorem for small cliques in inhomogeneous random graphs. *Journal of Graph Theory*, 97(4):578–599, 2021.
- [35] P. D. Hoff, A. E. Raftery, and M. S. Handcock. Latent space approaches to social network analysis. *Journal of The American Statistical Association*, 97(460):1090–1098, 2002.
- [36] P. W. Holland and S. Leinhardt. Transitivity in structural models of small groups. *Comparative Group Studies*, 2(2):107–124, 1971.
- [37] P. W. Holland, K. B. Laskey, and S. Leinhardt. Stochastic blockmodels: First steps. *Social Networks*, 5(2):109–137, 1983.
- [38] D. N. Hoover. Row-column exchangeability and a generalized model for probability. *Exchangeability in probability and statistics (Rome, 1981)*, pages 281–291, 1982.
- [39] K. H. Huang, M. Austern, and P. Orbanz. Gaussian universality for approximately polynomial functions of high-dimensional data. *arXiv:2403.10711*, 2024.
- [40] D. R. Hunter, M. S. Handcock, C. T. Butts, S. M. Goodreau, and M. Morris. ergm: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of Statistical Software*, 24(3):nihpa54860, 2008.
- [41] K. Itô. Multiple wiener integral. *Journal of the Mathematical Society of Japan*, 3(1):157–169, 1951.
- [42] S. Janson. A functional limit theorem for random graphs with applications to subgraph count statistics. *Random Structures & Algorithms*, 1(1):15–37, 1990.
- [43] S. Janson. Orthogonal decompositions and functional limit theorems for random graph statistics. *Memoirs Amer. Math. Soc.*, 111(534), 1994.

- [44] S. Janson. *Gaussian Hilbert spaces*, volume 129 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1997.
- [45] S. Janson and K. Nowicki. The asymptotic distributions of generalized U -statistics with applications to random graphs. *Probability Theory and Related Fields*, 90(3):341–375, 1991.
- [46] S. Janson, T. Luczak, and A. Rucinski. *Random Graphs*. John Wiley & Sons, 2011.
- [47] G. Kaur and A. Röllin. Higher-order fluctuations in dense random graph models. *Electronic Journal of Probability*, 26:1–36, 2021.
- [48] J. M. Klusowski and Y. Wu. Counting motifs with graph sampling. In *Conference On Learning Theory*, pages 1966–2011, 2018.
- [49] P. D. Lax. *Functional Analysis*. Pure and Applied Mathematics (New York). Wiley-Interscience [John Wiley & Sons], New York, 2002.
- [50] T. Le Minh. U -statistics on bipartite exchangeable networks. *ESAIM: Probability and Statistics*, 27: 576–620, 2023.
- [51] J. Lei. A goodness-of-fit test for stochastic block models. *The Annals of Statistics*, 44(1):401 – 424, 2016.
- [52] J. Lei. Network representation using graph root distributions. *The Annals of Statistics*, 49(2):745 – 768, 2021.
- [53] K. Levin and E. Levina. Bootstrapping networks with latent space structure. *arXiv:1907.10821*, 2019.
- [54] Q. Lin, R. Lunde, and P. Sarkar. Trading off accuracy for speedup: Multiplier bootstraps for subgraph counts. *arXiv:2009.06170v5*, 2022.
- [55] L. Lovász. *Large Networks and Graph Limits*, volume 60. American Mathematical Soc., 2012.
- [56] L. Lovász and B. Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933–957, 2006.
- [57] R. D. Luce and A. D. Perry. A method of matrix analysis of group structure. *Psychometrika*, 14(2): 95–116, 1949.
- [58] R. Lunde and P. Sarkar. Subsampling sparse graphons under minimal assumptions. *Biometrika*, 110 (1):15–32, 2023.
- [59] H. Lyu, F. Memoli, and D. Sivakoff. Sampling random graph homomorphisms and applications to network data analysis. *Journal of Machine Learning Research*, 24(9):1–79, 2023.
- [60] P. Maugis. Central limit theorems for local network statistics. *arXiv:2006.15738*, 2020.
- [61] P. Maugis, C. E. Priebe, S. C. Olhede, and P. J. Wolfe. Statistical inference for network samples using subgraph counts. *arXiv:1701.00505*, 2017.
- [62] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002.
- [63] E. Mossel, J. Neeman, and A. Sly. Consistency thresholds for the planted bisection model. *Electronic Journal of Probability*, 21(21):1–24, 2016.
- [64] S. Mukherjee. Degeneracy in sparse ergms with functions of degrees as sufficient statistics. *Bernoulli*, 26(2):1016–1043, 2020.
- [65] S. Mukherjee and Y. Xu. Statistics of the two star ergm. *Bernoulli*, 29(1):24–51, 2023.
- [66] K. Nowicki. Asymptotic normality of graph statistics. *Journal of Statistical Planning and Inference*, 21(2):209–222, 1989.
- [67] K. Nowicki and J. C. Wierman. Subgraph counts in random graphs using incomplete U -statistics methods. *Discrete Mathematics*, 72(1-3):299–310, 1988.
- [68] S. Ouadah, P. Latouche, and S. Robin. Motif-based tests for bipartite networks. *Electronic Journal of Statistics*, 16(1):293–330, 2022.
- [69] M. Penrose. *Random Geometric Graphs*, volume 5. OUP Oxford, 2003.
- [70] N. Pržulj. Biological network comparison using graphlet degree distribution. *Bioinformatics*, 23(2): e177–e183, 2007.
- [71] M. Reed and B. Simon. *Methods of Modern Mathematical Physics. vol. 1. Functional Analysis*. Academic New York, 1980.
- [72] M. Renardy and R. C. Rogers. *An Introduction to Partial Differential Equations*, volume 13 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition, 2004. ISBN 0-387-00444-0.

- [73] P. Rubin-Delanchy, J. Cape, M. Tang, and C. E. Priebe. A statistical interpretation of spectral embedding: The generalised random dot product graph. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84(4):1446–1473, 2022.
- [74] M. Rubinov and O. Sporns. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, 52(3):1059–1069, 2010.
- [75] A. Ruciński. When are small subgraphs of a random graph normally distributed? *Probability Theory and Related Fields*, 78(1):1–10, 1988.
- [76] C. R. Shalizi and A. Rinaldo. Consistency under sampling of exponential random graph models. *The Annals of statistics*, 41(2):508, 2013.
- [77] M. Shao, D. Xia, Y. Zhang, Q. Wu, and S. Chen. Higher-order accurate two-sample network inference and network hashing. *arXiv:2208.07573*, 2022.
- [78] S. S. Shen-Orr, R. Milo, S. Mangan, and U. Alon. Network motifs in the transcriptional regulation network of escherichia coli. *Nature Genetics*, 31(1):64–68, 2002.
- [79] M. Tang, A. Athreya, D. L. Sussman, V. Lyzinski, and C. E. Priebe. A nonparametric two-sample hypothesis testing problem for random graphs. *Bernoulli*, 23(3):1599 – 1630, 2017.
- [80] S. Wasserman and K. Faust. *Social Network Analysis: Methods and applications*. Cambridge university press, 1994.
- [81] D. J. Watts and S. H. Strogatz. Collective dynamics of ‘small-world’ networks. *nature*, 393(6684):440–442, 1998.
- [82] A. E. Wegner, L. Ospina-Forero, R. E. Gaunt, C. M. Deane, and G. Reinert. Identifying networks with common organizational principles. *Journal of Complex Networks*, 6(6):887–913, 2018.
- [83] W. Xu and G. Reinert. A stein goodness-of-test for exponential random graph models. In *International Conference on Artificial Intelligence and Statistics*, pages 415–423. PMLR, 2021.
- [84] Y. Xu and S. Mukherjee. Signal detection in degree corrected ERGMs. *arXiv:2108.09255*, 2021.
- [85] Y. Zhang and D. Xia. Edgeworth expansions for network moments. *The Annals of Statistics*, 50(2):726–753, 2022.
- [86] Z.-S. Zhang. Berry–Esseen bounds for generalized U -statistics. *Electronic Journal of Probability*, 27:1–36, 2022.

APPENDIX A. PROOF OF THEOREM 2.1

For a graphon W and a non-empty simple graph $H = (V(H), E(H))$, the number of copies $X(H, G_n)$ (recall (1.1)) can be expressed as a generalized U -statistic as follows:

$$\begin{aligned} X(H, G_n) &= \frac{\binom{n}{|V(H)|}}{|\text{Aut}(H)|} t(H, W) \\ &= \sum_{1 \leq i_1 < \dots < i_{|V(H)|} \leq n} f^{(H)}(U_{i_1}, \dots, U_{i_{|V(H)|}}, Y_{i_1 i_2} \dots, Y_{i_{|V(H)|-1} i_{|V(H)|}}), \end{aligned} \quad (\text{A.1})$$

where

$$\begin{aligned} f^{(H)}(U_1, \dots, U_{|V(H)|}, Y_{12}, \dots, Y_{|V(H)|-1, |V(H)|}) \\ = \sum_{H' \in \mathcal{G}_H} \prod_{(a,b) \in E(H')} \mathbf{1}\{Y_{ab} \leq W(U_a, U_b)\} - |\mathcal{G}_H| t(H, W) \end{aligned} \quad (\text{A.2})$$

and $\mathcal{G}_H := \mathcal{G}_H(\{1, 2, \dots, |V(H)|\})$. In this section, using the representation in (A.1), we will derive the joint distribution of

$$\mathbf{X}(\mathcal{H}, G_n) := (X(H_1, G_n), X(H_2, G_n), \dots, X(H_r, G_n)),$$

for a collection non-empty simple graphs $\mathcal{H} := \{H_1, H_2, \dots, H_r\}$, where $H_i = (V(H_i), E(H_i))$ and $V(H_i) = \{1, 2, \dots, |V(H_i)|\}$, for $1 \leq i \leq r$.

We begin recalling the framework of generalized U -statistics developed in [45] in the following section.

A.1. Orthogonal Decomposition of Generalized U -Statistics. Suppose $\{U_i : 1 \leq i \leq n\}$ and $\{Y_{ij} : 1 \leq i < j \leq n\}$ are i.i.d. sequences of $U[0, 1]$ random variables. Fix $R \geq 1$ and denote by K_R the complete graph on the set of vertices $\{1, 2, \dots, R\}$ and let $G = (V(G), E(G))$ be a subgraph of K_R . Let $\mathcal{F}_G$ be the σ -algebra generated by the collections $\{U_i\}_{i \in V(G)}$ and $\{Y_{ij}\}_{(i,j) \in E(G)}$, and let $L^2(G) = L^2(\mathcal{F}_G)$ be the space of all square integrable random variables that are functions of $\{U_i : i \in V(G)\}$ and $\{Y_{ij} : (i, j) \in E(G)\}$. Now, consider the following subspace of $L^2(G)$:

$$M_G := \{Z \in L^2(G) : \mathbb{E}[ZV] = 0 \text{ for every } V \in L^2(F) \text{ such that } F \subset G\}. \quad (\text{A.3})$$

(For the empty graph, $M_\emptyset$ is the space of all constants.) Equivalently, $Z \in M_G$ if and only if $Z \in L^2(G)$ and

$$\mathbb{E}[Z \mid X_i, Y_{ij} : i \in V(F), (i, j) \in E(F)] = 0, \quad \text{for all } F \subset G.$$

Then, one has the following orthogonal decomposition (see [45, Lemma 1])

$$L^2(G) = \bigoplus_{F \subseteq G} M_F, \quad (\text{A.4})$$

that is, $L^2(G)$ is the orthogonal direct sum of M_F for all subgraphs $F \subseteq G$. This allows us to decompose any function in $L^2(G)$ as the sum of its projections onto M_F for $F \subseteq G$. For any closed subspace M of $L^2(K_R)$, denote the orthogonal projection onto M by P_M .

Now, consider a *symmetric* function f defined on $L^2(K_R)$, that is,

$$\begin{aligned} f &= f(U_1, U_2, \dots, U_R, Y_{12}, \dots, Y_{R-1 \ R}) \\ &= f(U_{\sigma(1)}, U_{\sigma(2)}, \dots, U_{\sigma(R)}, Y_{\sigma(1)\sigma(2)}, \dots, Y_{\sigma(R-1) \ \sigma(R)}). \end{aligned} \quad (\text{A.5})$$

for any permutation σ of $\{1, 2, \dots, R\}$.

Then f can be decomposed as

$$f = \sum_{G \subseteq K_R} f_G, \quad (\text{A.6})$$

where $f_G = P_{M_G} f$ is the orthogonal projection of f onto M_G . Further, for $1 \leq s \leq R$, define

$$f_{(s)} := \sum_{G \subseteq K_R : |V(G)|=s} f_G. \quad (\text{A.7})$$

The smallest positive d such that $f_{(d)} \neq 0$ almost surely is called the *principal degree* of f . It is easy to observe that for any $G \subseteq K_R$,

$$P_{L^2(G)} = \mathbb{E}[\cdot | \mathcal{F}_G]. \quad (\text{A.8})$$

Moreover, by (A.4) we have,

$$P_{L^2(G)} = \sum_{F \subseteq G} P_{M_F}. \quad (\text{A.9})$$

For $f \in L^2(K_R)$ define

$$S_{n,R}(f) = \sum_{1 \leq i_1 < i_2 < \dots < i_R \leq n} f(U_{i_1}, U_{i_2}, \dots, U_{i_R}, Y_{i_1 i_2}, \dots, Y_{i_{R-1} \ i_R}), \quad (\text{A.10})$$

and the symmetrized version

$$\tilde{S}_{n,R}(f) = \sum_{1 \leq i_1 \neq i_2 \neq \dots \neq i_R \leq n} f(U_{i_1}, U_{i_2}, \dots, U_{i_R}, Y_{i_1 i_2}, \dots, Y_{i_{R-1} \ i_R}), \quad (\text{A.11})$$

where $Y_{ji} := Y_{ij}$ for $1 \leq i < j \leq n$. The symmetry of f , the decomposition (A.6), and the linearity of $S_n^*(\cdot)$ implies that

$$S_{n,R}(f) = \frac{1}{R!} \tilde{S}_{n,R}(f) = \frac{1}{R!} \sum_{G \subseteq K_R} \tilde{S}_{n,R}(f_G). \quad (\text{A.12})$$

The symmetry of f also implies that if G_1 and G_2 are isomorphic subgraphs of K_R , then $\tilde{S}_{n,R}(f_{G_1}) = \tilde{S}_{n,R}(f_{G_2})$. Hence, from (A.12),

$$S_{n,R}(f) = \sum_{s=0}^R \sum_{G \in \Gamma_s} \frac{\tilde{S}_{n,R}(f_G)}{(R-s)! |\text{Aut}(G)|}, \quad (\text{A.13})$$

where Γ_s is the collection of non-isomorphic graphs with s vertices.

The following result from [45] gives the leading order in the expansion (A.13) for symmetric functions f with principal degree d . We include the proof for the sake of completeness:

Proposition A.1 ([45]). *Suppose $f \in L^2(K_R)$ is symmetric and has principal degree d . Then*

$$\mathbb{E} \left[\left| S_{n,R}(f) - \sum_{G \in \Gamma_d} \frac{\tilde{S}_{n,R}(f_G)}{(R-d)! |\text{Aut}(G)|} \right|^2 \right] = O(n^{2R-d-1}). \quad (\text{A.14})$$

Proof. Since f has principal degree d , by (A.13) and (A.7),

$$S_{n,R}(f) = \sum_{s=d}^R \sum_{G \in \Gamma_s} \frac{\tilde{S}_{n,R}(f_G)}{(R-s)! |\text{Aut}(G)|}.$$

Hence,

$$\mathbb{E} \left[\left| S_{n,R}(f) - \sum_{G \in \Gamma_d} \frac{\tilde{S}_{n,R}(f_G)}{(R-d)! |\text{Aut}(G)|} \right|^2 \right] \lesssim_R \sum_{s=d+1}^R \sum_{G \in \Gamma_s} \frac{\mathbb{E}[|\tilde{S}_{n,R}(f_G)|^2]}{(R-s)!^2 |\text{Aut}(G)|^2}. \quad (\text{A.15})$$

By [45, Lemma 4], for $G \in \Gamma_s$,

$$\mathbb{E}[|\tilde{S}_{n,R}(f_G)|^2] = \frac{n!(n-s)!}{(n-R)!^2} \mathbb{E}[f_G^2] \lesssim_R n^{2R-s} \mathbb{E}[f^2],$$

where the second inequality uses $\mathbb{E}[f_G^2] \leq \mathbb{E}[f^2]$ (recall the orthogonal decomposition (A.6)). Applying the above bound in (A.15), the result in (A.17) follows. $\square$

Using the above framework, we now proceed with the proof of Theorem 2.1. The proof is organized as follows:

- In Section A.2 we show that the asymptotic distribution of $\mathbf{Z}(\mathcal{H}, G_n)$ can be expressed as infinite linear or quadratic forms in i.i.d. Gaussian variables depending on whether H is W -irregular or W -regular, respectively (see Proposition A.2).
- In Section A.3 we identify the limit obtained in Proposition A.2 with limit in Theorem 2.1 (see Proposition A.3 and Proposition A.4).

A.2. Asymptotic Expansion of $\mathbf{Z}(\mathcal{H}, G_n)$. Recall that $\mathcal{H} = \{H_1, H_2, \dots, H_r\}$ is a collection of non-empty subgraphs such that W is H_i -irregular for $1 \leq i \leq q$ and W is H_i -regular for $q+1 \leq i \leq r$. Let $f_i := f^{(H_i)}$ denote the function defined in (A.2) with H replaced by H_i , for $1 \leq i \leq r$. It follows from [8, Lemma 5.4], that

$$\text{Var}[(f_i)_{(1)}] = 0 \text{ if and only if } W \text{ is } H_i\text{-regular.}$$

Since $\mathbb{E}[(f_i)_{(1)}] = 0$, this implies, $(f_i)_{(1)} = 0$ almost surely if and only if W is H_i -regular. Hence, the principal degree of f_i is 1 for $1 \leq i \leq q$ and the principal degree of f_i is 2 for $q+1 \leq i \leq r$. (Technically, it is possible that the principal degree of f_i , for some $q+1 \leq i \leq r$, is greater than 2 (see [8, Section 4.3] for an example). Here, we will assume that the principal degree of f_i is equal to 2 for all $q+1 \leq i \leq r$, with the understanding that the limit given by Theorem 2.1 in this case can be degenerate if the principal degree is larger.)

Note that, for $1 \leq i \leq r$, $f_i \in L^2(K_{|V(H_i)|})$, is symmetric and has 0 mean. In particular,

$$S_{n,|V(H_i)|}(f_i) = X(H_i, G_n) - \frac{(n)^{|V(H_i)|}}{|\text{Aut}(H_i)|} t(H_i, W), \quad (\text{A.16})$$

for $1 \leq i \leq r$. Now, to apply Proposition A.1 note that

$$\Gamma_1 = \{K_{\{1\}}\} \text{ and } \Gamma_2 = \{E_{\{1,2\}}, K_{\{1,2\}}\},$$

where

- $K_{\{1\}}$ is the graph with a single vertex 1,
- $E_{\{1,2\}} = (\{1,2\}, \emptyset)$ is the graph with two vertices $\{1,2\}$ and no edge between them,
- $K_{\{1,2\}} = (\{1,2\}, \{(1,2)\})$ is the complete graph on two vertices $\{1,2\}$.

Then by Proposition A.1 and Markov's inequality the following hold:

- For $1 \leq i \leq q$,

$$\left| S_{n,|V(H_i)|}(f_i) - \frac{\tilde{S}_{n,|V(H_i)|}((f_i)_{K_{\{1\}}})}{(|V(H_i)| - 1)!} \right| = O_P(n^{|V(H_i)|-1}). \quad (\text{A.17})$$

- For $q+1 \leq i \leq r$,

$$\left| S_{n,|V(H_i)|}(f_i) - \frac{\tilde{S}_{n,|V(H_i)|}((f_i)_{E_{\{1,2\}}}) + \tilde{S}_{n,|V(H_i)|}((f_i)_{K_{\{1,2\}}})}{2(|V(H_i)| - 2)!} \right| = O_P(n^{|V(H_i)|-\frac{3}{2}}). \quad (\text{A.18})$$

Now, define

$$T(H_i, G_n) := \begin{cases} \frac{\tilde{S}_{n,|V(H_i)|}((f_i)_{K_{\{1\}}})}{(|V(H_i)| - 1)! n^{|V(H_i)|-\frac{1}{2}}} & \text{for } 1 \leq i \leq q, \\ \frac{\tilde{S}_{n,|V(H_i)|}((f_i)_{E_{\{1,2\}}}) + \tilde{S}_{n,|V(H_i)|}((f_i)_{K_{\{1,2\}}})}{2(|V(H_i)| - 2)! n^{|V(H_i)|-1}} & \text{for } q+1 \leq i \leq r. \end{cases} \quad (\text{A.19})$$

Then recalling the definition of $Z(H_i, G_n)$ from (2.9) and using from (A.16), (A.17), and (A.18), it follows that

$$Z(H_i, G_n) = T(H_i, G_n) + o_P(1).$$

Hence, recalling (2.10)

$$\mathbf{Z}(\mathcal{H}, G_n) = \mathbf{T}(\mathcal{H}, G_n) + o_P(1), \quad (\text{A.20})$$

where

$$\mathbf{T}(\mathcal{H}, G_n) = (T(H_1, G_n), T(H_2, G_n), \dots, T(H_r, G_n))^\top. \quad (\text{A.21})$$

The result in (A.20) shows that to obtain the asymptotic joint distribution of $\mathbf{Z}(\mathcal{H}, G_n)$ it suffices to obtain the joint distribution of $\mathbf{T}(\mathcal{H}, G_n)$. The first step towards this is to compute the projections $(f_i)_{K_{\{1\}}}$, $(f_i)_{E_{\{1,2\}}}$, and $(f_i)_{K_{\{1,2\}}}$. We begin with a few definitions: A function $f \in L^2(G)$ is said to be G -symmetric if (A.5) holds whenever σ is an automorphism of G . Also, for two functions h_1, h_2 we define,

$$h_1 \otimes h_2(x, y) := h_1(x)h_2(y).$$

Now, let $\{\phi_s\}_{s \geq 1}$ be an orthonormal basis of $M_{K_{\{1\}}}$. Then $\{\phi_s \otimes \phi_t\}_{s,t \geq 1}$ is an orthonormal set whose span contains the $E_{\{1,2\}}$ -symmetric functions in $M_{E_{\{1,2\}}}$. Also, let $\{\psi_s\}_{s \geq 1}$ be an orthonormal basis of the subspace of $K_{\{1,2\}}$ -symmetric functions in $M_{K_{\{1,2\}}}$. Then using results in [8, 45] we have the following lemma:

Lemma A.1. *For $1 \leq i \leq q$, the projection of f_i on $K_{\{1\}}$ is given by:*

$$(f_i)_{K_{\{1\}}} = \mathbb{E}[f_i|U_1] = \sum_{s \geq 1} \mathbb{E}[f_i \phi_s] \phi_s. \quad (\text{A.22})$$

Moreover, for $q+1 \leq i \leq r$, the following hold:

- The projection of f_i on $E_{\{1,2\}}$ is given by:

$$(f_i)_{E_{\{1,2\}}} = \mathbb{E}[f_i|U_1, U_2] = \sum_{s,t \geq 1} \mathbb{E}[f_i(\phi_s \otimes \phi_t)] \phi_s \otimes \phi_t, \quad (\text{A.23})$$

- The projection of f_i on $K_{\{1,2\}}$ is given by:

$$(f_i)_{K_{\{1,2\}}} = \mathbb{E}[f_i|U_1, U_2, Y_{12}] - \mathbb{E}[f_i|U_1, U_2] = \sum_{s \geq 1} \mathbb{E}[f_i \psi_s] \psi_s. \quad (\text{A.24})$$

Proof. For $1 \leq i \leq q$, by (A.8) and (A.9),

$$(f_i)_{K_{\{1\}}} := P_{M_{K_{\{1\}}}}(f_i)_{K_{\{1\}}} = P_{L^2(K_{\{1\}})} f_i - P_{M_{\emptyset}} f_i = \mathbb{E}[f_i|U_1] - \mathbb{E}[f_i] = \mathbb{E}[f_i|U_1].$$

This proves the first equality in (A.22). To establish the second equality, expand $(f_i)_{K_{\{1\}}}$ in the basis $\{\phi_s\}_{s \geq 1}$ as follows (see [49, Chapter 6]):

$$(f_i)_{K_{\{1\}}} = \sum_{s \geq 1} \mathbb{E}[(f_i)_{K_{\{1\}}} \phi_s] \phi_s. \quad (\text{A.25})$$

By the first equality in (A.22),

$$\mathbb{E}[(f_i)_{K_{\{1\}}} \phi_s] = \mathbb{E}[(f_i)_{K_{\{1\}}}(U_1) \phi_s(U_1)] = \mathbb{E}[\mathbb{E}[f_i|U_1] \phi_s(U_1)] = \mathbb{E}[f_i \phi_s].$$

Applying the above identity in (A.25) the second equality in (A.22) follows.

Next, we will prove (A.23). By (A.8) and (A.9),

$$\begin{aligned} (f_i)_{E_{\{1,2\}}} &= P_{M_{E_{\{1,2\}}}} f_i = P_{L^2(E_{\{1,2\}})} f_i - P_{M_{K_{\{1\}}}} f_i - P_{M_{K_{\{2\}}}} f_i - P_{M_{\emptyset}} f_i \\ &= P_{L^2(E_{\{1,2\}})} f_i - P_{L^2(K_{\{1\}})} f_i - P_{L^2(K_{\{2\}})} f_i + P_{M_{\emptyset}} f_i \\ &= \mathbb{E}[f_i|U_1, U_2] - \mathbb{E}[f_i|U_1] - \mathbb{E}[f_i|U_2] + \mathbb{E}[f_i] \\ &= \mathbb{E}[f_i|U_1, U_2], \end{aligned} \quad (\text{A.26})$$

where the last step follows by noting that $\mathbb{E}[f_i] = 0$ and $\mathbb{E}[f_i|U_j] = 0$, for $j \in \{1, 2\}$, since W is H_i -regular, for $q+1 \leq i \leq r$. This proves the first equality in (A.23). To establish the second equality in (A.23), note that

$$\mathbb{E}[(f_i)_{E_{\{1,2\}}}(\phi_s \otimes \phi_t)] = \mathbb{E}[\mathbb{E}[f_i|U_1, U_2] \phi_s(U_1) \phi_t(U_2)] = \mathbb{E}[f_i(\phi_s \otimes \phi_t)].$$

Hence,

$$(f_i)_{E_{\{1,2\}}} = \sum_{s,t \geq 1} \mathbb{E}[(f_i)_{E_{\{1,2\}}}(\phi_s \otimes \phi_t)](\phi_s \otimes \phi_t) = \sum_{s,t \geq 1} \mathbb{E}[f_i(\phi_s \otimes \phi_t)](\phi_s \otimes \phi_t).$$

Finally, we prove (A.24). For this note that

$$\begin{aligned} (f_i)_{K_{\{1,2\}}} &= P_{M_{K_{\{1,2\}}}} f_i = P_{L^2(K_{\{1,2\}})} f_i - P_{M_{E_{\{1,2\}}}} f_i - P_{M_{K_{\{1\}}}} f_i - P_{M_{K_{\{2\}}}} f_i - P_{M_{\emptyset}} f_i \\ &= P_{L^2(K_{\{1,2\}})} f_i - P_{L^2(E_{\{1,2\}})} f_i \\ &= \mathbb{E}[f_i|U_1, U_2, Y_{12}] - \mathbb{E}[f_i|U_1, U_2]. \end{aligned} \quad (\text{A.27})$$

This proves the first equality in (A.24). For the second inequality note that $\mathbb{E}[\mathbb{E}[f_i|U_1, U_2]\psi_s] = 0$, since $\mathbb{E}[f_i|U_1, U_2] \in M_{E_{\{1,2\}}}$ and $M_{E_{\{1,2\}}}$ is orthogonal to $M_{K_{\{1,2\}}}$. This implies,

$$\mathbb{E}[(f_i)_{K_{\{1,2\}}} \psi_s] = \mathbb{E}[\mathbb{E}[f_i|U_1, U_2, Y_{12}] \psi_s] = \mathbb{E}[f_i \psi_s]$$

and, hence,

$$(f_i)_{K_{\{1,2\}}} = \sum_{s \geq 1} \mathbb{E}[(f_i)_{K_{\{1,2\}}} \psi_s] \psi_s = \sum_{s \geq 1} \mathbb{E}[f_i \psi_s] \psi_s. \quad (\text{A.28})$$

This completes the proof of (A.24). $\square$

Lemma A.1 and [49, Chapter 6, Lemma 8] we now can compute the L^2 norms of the projections, denoted by $\|\cdot\|_2$, as follows: For $1 \leq i \leq q$,

$$\|(f_i)_{K_{\{1\}}}\|_2^2 = \sum_{s \geq 1} \mathbb{E}[f_i \phi_s]^2 \leq \|f_i\|_2^2, \quad (\text{A.29})$$

where the last step follows by Bessel's Inequality. Similarly, for $q+1 \leq i \leq r$,

$$\|(f_i)_{E_{\{1,2\}}}\|_2^2 = \sum_{s,t \geq 1} \mathbb{E}[f_i(\phi_s \otimes \phi_t)]^2 \leq \|f_i\|_2^2 \text{ and } \|(f_i)_{K_{\{1,2\}}}\|_2^2 = \sum_{s \geq 1} \mathbb{E}[f_i \psi_s]^2 \leq \|f_i\|_2^2. \quad (\text{A.30})$$

Next, using Lemma A.1, the linearity of $\tilde{S}_{n,(\cdot)}$, and a standard truncation argument we obtain the expansions of $\tilde{S}_{n,|V(H_i)|}((f_i)_{K_{\{1\}}})$, $\tilde{S}_{n,|V(H_i)|}((f_i)_{E_{\{1,2\}}})$, and $\tilde{S}_{n,|V(H_i)|}((f_i)_{K_{\{1,2\}}})$. Specifically, for $1 \leq i \leq q$, from (A.22) we have

$$\tilde{S}_{n,|V(H_i)|}((f_i)_{K_{\{1\}}}) = \sum_{s \geq 1} \mathbb{E}[f_i \phi_s] \tilde{S}_{n,|V(H_i)|}(\phi_s), \quad (\text{A.31})$$

for $1 \leq i \leq q$, where the equality hold almost surely. Similarly, for $q+1 \leq i \leq r$, by (A.23) and (A.24) we have the following:

$$\begin{aligned} \tilde{S}_{n,|V(H_i)|}((f_i)_{E_{\{1,2\}}}) &= \sum_{s \geq 1} \mathbb{E}[f_i(\phi_s \otimes \phi_s)] \tilde{S}_{n,|V(H_i)|}(\phi_s \otimes \phi_s) \\ &\quad + 2 \sum_{s < t} \mathbb{E}[f_i(\phi_s \otimes \phi_t)] \tilde{S}_{n,|V(H_i)|}(\phi_s \otimes \phi_t), \end{aligned} \quad (\text{A.32})$$

$$\tilde{S}_{n,|V(H_i)|}((f_i)_{K_{\{1,2\}}}) = \sum_{s \geq 1} \mathbb{E}[f_i \psi_s] \tilde{S}_{n,|V(H_i)|}(\psi_s). \quad (\text{A.33})$$

Using the above expansions we can now derive the asymptotic distribution of $\mathbf{T}(\mathcal{H}, G_n)$ and hence, that of $\mathbf{Z}(\mathcal{H}, G_n)$.

Proposition A.2. *For $\mathbf{T}(\mathcal{H}, G_n)$ as defined in (A.21) and (A.19), the following hold as $n \rightarrow \infty$:*

$$\mathbf{T}(\mathcal{H}, G_n) \xrightarrow{D} \mathbf{T}(\mathcal{H}, W),$$

where $\mathbf{T}(\mathcal{H}, W) = (T_1, T_2, \dots, T_r)^\top$ with

$$T_i = \begin{cases} \frac{1}{(|V(H_i)| - 1)!} \sum_{s \geq 1} \mathbb{E}[f_i \phi_s] \eta_s & \text{for } 1 \leq i \leq q, \\ \frac{1}{2(|V(H_i)| - 2)!} \left\{ \sum_{s \geq 1} \mathbb{E}[f_i(\phi_s \otimes \phi_s)] (\eta_s^2 - 1) + 2 \sum_{s < t} \mathbb{E}[f_i(\phi_s \otimes \phi_t)] \eta_s \eta_t \right. \\ \quad \left. + \sum_{s \geq 1} \mathbb{E}[f_i \psi_s] \tilde{\eta}_s \right\} & \text{for } q+1 \leq i \leq r, \end{cases}$$

and $\{\eta_s\}_{s \geq 1}$ and $\{\tilde{\eta}_s\}_{s \geq 1}$ are independent collections of $N(0, 1)$ and $N(0, 2)$ random variables, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, respectively. This implies, from (A.20),

$$\mathbf{Z}(\mathcal{H}, G_n) \xrightarrow{D} \mathbf{T}(\mathcal{H}, W).$$

A.2.1. *Proof of Proposition A.2.* Fix $L \geq 1$ and define the truncated version of $\tilde{S}_{n, |V(H_i)|}((f_i)_{K_{\{1\}}})$ (recall (A.31)) as follows:

$$\tilde{S}_{n, |V(H_i)|}^{(L)}((f_i)_{K_{\{1\}}}) = \sum_{s=1}^L \mathbb{E}[f_i \phi_s] \tilde{S}_{n, |V(H_i)|}(\phi_s),$$

for $1 \leq i \leq q$. Similarly, for $q+1 \leq i \leq r$, recalling (A.32) and (A.33) define the truncated versions:

$$\begin{aligned} \tilde{S}_{n, |V(H_i)|}^{(L)}((f_i)_{E_{\{1,2\}}}) &= \sum_{s=1}^L \mathbb{E}[f_i(\phi_s \otimes \phi_s)] \tilde{S}_{n, |V(H_i)|}(\phi_s \otimes \phi_s) \\ &\quad + 2 \sum_{1 \leq s < t \leq L} \mathbb{E}[f_i(\phi_s \otimes \phi_t)] \tilde{S}_{n, |V(H_i)|}(\phi_s \otimes \phi_t), \\ \tilde{S}_{n, |V(H_i)|}^{(L)}((f_i)_{K_{\{1,2\}}}) &= \sum_{s=1}^L \mathbb{E}[f_i \psi_s] \tilde{S}_{n, |V(H_i)|}(\psi_s). \end{aligned}$$

and for $q+1 \leq i \leq r$. Now, recalling (A.19), define the truncated version of $T(H_i, G_n)$ as follows:

$$T^{(L)}(H_i, G_n) := \begin{cases} \frac{\tilde{S}_{n, |V(H_i)|}^{(L)}((f_i)_{K_{\{1\}}})}{(|V(H_i)| - 1)! |n|^{|V(H_i)| - \frac{1}{2}}} & \text{for } 1 \leq i \leq q, \\ \frac{\tilde{S}_{n, |V(H_i)|}^{(L)}((f_i)_{E_{\{1,2\}}}) + \tilde{S}_{n, |V(H_i)|}^{(L)}((f_i)_{K_{\{1,2\}}})}{2(|V(H_i)| - 2)! |n|^{|V(H_i)| - 1}} & \text{for } q+1 \leq i \leq r, \end{cases} \quad (\text{A.34})$$

and

$$\mathbf{T}^{(L)}(\mathcal{H}, G_n) = (T^{(L)}(H_1, G_n), T^{(L)}(H_2, G_n), \dots, T^{(L)}(H_r, G_n))^\top. \quad (\text{A.35})$$

The following lemma shows that $\mathbf{T}^{(L)}(\mathcal{H}, G_n)$ converges to a truncated version of $\mathbf{T}(\mathcal{H}, W)$.

Lemma A.2. Fix $L \geq 1$ and let $\mathbf{T}^{(L)}(\mathcal{H}, G_n)$ be as defined in (A.35). Then the following hold as $n \rightarrow \infty$:

$$\mathbf{T}^{(L)}(\mathcal{H}, G_n) \xrightarrow{D} \mathbf{T}^{(L)}(\mathcal{H}, W),$$

where $\mathbf{T}^{(L)}(\mathcal{H}, W) = (T_1^{(L)}, T_2^{(L)}, \dots, T_r^{(L)})^\top$ with

$$T_i^{(L)} = \begin{cases} \frac{1}{(|V(H_i)| - 1)!} \sum_{s=1}^L \mathbb{E}[f_i \phi_s] \eta_s & \text{for } 1 \leq i \leq q, \\ \frac{1}{2(|V(H_i)| - 2)!} \left\{ \sum_{s=1}^L \mathbb{E}[f_i(\phi_s \otimes \phi_s)] (\eta_s^2 - 1) + 2 \sum_{1 \leq s < t \leq L} \mathbb{E}[f_i(\phi_s \otimes \phi_t)] \eta_s \eta_t \right. \\ \quad \left. + \sum_{s=1}^L \mathbb{E}[f_i \psi_s] \tilde{\eta}_s \right\} & \text{for } q+1 \leq i \leq r, \end{cases}$$

where $\{\eta_s\}_{s \geq 1}$ and $\{\tilde{\eta}_s\}_{s \geq 1}$ are independent collections of $N(0, 1)$ and $N(0, 2)$ random variables, respectively.

Proof. Recalling the definition of $\tilde{S}_{n,\cdot}(\cdot)$ from (A.11) we get

$$\frac{1}{n^{|V(H_i)|-\frac{1}{2}}} \tilde{S}_{n,|V(H_i)|}(\phi_s) = \frac{1+o(1)}{\sqrt{n}} \sum_{i=1}^n \phi_s(U_i), \quad (\text{A.36})$$

for $s \geq 1$ and $1 \leq i \leq q$. Similarly, for $s, t \geq 1$ and $q+1 \leq i \leq r$,

$$\frac{1}{n^{|V(H_i)|-1}} \tilde{S}_{n,|V(H_i)|}(\phi_s \otimes \phi_t) = \frac{1+o(1)}{n} \sum_{1 \leq i \neq j \leq n} \phi_s(U_i) \phi_t(U_j), \quad (\text{A.37})$$

and

$$\frac{1}{n^{|V(H_i)|-1}} \tilde{S}_{n,|V(H_i)|}(\psi_s) = \frac{1+o(1)}{n} \sum_{1 \leq i \neq j \leq n} \psi_s(U_i, U_j, Y_{ij}). \quad (\text{A.38})$$

Now, by [45, Lemma 8] the collection,

$$\left\{ \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^n \phi_s(U_i) \right\}_{s=1}^L, \left\{ \frac{1}{n} \sum_{1 \leq i \neq j \leq n} \phi_s(U_i) \phi_t(U_j) \right\}_{s=1}^L, \left\{ \frac{1}{n} \sum_{1 \leq i \neq j \leq n} \psi_s(U_i, U_j, Y_{ij}) \right\}_{s=1}^L \right\}$$

converges jointly to

$$\left\{ \{\eta_s\}_{1 \leq s \leq L}, \{\eta_s \eta_t - \mathbb{E}[\eta_s \eta_t]\}_{1 \leq s, t \leq L}, \{\tilde{\eta}_s\}_{1 \leq s \leq L} \right\}.$$

The result in Lemma A.2 then follows by recalling the definition of $\mathbf{T}^{(L)}(\mathcal{H}, G_n)$ from (A.35) and (A.34) and the decompositions from (A.36), (A.37) and (A.38). $\square$

Now, to complete the proof of Proposition A.2 it suffices to show the following:

- (1) $\mathbf{T}^{(L)}(\mathcal{H}, G_n)$ and $\mathbf{T}(\mathcal{H}, G_n)$ are asymptotically close and
- (2) $\mathbf{T}^{(L)}(\mathcal{H}, W)$ converges to $\mathbf{T}(\mathcal{H}, W)$, as $L \rightarrow \infty$.

These are established in the following 2 lemmas, respectively.

Lemma A.3. *Let $\mathbf{T}(\mathcal{H}, G_n)$ and $\mathbf{T}^{(L)}(\mathcal{H}, G_n)$ be defined in (A.21) and (A.35), respectively. Then*

$$\lim_{L \rightarrow \infty} \sup_{n \geq 1} \mathbb{E} \left[\left\| \mathbf{T}(\mathcal{H}, G_n) - \mathbf{T}^{(L)}(\mathcal{H}, G_n) \right\|_2^2 \right] = 0.$$

Proof. Note that for $1 \leq i \leq q$,

$$\mathbb{E} \left[|\mathbf{T}(H_i, G_n) - \mathbf{T}^{(L)}(H_i, G_n)|^2 \right] = \frac{1}{n^{2|V(H_i)|-2}} \mathbb{E} \left[\left| \sum_{s=L+1}^N \mathbb{E}[f_i \phi_s] \tilde{S}_{n,|V(H_i)|}(\phi_s) \right|^2 \right]. \quad (\text{A.39})$$

By the orthogonality of the basis $\{\phi_s\}_{s \geq 1}$, it is easy to verify that $\mathbb{E}[\tilde{S}_{n,\cdot}(\phi_s) \tilde{S}_{n,\cdot}(\phi_t)] = 0$, for $s \neq t$. Moreover, from (A.36), it follows that $\frac{1}{n^{2|V(H_i)|-2}} \mathbb{E}[\tilde{S}_{n,|V(H_i)|}(\phi_s)^2] = 1 + o(1)$, for $s \geq 1$. Hence, from (A.39) and (A.29),

$$\lim_{L \rightarrow \infty} \sup_{n \geq 1} \mathbb{E} \left[|\mathbf{T}(H_i, G_n) - \mathbf{T}^{(L)}(H_i, G_n)|^2 \right] \lesssim_{H_i} \lim_{L \rightarrow \infty} \sum_{m=L+1}^{\infty} \mathbb{E}[f_i \phi_s]^2 = 0. \quad (\text{A.40})$$

Similarly, for $q+1 \leq i \leq r$ it can be shown that

$$\lim_{L \rightarrow \infty} \sup_{n \geq 1} \mathbb{E} \left[|\mathbf{T}(H_i, G_n) - \mathbf{T}^{(L)}(H_i, G_n)|^2 \right]$$

$$\begin{aligned} & \lesssim_{H_i} \lim_{L \rightarrow \infty} \left\{ \sum_{s=L+1}^{\infty} \mathbb{E}[f_i(\phi_s \otimes \phi_s)]^2 + \sum_{t=L+1}^{\infty} \sum_{1 \leq s < t} \mathbb{E}[f_i(\phi_s \otimes \phi_t)]^2 + \sum_{s=L+1}^{\infty} \mathbb{E}[f_i \psi_s]^2 \right\} \\ & = 0, \end{aligned} \quad (\text{A.41})$$

by (A.30).

Combining (A.40) and (A.41) the proof of Lemma A.3 follows. $\square$

Lemma A.4. *Let $\mathbf{T}(\mathcal{H}, W) = (T_1, T_2, \dots, T_r)^\top$ and $\mathbf{T}^{(L)}(\mathcal{H}, W) = (T_1^{(L)}, T_2^{(L)}, \dots, T_r^{(L)})^\top$ be as defined in Proposition A.2 and Lemma A.2, respectively. Then*

$$\lim_{L \rightarrow \infty} \mathbb{E} \left[\|\mathbf{T}^{(L)}(\mathcal{H}, W) - \mathbf{T}(\mathcal{H}, W)\|_2^2 \right] = 0.$$

Proof. For $1 \leq i \leq q$, by [49, Lemma 8] we get,

$$\mathbb{E}|T_i^{(L)} - T_i| \lesssim_{H_i} \mathbb{E} \left| \sum_{s=L+1}^{\infty} \mathbb{E}[f_i \psi_s] \eta_s \right|^2 \leq \sum_{s=L+1}^{\infty} \mathbb{E}[f_i \psi_s]^2 \rightarrow 0, \quad (\text{A.42})$$

as $L \rightarrow \infty$, by (A.29). Similarly, for $q+1 \leq i \leq r$,

$$\mathbb{E} \left| \sum_{s=L+1}^{\infty} \mathbb{E}[f_i \phi_s] \tilde{\eta}_s \right|^2 \rightarrow 0, \quad (\text{A.43})$$

as $L \rightarrow \infty$, by (A.30). Also, since $\{\eta_s^2 - 1 : s \geq 1\}$ are orthogonal and $\mathbb{E}(\eta_s^2 - 1)^2 = 2$,

$$\mathbb{E} \left| \sum_{s=L+1}^{\infty} \mathbb{E}[f_i(\phi_s \otimes \phi_s)] (\eta_s^2 - 1) \right|^2 \leq 2 \sum_{s=L+1}^{\infty} \mathbb{E}[f_i(\phi_s \otimes \phi_s)]^2 \rightarrow 0. \quad (\text{A.44})$$

Once again by definition $\{\eta_s \eta_t : s < t\}$ are orthogonal and $\mathbb{E} \eta_s^2 \eta_t^2 = 1$. Hence, as above,

$$\mathbb{E} \left| \sum_{t=1}^{\infty} \sum_{1 \leq s < t} \mathbb{E}[f_i(\phi_s \otimes \phi_t)] \eta_s \eta_t - \sum_{t=L+1}^{\infty} \sum_{1 \leq s < t} \mathbb{E}[f_i(\phi_s \otimes \phi_t)] \eta_s \eta_t \right| \rightarrow 0, \quad (\text{A.45})$$

as $L \rightarrow \infty$. Combining (A.43), (A.44), and (A.45), we get $\mathbb{E}|T_i^{(L)} - T_i|^2 \rightarrow 0$, as $L \rightarrow \infty$, for $q+1 \leq i \leq r$. This together with (A.42) completes the proof of Lemma A.4. $\square$

Combining Lemma A.2, Lemma A.3, and Lemma A.4 along with [45, Lemma 6] the result in Proposition A.2 follows. $\square$

A.3. Equivalence of $\mathbf{T}(\mathcal{H}, W)$ and $\mathbf{Z}(\mathcal{H}, W)$. For $1 \leq i \leq q$, define

$$Q_i = \frac{1}{(|V(H_i)| - 1)!} \sum_{s \geq 1} \mathbb{E}[f_i \phi_s] \eta_s. \quad (\text{A.46})$$

Also, for $q+1 \leq i \leq r$, define

$$R_i = \frac{1}{2(|V(H_i)| - 2)!} \left\{ \sum_{s \geq 1} \mathbb{E}[f_i(\phi_s \otimes \phi_s)] (\eta_s^2 - 1) + 2 \sum_{s < t} \mathbb{E}[f_i(\phi_s \otimes \phi_t)] \eta_s \eta_t \right\}, \quad (\text{A.47})$$

$$\tilde{R}_i = \frac{1}{2(|V(H_i)| - 2)!} \sum_{s \geq 1} \mathbb{E}[f_i \psi_s] \tilde{\eta}_s, \quad (\text{A.48})$$

where $\{\eta_s\}_{s \geq 1}$ and $\{\tilde{\eta}_s\}_{s \geq 1}$ are independent collections of $N(0, 1)$ and $N(0, 2)$ random variables, respectively. Then recalling the definition of T_i from Proposition A.2, note that

$$T_i = \begin{cases} Q_i & \text{for } 1 \leq i \leq q, \\ R_i + \tilde{R}_i & \text{for } q+1 \leq i \leq r. \end{cases}$$

Denote $\mathbf{Q} = (Q_1, Q_2, \dots, Q_q)^\top$, $\mathbf{R} = (R_{q+1}, R_{q+2}, \dots, R_r)^\top$ and $\tilde{\mathbf{R}} = (\tilde{R}_{q+1}, \tilde{R}_{q+2}, \dots, \tilde{R}_r)^\top$. In the following 2 propositions we identify $\mathbf{Q}$, $\mathbf{R}$, $\tilde{\mathbf{R}}$ with their corresponding components in $\mathbf{Z}(\mathcal{H}, W)$ (as defined in Theorem 2.1).

Proposition A.3. *Let $\tilde{\mathbf{R}} = (\tilde{R}_{q+1}, \tilde{R}_{q+2}, \dots, \tilde{R}_r)^\top$ be as defined in (A.48). Then the following hold:*

$$\tilde{\mathbf{R}} \sim N_{r-q}(\mathbf{0}, \Sigma),$$

where Σ is as defined in Theorem 2.1.

Proposition A.4. *Let $\mathbf{Q} = (Q_1, Q_2, \dots, Q_q)^\top$ and $\mathbf{R} = (R_{q+1}, R_{q+2}, \dots, R_r)^\top$ be as defined in (A.46) and (A.48), respectively. Suppose $\{B_t\}_{t \in [0,1]}$ be the standard Brownian motion in $[0, 1]$. Then the following hold:*

$$\mathbf{Q} \stackrel{D}{=} \left(\frac{1}{|\text{Aut}(H_i)|} \int_0^1 \left\{ \sum_{a=1}^{|V(H_i)|} (t_a(x, H_i, W) - t(H_i, W)) \right\} dB_x \right)_{1 \leq i \leq q}, \quad (\text{A.49})$$

and

$$\mathbf{R} \stackrel{D}{=} \left(\int_0^1 \int_0^1 \left\{ W_{H_i}(x, y) - \frac{|V(H_i)|(|V(H_i) - 1|)}{2|\text{Aut}(H_i)|} t(H_i, W) \right\} dB_x dB_y \right)_{q+1 \leq i \leq r}. \quad (\text{A.50})$$

The proofs of Proposition A.3 and Proposition A.4 are given in Section A.3.1 and Section A.3.2, respectively. These 2 results combined establishes the equivalence of $\mathbf{T}(\mathcal{H}, W)$ and $\mathbf{Z}(\mathcal{H}, W)$.

A.3.1. *Proof of Proposition A.3.* Fix $L \geq 1$ and define the truncated version of $\tilde{R}_i$ as follows:

$$\tilde{R}_i^{(L)} = \frac{1}{2(|V(H_i)| - 2)!} \sum_{s=1}^L \mathbb{E}[f_i \psi_s] \tilde{\eta}_s,$$

for $q+1 \leq i \leq r$. Denote $\tilde{\mathbf{R}}^{(L)} = (\tilde{R}_{q+1}^{(L)}, \tilde{R}_{q+2}^{(L)}, \dots, \tilde{R}_r^{(L)})^\top$. Then from (A.30) it follows that

$$\lim_{L \rightarrow \infty} \mathbb{E} \left\| \tilde{\mathbf{R}} - \tilde{\mathbf{R}}^{(L)} \right\|_2^2 = 0. \quad (\text{A.51})$$

Since the collection $\{\eta_s\}_{s \geq 1}$ are independent $N(0, 2)$,

$$\tilde{\mathbf{R}}^{(L)} \stackrel{D}{=} N_{r-q}(0, \Gamma^{(L)}),$$

where $\Gamma^{(L)} = ((\gamma_{ij}^{(L)}))_{q+1 \leq i, j \leq r}$ is given by

$$\gamma_{ij}^{(L)} = \begin{cases} \frac{1}{2(|V(H_i)| - 2)!^2} \sum_{s=1}^L \mathbb{E}[f_i \psi_s]^2 & \text{for } q+1 \leq i = j \leq r, \\ \frac{1}{2(|V(H_i)| - 2)!(|V(H_j)| - 2)!} \sum_{s=1}^L \mathbb{E}[f_i \psi_s] \mathbb{E}[f_j \psi_s] & \text{for } q+1 \leq i \neq j \leq r. \end{cases}$$

By computing characteristic functions and recalling (A.30) one has $\tilde{\mathbf{R}}^{(L)} \xrightarrow{D} N_{r-q}(0, \Gamma)$, as $L \rightarrow \infty$, where $\Gamma = ((\gamma_{ij}))_{q+1 \leq i, j \leq r}$ is given by

$$\gamma_{ij} = \begin{cases} \frac{1}{2(|V(H_i)| - 2)!^2} \mathbb{E}[(f_i)_{K_{\{1,2\}}}^2] & \text{for } q+1 \leq i = j \leq r, \\ \frac{1}{2(|V(H_i)| - 2)!(|V(H_j)| - 2)!} \mathbb{E}[(f_i)_{K_{\{1,2\}}} (f_j)_{K_{\{1,2\}}}] & \text{for } q+1 \leq i \neq j \leq r. \end{cases}$$

Here, we use the identity $\sum_{s=1}^{\infty} \mathbb{E}[f_i \psi_s]^2 = \mathbb{E}[(f_i)_{K_{\{1,2\}}}^2]$ (recall (A.30)) and

$$\sum_{s=1}^{\infty} \mathbb{E}[f_i \psi_s] \mathbb{E}[f_j \psi_s] = \mathbb{E}[(f_i)_{K_{\{1,2\}}}(f_j)_{K_{\{1,2\}}}],$$

which follows from the expansion (A.28) and the orthogonality of the functions $\{\psi_s\}_{s \geq 1}$.

The proof of Proposition now follows from Lemma A.5, which shows that the matrix Γ is same as the matrix Σ in Theorem 2.1. $\square$

Lemma A.5. *For all $q+1 \leq i, j \leq r$,*

$$\mathbb{E}[(f_i)_{K_{\{1,2\}}} f_{j,K_{\{1,2\}}}] = \frac{(|V(H_i)| - 2)! (|V(H_j)| - 2)!}{|\text{Aut}(H_i)| |\text{Aut}(H_j)|} \sum_{\substack{(a,b) \in E^+(H_i) \\ (c,d) \in E^+(H_j)}} \left(t \left(H_i \ominus_{(a,b),(c,d)} H_j, W \right) - t \left(H_i \oplus_{(a,b),(c,d)} H_j, W \right) \right).$$

Proof. For $q+1 \leq i \leq r$, from Lemma A.1 we have

$$\begin{aligned} (f_i)_{K_{\{1,2\}}} &= \mathbb{E}[f_i | U_1, U_2, Y_{12}] - \mathbb{E}[f_i | U_1, U_2] \\ &= \sum_{H' \in \mathcal{G}_{H_i, \{1,2\}}} t_{1,2}^-(U_1, U_2, H', W) (\mathbf{1}\{Y_{12} \leq W(U_1, U_2)\} - W(U_1, U_2)). \end{aligned}$$

where $\mathcal{G}_{H_i, \{1,2\}} := \{H' \in \mathcal{G}_{H_i} : (1,2) \in E(H')\}$ for all $q+1 \leq i \leq r$ and,

$$t_{1,2}^-(U_1, U_2, H', W) = \mathbb{E} \left[\prod_{(i,j) \in E(H') \setminus \{(1,2)\}} W(U_i, U_j) \middle| U_1, U_2 \right] \text{ for all } H' \in \mathcal{G}_{H_i, \{1,2\}}.$$

Then for $q+1 \leq i, j \leq r$,

$$\begin{aligned} &\mathbb{E}[(f_i)_{K_{\{1,2\}}} f_{j,K_{\{1,2\}}}] \\ &= \sum_{\substack{H_1 \in \mathcal{G}_{H_i, \{1,2\}} \\ H_2 \in \mathcal{G}_{H_j, \{1,2\}}}} \mathbb{E} \left[t_{1,2}^-(U_1, U_2, H_1, W) t_{1,2}^-(U_1, U_2, H_2, W) (\mathbf{1}\{Y_{12} \leq W(U_1, U_2)\} - W(U_1, U_2))^2 \right] \\ &= \sum_{\substack{H_1 \in \mathcal{G}_{H_i, \{1,2\}} \\ H_2 \in \mathcal{G}_{H_j, \{1,2\}}}} \mathbb{E} \left[t_{1,2}^-(U_1, U_2, H_1, W) t_{1,2}^-(U_1, U_2, H_2, W) W(U_1, U_2) (1 - W(U_1, U_2)) \right] \\ &= \sum_{\substack{H_1 \in \mathcal{G}_{H_i, \{1,2\}} \\ H_2 \in \mathcal{G}_{H_j, \{1,2\}}}} \left(t \left(H_1 \ominus_{(1,2),(1,2)} H_2, W \right) - t \left(H_1 \oplus_{(1,2),(1,2)} H_2, W \right) \right), \end{aligned} \tag{A.52}$$

recalling the join operations from Definition 2.3. Now, considering $V_{H_\ell}^2 = \{(a,b) \in V(H_\ell) : a \neq b\}$ for $\ell \in \{i, j\}$ define,

$$\underline{t} \left(H_1 \ominus_{(a,b),(c,d)} H_2, W \right) = t \left(H_1 \ominus_{(a,b),(c,d)} H_2, W \right) \mathbf{1}\{(a,b) \in E^+(H_1) \text{ and } (c,d) \in E^+(H_2)\}$$

and similarly,

$$\underline{t} \left(H_1 \bigoplus_{(a,b),(c,d)} H_2, W \right) = t \left(H_1 \bigoplus_{(a,b),(c,d)} H_2, W \right) \mathbf{1}\{(a,b) \in E^+(H_1) \text{ and } (c,d) \in E^+(H_2)\}.$$

where $(a,b) \in V_{H_i}^2$, $(c,d) \in V_{H_j}^2$ and $H_1 \in \mathcal{G}_{H_i}$ and $H_2 \in \mathcal{G}_{H_j}$. Then we can rewrite (A.52) as,

$$\mathbb{E} \left[(f_i)_{K_{\{1,2\}}} f_j, K_{\{1,2\}} \right] = \sum_{\substack{H_1 \in \mathcal{G}_{H_i} \\ H_2 \in \mathcal{G}_{H_j}}} \left(\underline{t} \left(H_1 \bigoplus_{(1,2),(1,2)} H_2, W \right) - \underline{t} \left(H_1 \bigoplus_{(1,2),(1,2)} H_2, W \right) \right) \quad (\text{A.53})$$

Now consider the permutations $\pi_{(a,b)} : V(H_i) \rightarrow V(H_i)$ and $\pi'_{(c,d)} : V(H_j) \rightarrow V(H_j)$ such that $\pi_{(a,b)}(a) = \pi'_{(c,d)}(c) = 1$ and $\pi_{(a,b)}(b) = \pi'_{(c,d)}(d) = 2$. Then

$$\begin{aligned} & \sum_{\substack{(a,b) \in V_{H_i}^2 \\ (c,d) \in V_{H_j}^2}} \sum_{\substack{H_1 \in \mathcal{G}_{H_i} \\ H_2 \in \mathcal{G}_{H_j}}} \underline{t} \left(H_1 \bigoplus_{(a,b),(c,d)} H_2, W \right) \\ &= \sum_{\substack{(a,b) \in V_{H_i}^2 \\ (c,d) \in V_{H_j}^2}} \sum_{\substack{H_1 \in \mathcal{G}_{H_i} \\ H_2 \in \mathcal{G}_{H_j}}} \underline{t} \left(\pi_{(a,b)}(H_1) \bigoplus_{(1,2),(1,2)} \pi'_{(c,d)}(H_2), W \right) \\ &= |V_{H_i}^2| |V_{H_j}^2| \sum_{\substack{H_1 \in \mathcal{G}_{H_i} \\ H_2 \in \mathcal{G}_{H_j}}} \underline{t} \left(H_1 \bigoplus_{(1,2),(1,2)} H_2, W \right) \end{aligned} \quad (\text{A.54})$$

where the last equality follows by observing that

$$(H_1, H_2) \rightarrow (\pi_{(a,b)}(H_1), \pi'_{(c,d)}(H_2))$$

is a bijection from $\mathcal{G}_{H_i} \times \mathcal{G}_{H_j}$ to itself for all $(a,b) \in V_{H_i}^2$ and $(c,d) \in V_{H_j}^2$. By considering isomorphisms τ_{H_1} and τ'_{H_1} for $H_1 \in \mathcal{G}_{H_i}$ and $H_2 \in \mathcal{G}_{H_j}$ such that $\tau_{H_1}(H_1) = H_i$ and $\tau'_{H_1}(H_2) = H_j$ a similar argument as above shows that,

$$\sum_{\substack{(a,b) \in V_{H_i}^2 \\ (c,d) \in V_{H_j}^2}} \sum_{\substack{H_1 \in \mathcal{G}_{H_i} \\ H_2 \in \mathcal{G}_{H_j}}} \underline{t} \left(H_1 \bigoplus_{(a,b),(c,d)} H_2, W \right) = |\mathcal{G}_{H_i}| |\mathcal{G}_{H_j}| \sum_{\substack{(a,b) \in V_{H_i}^2 \\ (c,d) \in V_{H_j}^2}} \underline{t} \left(H_i \bigoplus_{(a,b),(c,d)} H_j, W \right). \quad (\text{A.55})$$

Thus, combining (A.54) and (A.55) we find,

$$\sum_{\substack{H_1 \in \mathcal{G}_{H_i} \\ H_2 \in \mathcal{G}_{H_j}}} \underline{t} \left(H_1 \bigoplus_{(1,2),(1,2)} H_2, W \right) = \frac{|\mathcal{G}_{H_i}| |\mathcal{G}_{H_j}|}{|V_{H_i}^2| |V_{H_j}^2|} \sum_{\substack{(a,b) \in V_{H_i}^2 \\ (c,d) \in V_{H_j}^2}} \underline{t} \left(H_i \bigoplus_{(a,b),(c,d)} H_j, W \right). \quad (\text{A.56})$$

Similarly,

$$\sum_{\substack{H_1 \in \mathcal{G}_{H_i} \\ H_2 \in \mathcal{G}_{H_j}}} \underline{t} \left(H_1 \bigoplus_{(1,2),(1,2)} H_2, W \right) = \frac{|\mathcal{G}_{H_i}| |\mathcal{G}_{H_j}|}{|V_{H_i}^2| |V_{H_j}^2|} \sum_{\substack{(a,b) \in V_{H_i}^2 \\ (c,d) \in V_{H_j}^2}} \underline{t} \left(H_i \bigoplus_{(a,b),(c,d)} H_j, W \right). \quad (\text{A.57})$$

Notice that by definition,

$$\begin{aligned} & \sum_{\substack{(a,b) \in V_{H_i}^2 \\ (c,d) \in V_{H_j}^2}} \left(\underline{t} \left(H_i \ominus_{(a,b),(c,d)} H_j, W \right) - \underline{t} \left(H_i \oplus_{(a,b),(c,d)} H_j, W \right) \right) \\ &= \sum_{\substack{(a,b) \in E^+(H_i) \\ (c,d) \in E^+(H_j)}} \left(t \left(H_i \ominus_{(a,b),(c,d)} H_j, W \right) - t \left(H_i \oplus_{(a,b),(c,d)} H_j, W \right) \right). \end{aligned} \quad (\text{A.58})$$

Recall that $|\mathcal{G}_{H_\ell}| = \frac{|V(H_\ell)|!}{|\text{Aut}(H_\ell)|}$, for $\ell \in \{i, j\}$. Then observing that $|V_{H_\ell}^2| = |V(H_\ell)|(|V(H_\ell)| - 1)$, for $\ell \in \{i, j\}$, and using (A.56), (A.57) in combination with (A.53) and (A.58) completes the proof of Lemma A.5. $\square$

A.3.2. *Proof of Proposition A.4.* For $1 \leq i \leq q$, using the expansion of $(f_i)_{K_{\{1\}}}$ in (A.22) and Proposition G.1 it follows that

$$I_1((f_i)_{K_{\{1\}}}) \stackrel{a.s.}{=} \sum_{s \geq 1} \mathbb{E}[f_i \phi_s] I_1(\phi_s), \quad (\text{A.59})$$

where $I_1(\cdot)$ is the 1-dimensional stochastic integral as defined in Section G. Note that $\{I_1(\phi_s)\}_{s \geq 1}$ is a collection of independent $N(0, 1)$ random variables. Hence,

$$\frac{1}{(|V(H_i)| - 1)!} I_1((f_i)_{K_{\{1\}}}) \stackrel{D}{=} \frac{1}{(|V(H_i)| - 1)!} \sum_{s \geq 1} \mathbb{E}[f_i \phi_s] \eta_s = Q_i,$$

for Q_i as defined in (A.46). Now, recalling (A.22) and Definition 2.1 note that

$$\begin{aligned} (f_i)_{K_{\{1\}}}(x) &= \mathbb{E}[f_i | U_1 = x] \\ &= \sum_{H' \in \mathcal{G}_{H_i}} t_1(x, H', W) - |\mathcal{G}_{H_i}| t(H_i, W) \\ &= \frac{|\mathcal{G}_{H_i}|}{|V(H_i)|} \sum_{a \in V(H_i)} (t_a(x, H_i, W) - t(H_i, W)), \end{aligned} \quad (\text{A.60})$$

where the last equality follows by arguments similar to proof of (A.56). Hence, using (A.60) in (A.59) and recalling that $|\mathcal{G}_{H_i}| = \frac{|V(H_i)|!}{|\text{Aut}(H_i)|}$ gives,

$$Q_i \stackrel{D}{=} \frac{1}{|\text{Aut}(H_i)|} \int_0^1 \left\{ \sum_{a=1}^{|V(H_i)|} (t_a(x, H_i, W) - t(H_i, W)) \right\} dB_x,$$

for $1 \leq i \leq q$. This shows (A.49).

Now, suppose $q + 1 \leq i \leq r$. Then from the expansion of $(f_i)_{E_{\{1,2\}}}$ in (A.23) and Proposition G.1,

$$\begin{aligned} I_2((f_i)_{E_{\{1,2\}}}) &\stackrel{a.s.}{=} \sum_{s \geq 1} \mathbb{E}[f_i(\phi_s \otimes \phi_s)] I_2(\phi_s \otimes \phi_s) + 2 \sum_{s < t} \mathbb{E}[f_i(\phi_s \otimes \phi_t)] I_2(\phi_s \otimes \phi_t) \\ &= \sum_{s \geq 1} \mathbb{E}[f_i(\phi_s \otimes \phi_s)] (I_1(\phi_s)^2 - 1) + 2 \sum_{s < t} \mathbb{E}[f_i(\phi_s \otimes \phi_t)] I_1(\phi_s) I_1(\phi_t), \end{aligned}$$

by (G.3), since $\mathbb{E}[I_1(\phi_s)] = 0$ and $\mathbb{E}[I_1(\phi_s)^2] = 1$. As before, noting that $\{I_1(\phi_s)\}_{s \geq 1}$ is a collection of independent $N(0, 1)$ random variables and recalling (A.47) gives,

$$\begin{aligned} & \frac{1}{2(|V(H_i)| - 2)!} I_2 \left((f_i)_{E_{\{1,2\}}} \right) \\ & \stackrel{D}{=} \frac{1}{2(|V(H_i)| - 2)!} \left\{ \sum_{s \geq 1} \mathbb{E}[f_i(\phi_s \otimes \phi_s)] (\eta_s^2 - 1) + 2 \sum_{s < t} \mathbb{E}[f_i(\phi_s \otimes \phi_t)] \eta_s \eta_t \right\} \\ & = R_i. \end{aligned} \tag{A.61}$$

Recalling Definition 2.2 and Lemma A.6 we have for all $q + 1 \leq i \leq r$,

$$(f_i)_{E_{\{1,2\}}}(x, y) = 2(|V(H_i)| - 2)! W_{H_i}(x, y) - \frac{|V(H_i)|!}{|\text{Aut}(H_i)|} t(H_i, W)$$

for almost every $(x, y) \in [0, 1]^2$. Thus, for $q + 1 \leq i \leq r$,

$$\begin{aligned} & I_2 \left((f_i)_{E_{\{1,2\}}} \right) \\ & = 2(|V(H_i)| - 2)! \int_{[0,1]^2} \left\{ W_{H_i}(x, y) - \frac{|V(H_i)| (|V(H_i)| - 1)}{2|\text{Aut}(H_i)|} t(H_i, W) \right\} dB_x dB_y. \end{aligned} \tag{A.62}$$

Combining (A.61) and (A.62), the result in (A.50) follows. This completes the proof of Proposition A.4. $\square$

Lemma A.6. For $q + 1 \leq i \leq r$,

$$(f_i)_{E_{\{1,2\}}}(x, y) = \frac{(|V(H_i)| - 2)!}{|\text{Aut}(H_i)|} \sum_{1 \leq a \neq b \leq |V(H)|} (t_{ab}(x, y, H_i, W) - t(H_i, W)),$$

for almost every $(x, y) \in [0, 1]^2$.

Proof. From (A.23) and Definition 2.2 we have,

$$\begin{aligned} (f_i)_{E_{\{1,2\}}}(x, y) &= \mathbb{E}[f_i | U_1 = x, U_2 = y] \\ &= \sum_{H' \in \mathcal{G}_{H_i}} t_{1,2}(x, y, H', W) - |\mathcal{G}_{H_i}| t(H_i, W), \end{aligned} \tag{A.63}$$

for almost every $(x, y) \in [0, 1]^2$. Denote by $S_{|V(H_i)|}$ the set of all $|V(H_i)|!$ permutations of $V(H_i)$. Then it is easy to observe that

$$\sum_{\xi \in S_{|V(H_i)|}} t_{12}(x, y, \xi(H_i), W) = |\text{Aut}(H)| \sum_{H' \in \mathcal{G}_{H_i}} t_{12}(x, y, H', W). \tag{A.64}$$

where $\xi(H)$ is the graph obtained by permuting the vertex labels of H according to the permutation ξ . Also,

$$\begin{aligned} \sum_{\xi \in S_{|V(H_i)|}} t_{12}(x, y, \xi(H_i), W) &= \sum_{1 \leq a \neq b \leq |V(H_i)|} \sum_{\substack{\xi \in S_{|V(H_i)|} \\ \xi(a)=1, \xi(b)=2}} t_{12}(x, y, \xi(H_i), W) \\ &= \sum_{1 \leq a \neq b \leq |V(H_i)|} \sum_{\substack{\xi \in S_{|V(H_i)|} \\ \xi(a)=1, \xi(b)=2}} t_{\xi^{-1}(1)\xi^{-1}(2)}(x, y, H_i, W) \\ &= \sum_{1 \leq a \neq b \leq |V(H_i)|} \sum_{\substack{\xi \in S_{|V(H_i)|} \\ \xi(a)=1, \xi(b)=2}} t_{a,b}(x, y, H_i, W) \end{aligned}$$

$$= (|V(H_i)| - 2)! \sum_{1 \leq a \neq b \leq |V(H_i)|} t_{a,b}(x, y, H_i, W) \quad (\text{A.65})$$

Combining (A.64) and (A.65), we have,

$$\sum_{H' \in \mathcal{G}_{H_i}} t_{12}(x, y, H', W) = \frac{(|V(H_i)| - 2)!}{|\text{Aut}(H_i)|} \sum_{1 \leq a \neq b \leq |V(H_i)|} t_{a,b}(x, y, H_i, W). \quad (\text{A.66})$$

Thus combining (A.63) and (A.66) gives,

$$\begin{aligned} (f_i)_{E_{\{1,2\}}}(x, y) &= \frac{(|V(H_i)| - 2)!}{|\text{Aut}(H_i)|} \sum_{1 \leq a \neq b \leq |V(H_i)|} t_{a,b}(x, y, H_i, W) - |\mathcal{G}_{H_i}| t(H_i, W) \\ &= \frac{(|V(H_i)| - 2)!}{|\text{Aut}(H_i)|} \sum_{a \neq b} (t_{ab}(x, y, H_i, W) - t(H_i, W)) \end{aligned}$$

where the last equality follows by recalling that $|\mathcal{G}_{H_i}| = \frac{|V(H_i)|!}{|\text{Aut}(H_i)|}$, for all $q + 1 \leq i \leq r$. $\square$

A.4. Completing the Proof of Theorem 2.1. The result in Theorem 2.1 follows by combining Proposition A.2, Proposition A.3, Proposition A.4, and by noting that $\mathbf{T}(\mathcal{H}, W) = (\mathbf{Q}^\top, (\mathbf{R} + \tilde{\mathbf{R}})^\top)^\top$.

APPENDIX B. MOMENT GENERATING FUNCTION OF THE LIMITING DISTRIBUTION

In this section we derive the moment generating function (MGF) of the limiting distribution $\mathbf{Z}(\mathcal{H}, W)$ obtained in Theorem 2.1. We begin by introducing some notation: For any symmetric function $U : [0, 1]^2 \rightarrow \mathbb{R}$, for $L \geq 2$ define its L -th path composition as follows: For $x, y \in [0, 1]$,

$$U^{(L)}(x, y) = \int_{[0,1]^{L-1}} U(x, w_1) U(w_1, w_2) \cdots U(w_{L-1}, y) dw_1 dw_2 \cdots dw_{L-1}. \quad (\text{B.1})$$

$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_r)^\top \in \mathbb{R}^r$, define the functions $V_\alpha : [0, 1] \rightarrow \mathbb{R}$ and $U_\alpha : [0, 1]^2 \rightarrow \mathbb{R}$ as:

$$V_\alpha(x) := \sum_{i=1}^q \alpha_i \left[\frac{1}{|\text{Aut}(H_i)|} \sum_{a=1}^{|V(H_i)|} t_a(x, H_i, W) - \frac{|V(H_i)|}{|\text{Aut}(H_i)|} t(H_i, W) \right], \quad (\text{B.2})$$

and

$$U_\alpha(x, y) := \sum_{i=q+1}^r \alpha_i (W_{H_i}(x, y) - c_{H_i}(W)), \quad (\text{B.3})$$

where $c_{H_i}(W) = \frac{|V(H_i)|(|V(H_i)|-1)}{2|\text{Aut}(H_i)|} t(H_i, W)$ and W_{H_i} is as in Definition 2.2. We can now express the MGF of $\mathbf{Z}(\mathcal{H}, W)$, for $\mathcal{H} = \{H_1, H_2, \dots, H_r\}$ as in Theorem 2.1, as follows:

Proposition B.1. Fix $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_r)^\top \in \mathbb{R}^r$ and let $\mathcal{C} := \sum_{i=q+1}^r |\alpha_i| \frac{|V(H_i)| |V(H_i-1)|}{|\text{Aut}(H_i)|}$. Then for $|\theta| < \frac{1}{32\mathcal{C}}$,

$$\begin{aligned} &\log \mathbb{E} \left[e^{\theta \alpha^\top \mathbf{Z}(\mathcal{H}, W)} \right] \\ &= (\eta_\alpha + \tilde{\eta}_\alpha) \frac{\theta^2}{2} + \sum_{L=1}^{\infty} 2^{L-1} \theta^{L+2} \int_{[0,1]^2} V_\alpha(x) V_\alpha(y) U_\alpha^{(L)}(x, y) dx dy + \frac{1}{2} \sum_{L=3}^{\infty} \frac{(2\theta)^L}{L} \int_0^1 U_\alpha^{(L)}(x, x) dx, \end{aligned}$$

where $U_{\alpha}^{(L)}$ is the L -th path composition of $U^{(L)}$ as defined in (B.1) and

$$\eta_{\alpha} = \sum_{1 \leq i, j \leq q} \frac{\alpha_i \alpha_j}{|Aut(H_i)| |Aut(H_j)|} \sum_{\substack{a \in V(H_i) \\ b \in V(H_j)}} \left(t \left(H_i \oplus_{a,b} H_j, W \right) - t(H_i, W) t(H_j, W) \right),$$

$$\tilde{\eta}_{\alpha} = \sum_{q+1 \leq i, j \leq r} \frac{\alpha_i \alpha_j}{2 |Aut(H_i)| |Aut(H_j)|} \sum_{\substack{a \neq b \in V(H_i) \\ c \neq d \in V(H_j)}} t \left(H_i \ominus_{(a,b),(c,d)} H_j, W \right) - 2 \left(\sum_{i=q+1}^r \alpha_i c_{H_i}(W) \right)^2.$$

B.1. Proof of Proposition B.1. Recalling the definition of $\mathbf{Z}(\mathcal{H}, W)$ from Theorem 2.1 note that

$$\alpha^{\top} \mathbf{Z}(\mathcal{H}, W) = \sum_{i=q+1}^r \alpha_i G_i + \int \int U_{\alpha} dB_x dB_y + \int V_{\alpha} dB_x, \quad (\text{B.4})$$

where $\{G_i : q+1 \leq i \leq r\} \sim N_{r-q}(\mathbf{0}, \Sigma)$ with Σ as Definition 2.4 is independent of the standard Brownian motion $\{B_t : t \in [0, 1]\}$. Now, observe that $U_{\alpha} \in L^2([0, 1]^2)$ is symmetric and hence, the operator,

$$T_{U_{\alpha}} f(x) = \int U_{\alpha}(x, y) f(y) dy, \quad (\text{B.5})$$

where $f \in L^2[0, 1]$, is a self-adjoint Hilbert-Schmidt integral operator. Then by the spectral theorem (see [72, Theorem 8.94 and Theorem 8.83]) we can find a set of orthonormal eigenfunctions $\{\phi_s\}_{s \geq 1}$ corresponding to eigenvalues (with repetition) $\{\lambda_s\}_{s \geq 1}$ of $T_{U_{\alpha}}$ which forms a basis of $L^2[0, 1]$ and

$$U_{\alpha}(x, y) = \sum_{s=1}^{\infty} \lambda_s \phi_s(x) \phi_s(y), \quad (\text{B.6})$$

where the above sum converges in L^2 . Further, we assume that $\{\lambda_s\}_{s \geq 1}$ are arranged according to non-increasing order of magnitude and $\lim_{s \rightarrow \infty} \lambda_s = 0$. Moreover, by the orthonormality of the eigenvectors (see, for example, [49, Lemma 8, Chapter 6]),

$$\sum_{s=1}^{\infty} \lambda_s^2 = \|U_{\alpha}\|_2^2 < \infty. \quad (\text{B.7})$$

Also, since $V_{\alpha} \in L^2[0, 1]$, expanding V_{α} using the basis $\{\phi_s\}_{s \geq 1}$ we have the following,

$$V_{\alpha}(x) = \sum_{s=1}^{\infty} \gamma_s \phi_s(x), \quad (\text{B.8})$$

where once again the above sum converges in L^2 and

$$\sum_{s=1}^{\infty} \gamma_s^2 = \|V_{\alpha}\|_2^2 < \infty. \quad (\text{B.9})$$

Hence, recalling the expression of $\alpha^{\top} \mathbf{Z}(\mathcal{H}, W)$ from (B.4) along with Proposition G.1 and the expansions of U_{α} and V_{α} from (B.6) and from (B.8) respectively gives,

$$\alpha^{\top} \mathbf{Z}(\mathcal{H}, W) \stackrel{a.s.}{=} \sum_{i=q+1}^r \alpha_i G_i + \sum_{s=1}^{\infty} \lambda_s I_2(\phi_s \times \phi_s) + \sum_{s=1}^{\infty} \gamma_s I_1(\phi_s)$$

$$\begin{aligned}
&\stackrel{a.s.}{=} \sum_{i=q+1}^r \alpha_i G_i + \sum_{s=1}^{\infty} \lambda_s (I_1(\phi_s)^2 - 1) + \sum_{s=1}^{\infty} \gamma_s I_1(\phi_s) \quad (\text{by (G.3)}) \\
&\stackrel{D}{=} \sum_{i=q+1}^r \alpha_i G_i + \sum_{s=1}^{\infty} \lambda_s (\eta_s^2 - 1) + \sum_{s=1}^{\infty} \gamma_s \eta_s,
\end{aligned}$$

where $\{\eta_s\}_{s \geq 1}$ is an independent collection of standard Gaussian random variables which is also independent of $\{G_i\}_{q+1 \leq i \leq r}$. Now, for $K \geq 1$, define the truncated version of $\alpha^\top \mathbf{Z}(\mathcal{H}, W)$ as follows:

$$Y_{\alpha, K} := \sum_{i=q+1}^r \alpha_i G_i + \sum_{s=1}^K \lambda_s (\eta_s^2 - 1) + \sum_{s=1}^K \gamma_s \eta_s. \quad (\text{B.10})$$

We begin by computing the MGF of $Y_{\alpha, K}$ in the following lemma:

Lemma B.1. *Let $Y_{\alpha, K}$ be as defined above. Then for $|\theta| < \frac{1}{16\mathcal{C}}$, where $\mathcal{C}$ as in Proposition B.1, the MGF of $Y_{\alpha, K}$ is given by*

$$\log \mathbb{E} \left[e^{\theta Y_{\alpha, K}} \right] = \alpha_+^\top \Sigma \alpha_+ + \frac{\theta^2}{2} + \sum_{s=1}^K \sum_{L=1}^{\infty} 2^{L-2} \theta^{L+1} \gamma_s^2 \lambda_s^{L-1} + \frac{1}{2} \sum_{s=1}^K \sum_{L=2}^{\infty} \frac{(2\lambda_s \theta)^L}{L},$$

where $\alpha_+ = (\alpha_{q+1}, \dots, \alpha_r)^\top$.

Proof. For all $K \geq 1$ define,

$$Y_{\alpha}^{(1)} := \sum_{i=q+1}^r \alpha_i G_i, \quad Y_{\alpha, K}^{(2)} := \sum_{s=1}^K \lambda_s (\eta_s^2 - 1), \quad \text{and} \quad Y_{\alpha, K}^{(3)} := \sum_{s=1}^K \gamma_s \eta_s. \quad (\text{B.11})$$

by the independence of $Y_{\alpha}^{(1)}$ and $(Y_{\alpha, K}^{(2)}, Y_{\alpha, K}^{(3)})$ and the Cauchy-Schwartz inequality we have,

$$\mathbb{E} \left[e^{\theta Y_{\alpha, K}} \right] \leq \mathbb{E} \left[e^{\theta Y_{\alpha}^{(1)}} \right] \left(\mathbb{E} \left[e^{2\theta Y_{\alpha, K}^{(2)}} \right] \mathbb{E} \left[e^{2\theta Y_{\alpha, K}^{(3)}} \right] \right)^{\frac{1}{2}}.$$

Note that the MGFs of $Y_{\alpha}^{(1)}$ and $Y_{\alpha, K}^{(3)}$ exist for all θ . Also, by [34, Theorem 2.2],

$$\mathbb{E} \left[e^{2\theta Y_{\alpha, K}^{(2)}} \right] < \infty, \quad \text{for } |\theta| < \min \left\{ \frac{1}{16\sqrt{\sum_{s=1}^K \lambda_s^2}}, \frac{1}{16\mathcal{C}} \right\}.$$

Recalling (B.3) and (B.7) observe that,

$$\sum_{s=1}^K \lambda_s^2 \leq \|U_{\alpha}\|_2 \leq \sum_{i=q+1}^r |\alpha_i| \frac{|V(H_i)|(|V(H_i)| - 1)}{|\text{Aut}(H_i)|} = \mathcal{C},$$

since $|W_{H_i}| \leq \frac{|V(H_i)|(|V(H_i)| - 1)}{2|\text{Aut}(H_i)|}$, for $q+1 \leq i \leq r$. This implies,

$$\mathbb{E} \left[e^{\theta Y_{\alpha, K}} \right] < \infty, \quad \text{for } |\theta| < \frac{1}{16\mathcal{C}}.$$

Now, we proceed to compute the MGF of $Y_{\alpha, K}$. Observe that by independence,

$$\mathbb{E} \left[e^{\theta Y_{\alpha, K}} \right] = \mathbb{E} \left[e^{\theta Y_{\alpha}^{(1)}} \right] \mathbb{E} \left[e^{\theta (Y_{\alpha, K}^{(2)} + Y_{\alpha, K}^{(3)})} \right]. \quad (\text{B.12})$$

We will consider the following 2 cases:

Case 1: First, assume that $\lambda_s \neq 0$, for all $1 \leq s \leq K$. Then recalling (B.19) and completing the square,

$$\begin{aligned} Y_{\alpha,K}^{(2)} + Y_{\alpha,K}^{(3)} &= \sum_{s=1}^K \lambda_s \left(\eta_s^2 + \frac{\gamma_s \eta_s}{\lambda_s} \right) - \sum_{s=1}^K \lambda_s \\ &= \sum_{s=1}^K \lambda_s \left(\eta_s + \frac{\gamma_s}{2\lambda_s} \right)^2 - \sum_{s=1}^K \frac{\gamma_s^2}{4\lambda_s} - \sum_{s=1}^K \lambda_s \end{aligned} \quad (\text{B.13})$$

Recall that $\{\eta_s\}_{s \geq 1}$ are i.i.d. $N(0, 1)$. Hence, from the MGF of non-central chi-squared distribution and (B.7) we have,

$$\mathbb{E} \left[e^{\theta \sum_{s=1}^K \lambda_s (\eta_s + \frac{\gamma_s}{2\lambda_s})^2} \right] = \prod_{s=1}^K \frac{e^{\frac{\theta \gamma_s^2}{4\lambda_s(1-2\lambda_s\theta)}}}{\sqrt{1-2\lambda_s\theta}} \quad \text{for all } |\theta| < \frac{1}{16C} \leq \frac{1}{2 \max_{1 \leq s \leq K} |\lambda_s|}.$$

This implies, for all $|\theta| < \frac{1}{16C}$,

$$\begin{aligned} \log \mathbb{E} \left[e^{\theta \sum_{s=1}^K \lambda_s (\eta_s + \frac{\gamma_s}{2\lambda_s})^2} \right] &= \sum_{s=1}^K \frac{s\gamma_s^2}{4\lambda_s} (1-2\lambda_s\theta)^{-1} - \frac{1}{2} \sum_{s=1}^K \log(1-2\lambda_s\theta) \\ &= \sum_{s=1}^K \frac{s\gamma_s^2}{4\lambda_s} \sum_{L=0}^{\infty} (2\lambda_s\theta)^L + \frac{1}{2} \sum_{s=1}^K \sum_{L=1}^{\infty} \frac{(2\lambda_s\theta)^L}{L} \\ &= \sum_{s=1}^K \sum_{L=0}^{\infty} 2^{L-2} \theta^{L+1} \gamma_s^2 \lambda_s^{L-1} + \sum_{s=1}^K \lambda_s \theta + \frac{1}{2} \sum_{s=1}^K \sum_{L=2}^{\infty} \frac{(2\lambda_s\theta)^L}{L}. \end{aligned}$$

Recalling (B.13), this implies that

$$\log \mathbb{E} \left[e^{\theta(Y_{\alpha,K}^{(2)} + Y_{\alpha,K}^{(3)})} \right] = \sum_{s=1}^K \sum_{L=1}^{\infty} 2^{L-2} \theta^{L+1} \gamma_s^2 \lambda_s^{L-1} + \frac{1}{2} \sum_{s=1}^K \sum_{L=2}^{\infty} \frac{(2\lambda_s\theta)^L}{L}. \quad (\text{B.14})$$

Case 2: There exists $1 \leq t \leq K$ such that $\lambda_t = 0$. Recall that $\{\lambda_s\}_{s \geq 1}$ are arranged according to non-increasing order of magnitude. Hence, $\lambda_s = 0$ for all $t \leq s \leq K$. In this case,

$$Y_{\alpha,K}^{(2)} + Y_{\alpha,K}^{(3)} = Y_{\alpha,t-1}^{(2)} + Y_{\alpha,t-1}^{(3)} + \sum_{s=t}^K \gamma_s \eta_s.$$

Clearly, $Y_{\alpha,t-1}^{(2)} + Y_{\alpha,t-1}^{(3)}$ is independent of $\sum_{s=t}^K \gamma_s \eta_s$. Moreover, note that $\sum_{s=t}^K \gamma_s \eta_s \sim N(0, \sum_{s=t}^K \gamma_s^2)$. Hence, using (B.14) we have,

$$\begin{aligned} \log \mathbb{E} \left[e^{\theta(Y_{\alpha,K}^{(2)} + Y_{\alpha,K}^{(3)})} \right] &= \sum_{s=1}^{t-1} \sum_{L=1}^{\infty} 2^{L-2} \theta^{L+1} \gamma_s^2 \lambda_s^{L-1} + \frac{1}{2} \sum_{s=1}^{t-1} \sum_{L=2}^{\infty} \frac{(2\lambda_s\theta)^L}{L} + \frac{\theta^2}{2} \sum_{s=t}^K \gamma_s^2 \\ &= \sum_{s=1}^{t-1} \frac{1}{2} \theta^2 \gamma_s^2 + \sum_{s=1}^{t-1} \sum_{L=2}^{\infty} 2^{L-2} \theta^{L+1} \gamma_s^2 \lambda_s^{L-1} + \frac{1}{2} \sum_{s=1}^{t-1} \sum_{L=2}^{\infty} \frac{(2\lambda_s\theta)^L}{L} \\ &= \sum_{s=1}^{t-1} \sum_{L=1}^{\infty} 2^{L-2} \theta^{L+1} \gamma_s^2 \lambda_s^{L-1} + \frac{1}{2} \sum_{s=1}^{t-1} \sum_{L=2}^{\infty} \frac{(2\lambda_s\theta)^L}{L}. \end{aligned} \quad (\text{B.15})$$

Combining (B.14) and (B.15) we can conclude that for all $K \geq 1$,

$$\log \mathbb{E} \left[e^{\theta(Y_{\alpha,K}^{(2)} + Y_{\alpha,K}^{(3)})} \right] = \sum_{s=1}^K \sum_{L=1}^{\infty} 2^{L-2} \theta^{L+1} \gamma_s^2 \lambda_s^{L-1} + \frac{1}{2} \sum_{s=1}^K \sum_{L=2}^{\infty} \frac{(2\lambda_s\theta)^L}{L}. \quad (\text{B.16})$$

Finally, recall that $\mathbf{G} = (G_{q+1}, \dots, G_r)^\top \sim N_{r-q}(\mathbf{0}, \Sigma)$. This implies,

$$\log \mathbb{E} \left[e^{\theta Y_{\alpha}^{(1)}} \right] = \log \mathbb{E} \left[e^{\theta \alpha_+^\top \mathbf{G}} \right] = \frac{\theta^2}{2} \alpha_+^\top \Sigma \alpha_+. \quad (\text{B.17})$$

Now by (B.12), (B.16), and (B.17) we conclude that,

$$\log \mathbb{E} \left[e^{\theta Y_{\alpha, K}} \right] = \frac{\theta^2}{2} \alpha_+^\top \Sigma \alpha_+ + \sum_{s=1}^K \sum_{L=1}^{\infty} 2^{L-2} \theta^{L+1} \gamma_s^2 \lambda_s^{L-1} + \frac{1}{2} \sum_{s=1}^K \sum_{L=2}^{\infty} \frac{(2\lambda_s \theta)^L}{L},$$

for all $|\theta| < \frac{1}{16C}$. $\square$

Now, we compute the MGF of $\alpha^\top \mathbf{Z}(\mathcal{H}, W)$.

Lemma B.2. *The moment generating function of $\alpha^\top \mathbf{Z}(\mathcal{H}, W)$ exists for all $|\theta| < \frac{1}{32C}$ and is given by,*

$$\log \mathbb{E} \left[e^{\theta \alpha^\top \mathbf{Z}(\mathcal{H}, W)} \right] = \frac{\theta^2 c_2}{2} + \sum_{L=1}^{\infty} 2^{L-1} \theta^{L+2} \sum_{s=1}^{\infty} \gamma_s^2 \lambda_s^L + \frac{1}{2} \sum_{L=3}^{\infty} \frac{(2\theta)^L}{L} \sum_{s=1}^{\infty} \lambda_s^L, \quad (\text{B.18})$$

where $c_2 := \alpha_+^\top \Sigma \alpha_+ + \|\mathbf{V}_{\alpha}\|_2^2 + 2\|U_{\alpha}\|_2^2$.

Proof. Define,

$$Y_{\alpha}^{(1)} := \sum_{i=q+1}^r \alpha_i G_i, \quad Y_{\alpha}^{(2)} := \sum_{s=1}^{\infty} \lambda_s (\eta_s^2 - 1), \quad \text{and} \quad Y_{\alpha}^{(3)} := \sum_{s=1}^{\infty} \gamma_s \eta_s. \quad (\text{B.19})$$

Observe that $\{\frac{\eta_s^2 - 1}{\sqrt{2}}\}_{s \geq 1}$ and $\{\eta_s\}_{s \geq 1}$ are orthonormal. Hence, for $Y_{\alpha, K}$ as defined in (B.10) we have,

$$\begin{aligned} \mathbb{E}[(\alpha^\top \mathbf{Z}(\mathcal{H}, W) - Y_{\alpha, K})^2] &\leq \mathbb{E} \left[\sum_{s=K+1}^{\infty} \lambda_s (\eta_s^2 - 1) \right] + \mathbb{E} \left[\sum_{s=K+1}^{\infty} \gamma_s \eta_s \right] \\ &\leq 2 \sum_{s=K+1}^{\infty} \lambda_s^2 + \sum_{s=K+1}^{\infty} \gamma_s^2 \rightarrow 0, \end{aligned}$$

as $K \rightarrow \infty$, by (B.7) and (B.9). Thus,

$$e^{\theta Y_{\alpha, K}} \xrightarrow{P} e^{\theta \alpha^\top \mathbf{Z}(\mathcal{H}, W)}, \quad \text{for all } |\theta| < \frac{1}{32C}.$$

From the proof of Lemma B.1 it follows that $\mathbb{E}[e^{2\theta Y_{\alpha, K}}] < \infty$ for $|\theta| < \frac{1}{32C}$. Hence, $\{e^{\theta Y_{\alpha, K}} : K \geq 1\}$ is uniformly integral for $|\theta| < \frac{1}{32C}$ and

$$\lim_{K \rightarrow \infty} \log \mathbb{E} \left[e^{\theta Y_{\alpha, K}} \right] = \log \mathbb{E} \left[e^{\theta \alpha^\top \mathbf{Z}(\mathcal{H}, W)} \right], \quad \text{for all } |\theta| < \frac{1}{32C}. \quad (\text{B.20})$$

Now, note that

$$\begin{aligned} &\left| \alpha_+^\top \Sigma \alpha_+ \frac{\theta^2}{2} + \sum_{L=1}^{\infty} \sum_{s=1}^{\infty} 2^{L-2} \theta^{L+1} \gamma_s^2 \lambda_s^{L-1} + \frac{1}{2} \sum_{L=2}^{\infty} \sum_{s=1}^{\infty} \frac{(2\lambda_s \theta)^L}{L} - \log \mathbb{E} \left[e^{\theta Y_{\alpha, K}} \right] \right| \\ &\leq \left| \sum_{s=K+1}^{\infty} \sum_{L=1}^{\infty} 2^{L-2} \theta^{L+1} \gamma_s^2 \lambda_s^{L-1} + \frac{1}{2} \sum_{s=K+1}^{\infty} \sum_{L=2}^{\infty} \frac{(2\lambda_s \theta)^L}{L} \right| \\ &\leq \sum_{s=K+1}^{\infty} \sum_{L=1}^{\infty} 2^{L-2} |\theta|^{L+1} \gamma_s^2 |\lambda_s|^{L-1} + \frac{1}{2} \sum_{s=K+1}^{\infty} \sum_{L=2}^{\infty} \frac{|2\lambda_s \theta|^L}{L} \end{aligned}$$

$$\begin{aligned}
&\leq \frac{1}{8} \sum_{s=K+1}^{\infty} \sum_{L=1}^{\infty} \frac{2^{L+1}}{32^{L+1} \mathcal{C}^{L+1}} \gamma_s^2 |\lambda_s|^{L-1} + \frac{1}{32^2} \sum_{s=K+1}^{\infty} \frac{\lambda_s^2}{\mathcal{C}^2} + \frac{1}{2} \sum_{s=K+1}^{\infty} \sum_{L=3}^{\infty} \frac{2^L \mathcal{C}^L}{L s^{L/2} 32^L \mathcal{C}^L} \\
&\leq \frac{1}{8} \sum_{s=K+1}^{\infty} \gamma_s^2 \sum_{L=1}^{\infty} \frac{1}{16^{L+1} \mathcal{C}^2} + \sum_{s=K+1}^{\infty} \frac{\lambda_s^2}{\mathcal{C}^2} + \frac{1}{2} \sum_{s=K+1}^{\infty} \frac{1}{s^{\frac{3}{2}}} \sum_{L=3}^{\infty} \frac{1}{L 16^L} \rightarrow 0,
\end{aligned} \tag{B.21}$$

as $K \rightarrow \infty$. Thus, combining (B.20) and (B.21), for $|\theta| < \frac{1}{32\mathcal{C}}$,

$$\begin{aligned}
&\log \mathbb{E} \left[e^{\theta \alpha^\top \mathbf{Z}(\mathcal{H}, W)} \right] \\
&= \alpha_+^\top \Sigma \alpha_+ + \frac{\theta^2}{2} + \sum_{L=1}^{\infty} \sum_{s=1}^{\infty} 2^{L-2} \theta^{L+1} \gamma_s^2 \lambda_s^{L-1} + \frac{1}{2} \sum_{L=2}^{\infty} \sum_{s=1}^{\infty} \frac{(2\lambda_s \theta)^L}{L} \\
&= \frac{\theta^2}{2} \left(\alpha_+^\top \Sigma \alpha_+ + \sum_{s=1}^{\infty} \gamma_s^2 + 2 \sum_{s=1}^{\infty} \lambda_s^2 \right) + \sum_{L=2}^{\infty} \sum_{s=1}^{\infty} 2^{L-2} \theta^{L+1} \gamma_s^2 \lambda_s^{L-1} + \frac{1}{2} \sum_{L=3}^{\infty} \sum_{s=1}^{\infty} \frac{(2\lambda_s \theta)^L}{L} \\
&= \frac{\theta^2}{2} \left(\alpha_+^\top \Sigma \alpha_+ + \sum_{s=1}^{\infty} \gamma_s^2 + 2 \sum_{s=1}^{\infty} \lambda_s^2 \right) + \sum_{L=1}^{\infty} 2^{L-1} \theta^{L+2} \sum_{s=1}^{\infty} \gamma_s^2 \lambda_s^L + \frac{1}{2} \sum_{L=3}^{\infty} \frac{(2\theta)^L}{L} \sum_{s=1}^{\infty} \lambda_s^L.
\end{aligned}$$

Now, observing that $\sum_{s=1}^{\infty} \gamma_s^2 = \|\mathbf{V}_\alpha\|_2^2$ and $\sum_{s=1}^{\infty} \lambda_s^2 = \|\mathbf{U}_\alpha\|_2^2$, the result in Lemma B.2 follows. $\square$

Now we relate the terms in (B.18) with those in Proposition B.1. To begin with note that $\int_0^1 U_\alpha^{(L)}(x, x) dx$ can be interpreted as the density of the L -cycle for the function U_α . Since $U_\alpha : [0, 1]^2 \rightarrow \mathbb{R}$, is a symmetric and bounded function, it follows from arguments in [55, Section 7.5] that, for $L \geq 3$,

$$\sum_{s=1}^{\infty} \lambda_s^L = \int_0^1 U_\alpha^{(L)}(x, x) dx, \tag{B.22}$$

where $\{\lambda_s\}_{s \geq 1}$ are the eigenvalues, with eigenfunctions $\{\phi_s\}_{s \geq 1}$, for the operator T_{U_α} as defined in (B.5). This implies, by the spectral theorem,

$$U_\alpha^{(L)}(x, y) = \sum_{s=1}^{\infty} \lambda_s^L \phi_s(x) \phi_s(y), \tag{B.23}$$

Using this we relate $\sum_{s=1}^{\infty} \gamma_s^2 \lambda_s^L$ in terms of the functions U_α and V_α .

Lemma B.3. *For all $L \geq 1$,*

$$\sum_{s=1}^{\infty} \gamma_s^2 \lambda_s^L = \int_{[0,1]^2} V_\alpha(x) V_\alpha(y) U_\alpha^{(L)}(x, y) dx dy.$$

Proof. Since $V_\alpha(x) = \sum_{s=1}^{\infty} \gamma_s \phi_s(x)$ (recall (B.8)), we have $\gamma_s = \int_0^1 V_\alpha(x) \phi_s(x) dx$, for $s \geq 1$, by the orthonormality of the eigenvectors. Hence,

$$\sum_{s=1}^{\infty} \gamma_s^2 \lambda_s^L = \int_{[0,1]^2} \left(\sum_{s=1}^{\infty} \lambda_s^L \phi_s(x) \phi_s(y) \right) V_\alpha(x) V_\alpha(y) dx dy = \int_{[0,1]^2} V_\alpha(x) V_\alpha(y) U_\alpha^{(L)}(x, y) dx dy,$$

by (B.23). $\square$

Combining Lemma B.2, (B.22), and Lemma B.3 gives,

$$\log \mathbb{E} \left[e^{\theta \alpha^\top \mathbf{Z}(\mathcal{H}, W)} \right]$$

$$= \frac{\theta^2 c_2}{2} + \sum_{L=1}^{\infty} 2^{L-1} \theta^{L+2} \int_{[0,1]^2} V_{\alpha}(x) V_{\alpha}(y) U_{\alpha}^{(L)}(x, y) dx dy + \frac{1}{2} \sum_{L=3}^{\infty} \frac{(2\theta)^L}{L} \int_0^1 U_{\alpha}^{(L)}(x, x) dx,$$

where $c_2 = \alpha_+^{\top} \Sigma \alpha_+ + \|V_{\alpha}\|_2^2 + 2\|U_{\alpha}\|_2^2$. The result in Proposition B.1 now follows from the next lemma.

Lemma B.4. $\eta_{\alpha}, \tilde{\eta}_{\alpha}$ as defined in Proposition B.1 the following hold:

$$\|V_{\alpha}\|_2^2 = \eta_{\alpha} \quad \text{and} \quad 2\|U_{\alpha}\|_2^2 + \alpha_+^{\top} \Sigma \alpha_+ = \tilde{\eta}_{\alpha}.$$

Proof. Recalling the definition of V_{α} from (B.2) we have,

$$\begin{aligned} & \int_0^1 V_{\alpha}^2(x) dx \\ &= \sum_{1 \leq i, j \leq q} \kappa_i \kappa_j \int_0^1 \sum_{\substack{a \in V(H_i) \\ b \in V(H_j)}} (t_a(x, H_i, W) - t(H_i, W)) (t_a(x, H_j, W) - t(H_j, W)) dx, \end{aligned} \quad (\text{B.24})$$

where $\kappa_i := \frac{\alpha_i}{|\text{Aut}(H_i)|}$. Observe that,

$$\begin{aligned} & \int_0^1 (t_a(x, H_i, W) - t(H_i, W)) (t_a(x, H_j, W) - t(H_j, W)) dx \\ &= \int_0^1 t_a(x, H_i, W) t_b(x, H_j, W) - t(H_i, W) t_b(x, H_j, W) - t(H_j, W) t_a(x, H_j, W) + t(H_i, W) t(H_j, W) dx \\ &= t \left(H_i \bigoplus_{a,b} H_j, W \right) - t(H_i, W) t(H_j, W). \end{aligned} \quad (\text{B.25})$$

Combining (B.24) and (B.25) gives $\|V_{\alpha}\|_2^2 = \eta_{\alpha}$.

Next, recalling the definition of U_{α} from (B.3) we have,

$$\int_{[0,1]^2} U_{\alpha}^2 dx dy = \int_{[0,1]^2} \left(\sum_{i=q+1}^r \alpha_i W_{H_i}(x, y) \right)^2 dx dy - \left(\sum_{i=q+1}^r \alpha_i c(H_i, W) \right)^2, \quad (\text{B.26})$$

since $\int_{[0,1]^2} W_{H_i}(x, y) dx dy = c(H_i, W)$. By definition,

$$\int_{[0,1]^2} \left(\sum_{i=q+1}^r \alpha_i W_{H_i}(x, y) \right)^2 dx dy = \sum_{q+1 \leq i, j \leq r} \alpha_i \alpha_j \int_{[0,1]^2} W_{H_i}(x, y) W_{H_j}(x, y) dx dy, \quad (\text{B.27})$$

and

$$\begin{aligned} & \int_{[0,1]^2} W_{H_i}(x, y) W_{H_j}(x, y) dx dy \\ &= \frac{1}{4 |\text{Aut}(H_i)| |\text{Aut}(H_j)|} \sum_{\substack{a \neq b \in V(H_i) \\ c \neq d \in V(H_j)}} \int_{[0,1]^2} t_{a,b}(x, y, H_i, W) t_{c,d}(x, y, H_j, W) dx dy \\ &= \frac{1}{4 |\text{Aut}(H_i)| |\text{Aut}(H_j)|} \sum_{\substack{a \neq b \in V(H_i) \\ c \neq d \in V(H_j)}} t \left(H_i \bigoplus_{(a,b),(c,d)} H_j, W \right). \end{aligned} \quad (\text{B.28})$$

Finally, recalling the definition of Σ from Definition 2.4 note that

$$\alpha_+^{\top} \Sigma \alpha_+$$

$$\begin{aligned}
&= \sum_{q+1 \leq i, j \leq r} \frac{\kappa_i \kappa_j}{2} \sum_{\substack{(a,b) \in E^+(H_i) \\ (c,d) \in E^+(H_j)}} \left(t(H_i \ominus_{(a,b),(c,d)} H_j, W) - t(H_i \oplus_{(a,b),(c,d)} H_j, W) \right) \\
&= \sum_{q+1 \leq i, j \leq r} \frac{\kappa_i \kappa_j}{2} \sum_{\substack{a \neq b \in V(H_i) \\ c \neq d \in V(H_j)}} \left(t(H_i \ominus_{(a,b),(c,d)} H_j, W) - t(H_i \oplus_{(a,b),(c,d)} H_j, W) \right). \tag{B.29}
\end{aligned}$$

where the last equality follows from Remark B.1. Combining (B.26), (B.27), (B.28), and (B.29) we have $2\|U_\alpha\|_2^2 + \alpha_+^\top \Sigma \alpha_+ = \tilde{\eta} \alpha_+$. $\square$

Remark B.1. Note that both the weak and strong edge join operations can be extended to arbitrary $(a, b) \in V(H_1)^2$ and $(c, d) \in V(H_2)^2$, with $a \neq b$ and $c \neq d$ as follows: For the strong join we keep all edges, while for the weak join we keep the joined graph simple by merging any resulting double edge. In particular, if either $(a, b) \notin E^+(H_1)$ or $(c, d) \notin E^+(H_2)$, then the weak and strong edge joins are the same graph. This implies,

$$t(H_1 \ominus_{(a,b),(c,d)} H_2, W) = t(H_1 \oplus_{(a,b),(c,d)} H_2, W),$$

which explains the step in (B.29).

APPENDIX C. PROOF OF COROLLARY 2.2 AND COROLLARY 2.1

C.1. Proof of Corollary 2.2. Since W is H -irregular for all $H \in \mathcal{H}$, by Theorem 2.1,

$$\mathbf{Z}(\mathcal{H}, G_n) \xrightarrow{D} \mathbf{Z}(\mathcal{H}, W) = (Z(H_1, W), \dots, Z(H_r, W))$$

where

$$Z(H_i, W) = \int_0^1 \left\{ \frac{1}{|\text{Aut}(H_i)|} \sum_{a=1}^{|V(H_i)|} t_a(x, H_i, W) - \frac{|V(H_i)|}{|\text{Aut}(H_i)|} t(H_i, W) \right\} dB_x,$$

for $1 \leq i \leq r$. Since a linear stochastic integral has a centered Gaussian distribution,

$$\mathbf{Z}(\mathcal{H}, W) \stackrel{D}{=} N_r(\mathbf{0}, \Gamma),$$

where $\Gamma = ((\tau_{ij}))_{1 \leq i, j \leq r}$, with

$$\begin{aligned}
\tau_{ij} &= \text{Cov}(Z(H_i, W), Z(H_j, W)) \\
&= \frac{1}{|\text{Aut}(H_i)| |\text{Aut}(H_j)|} \left\{ \sum_{a=1}^{|V(H_i)|} \sum_{b=1}^{|V(H_j)|} \int_0^1 t_a(x, H_i, W) t_b(x, H_j, W) dx \right. \\
&\quad \left. - |V(H_i)| t(H_i, W) |V(H_j)| t(H_j, W) \right\}.
\end{aligned}$$

A direct computation shows that, for all $1 \leq i, j \leq r$,

$$\int_0^1 t_a(x, H_i, W) t_b(x, H_j, W) dx = t \left(H_i \oplus_{a,b} H_j, W \right).$$

This shows, $\tau_{ij} = \text{Cov}(Z(H_i, W), Z(H_j, W))$ equals to the expression in the statement of Corollary 2.2.

C.2. Proof of Corollary 2.1. Note that when W is H -irregular, the result is immediate from Corollary 2.2. Hence, suppose W is H -regular. In this case, from Theorem 2.1 we know that,

$$Z(H, G_n) \xrightarrow{D} G + \underbrace{\int_0^1 \int_0^1 \left\{ W_H(x, y) - \frac{|V(H)| (|V(H) - 1|)}{2|\text{Aut}(H)|} t(H, W) \right\} dB_x dB_y}_{Z(H, W)}, \quad (\text{C.1})$$

where $G \sim N(0, \eta_{H, W}^2)$, with

$$\eta_{H, W}^2 = \frac{1}{2|\text{Aut}(H)|^2} \sum_{(a, b), (c, d) \in E^+(H)} \left[t \left(H \ominus_{(a, b), (c, d)} H, W \right) - t \left(H \oplus_{(a, b), (c, d)} H, W \right) \right], \quad (\text{C.2})$$

and G is independent of the Brownian motion $\{B_t : 0 \leq t \leq 1\}$. Note that the expression of $\eta_{H, W}$ follows from Theorem 2.1 and (2.11). From (2.7) recall that,

$$d_{W_H} = \frac{|V(H)| (|V(H) - 1|)}{2|\text{Aut}(H)|} t(H, W),$$

is an eigenvalue of the kernel W_H with corresponding eigenfunction 1. Now, considering the spectral decomposition of W_H notice,

$$W_H(x, y) - \frac{|V(H)| (|V(H) - 1|)}{2|\text{Aut}(H)|} t(H, W) = \sum_{\lambda \in \text{Spec}^-(W_H)} \lambda \phi_\lambda(x) \phi_\lambda(y),$$

almost everywhere. Then,

$$Z(H, W) = \int_0^1 \int_0^1 \sum_{\lambda \in \text{Spec}^-(W_H)} \lambda \phi_\lambda(x) \phi_\lambda(y) dB_x dB_y = \sum_{\lambda \in \text{Spec}^-(W_H)} \int_0^1 \int_0^1 \lambda \phi_\lambda(x) \phi_\lambda(y) dB_x dB_y,$$

where the last equality follows by Proposition G.1. Now, using (G.3) and $\int_0^1 \phi_\lambda(x)^2 dx = 1$, we get

$$\int_0^1 \int_0^1 \lambda \phi_\lambda(x) \phi_\lambda(y) dB_x dB_y = \lambda \left[\left(\int_0^1 \phi_\lambda(x) dB_x \right)^2 - 1 \right], \text{ for all } \lambda \in \text{Spec}^-(W_H).$$

The orthonormality of the eigenvectors $\{\phi_\lambda\}_{\lambda \in \text{Spec}^-(W_H)}$ implies,

$$Z(H, W) = \sum_{\lambda \in \text{Spec}^-(W_H)} \lambda \left[\left(\int_0^1 \phi(x) dB_x \right)^2 - 1 \right] \stackrel{D}{=} \sum_{\lambda \in \text{Spec}^-(W_H)} \lambda (Z_\lambda^2 - 1), \quad (\text{C.3})$$

where $\{Z_\lambda : \lambda \in \text{Spec}^-(W_H)\}$ are i.i.d. $N(0, 1)$ which are independent from G . The proof of Corollary 2.1 is now complete by collecting (C.1), (C.2), and (C.3).

APPENDIX D. PROOF OF THEOREM 4.1

We begin by expressing the estimated distributions $\hat{Z}(\mathcal{H}, G_n)$ (recall (4.7)) in terms of stochastic integrals. For this, suppose $I_1, I_2, \dots, I_n$ be a partition of $[0, 1]$ into intervals of length $1/n$, that is, $I_s = [\frac{s-1}{n}, \frac{s}{n})$, for $1 \leq s \leq n$. Let $\eta_s = \int_{I_s} dB_s$, where $(B_t)_{t \in [0, 1]}$ is a standard Brownian motion on $[0, 1]$ independent of $\{G_n\}_{n \geq 1}$. Then $\{\eta_1, \eta_2, \dots, \eta_n\}$ is a collection of i.i.d. $N(0, 1/n)$ random variables. With notations as in (4.5), define

$$\hat{Z}'(H_i, G_n) := \begin{cases} \sum_{v=1}^n (\hat{t}(v, H_i, G_n) - \bar{t}(H_i, G_n)) \eta_v & \text{if } 1 \leq i \leq q, \\ \sum_{1 \leq u, v \leq n} (\hat{W}_{H_i}^{G_n}(u, v) - \bar{W}_{H_i}^{G_n}) \left(\eta_u \eta_v - \frac{\delta_{u,v}}{n} \right) & \text{if } q+1 \leq i \leq r, \end{cases} \quad (\text{D.1})$$

where, recall that, $\delta_{u,v} = \mathbf{1}\{u = v\}$, $\bar{t}(H_i, G_n) = \frac{1}{n} \sum_{v=1}^n \hat{t}(v, H_i, G_n)$, and $\bar{W}_{H_i}^{G_n} = \frac{1}{n^2} \sum_{1 \leq u, v \leq n} \hat{W}_{H_i}^{G_n}(u, v)$. Denote,

$$\hat{Z}'(\mathcal{H}, G_n) = (\hat{Z}'(H_1, G_n), \hat{Z}'(H_2, G_n), \dots, \hat{Z}'(H_r, G_n))^\top. \quad (\text{D.2})$$

Note that $\hat{Z}'(\mathcal{H}, G_n)$ has the same distribution as $\hat{Z}(\mathcal{H}, G_n)$ and $\hat{Z}'(\mathcal{H}, G_n)$ is defined on the same probability space as $\{B_t\}_{t \in [0,1]}$. Now, recalling (4.3), for $x \in [0, 1]$, define

$$\hat{t}(x, H_i, G_n) = \hat{t}([nx], H, G_n). \quad (\text{D.3})$$

Note that $\int_0^1 \hat{t}(x, H_i, G_n) dx = \bar{t}(H_i, G_n)$. Also, recalling (4.4), for $x, y \in [0, 1]$, define

$$\hat{W}_{H_i}^{G_n}(x, y) = \hat{W}_{H_i}^{G_n}([nx], [ny]). \quad (\text{D.4})$$

Observe that $\int_{[0,1]^2} \hat{W}_{H_i}^{G_n}(x, y) dx dy = \bar{W}_{H_i}^{G_n}$. Hence, (D.1) can be expressed as:

$$\hat{Z}'(H_i, G_n) := \begin{cases} \int_0^1 \left(\hat{t}(x, H_i, G_n) - \int_0^1 \hat{t}(x, H_i, G_n) dx \right) dB_x & \text{if } 1 \leq i \leq q, \\ \int_{[0,1]^2} \left(\hat{W}_{H_i}^{G_n}(x, y) - \int_{[0,1]^2} \hat{W}_{H_i}^{G_n}(x, y) dx dy \right) dB_x dB_y & \text{if } q+1 \leq i \leq r. \end{cases}$$

In the next lemma we show that distribution of $\hat{Z}'(H_i, G_n)$ remains unchanged in the limit when $\hat{t}(x, H_i, G_n)$ is replaced by $\bar{t}(x, H, W^{G_n})$ (recall (2.4)) and $\hat{W}_{H_i}^{G_n}(x, y)$ (recall (D.4)) is replaced by $W_{H_i}^{G_n}$, the 2-point conditional homomorphism kernel of the empirical graphon W^{G_n} (recall (2.5)). This is because the difference between all homomorphisms and injective homomorphisms is negligible in the limit.

Lemma D.1. *For $1 \leq i \leq r$, as $n \rightarrow \infty$,*

$$\mathbb{E} \left[\left| \hat{Z}'(H_i, G_n) - Y(H_i, W^{G_n}) \right|^2 \middle| G_n \right] \xrightarrow{a.s.} 0, \quad (\text{D.5})$$

where $Y(H_i, W^{G_n})$ is defined as follows:

- for $1 \leq i \leq q$,

$$Y(H_i, W^{G_n}) := \int_0^1 \left\{ \frac{1}{|\text{Aut}(H_i)|} \sum_{a=1}^{|V(H_i)|} t_a(x, H_i, W^{G_n}) - \frac{|V(H_i)|}{|\text{Aut}(H_i)|} t(H_i, W^{G_n}) \right\} dB_x,$$

- for $q+1 \leq i \leq r$,

$$Y(H_i, W^{G_n}) := \int_0^1 \int_0^1 \left\{ W_{H_i}^{G_n}(x, y) - \frac{|V(H_i)|(|V(H_i) - 1|)}{2|\text{Aut}(H_i)|} t(H_i, W^{G_n}) \right\} dB_x dB_y.$$

Proof. We start the proof by showing (D.5) for $1 \leq i \leq q$, that is, when W is H_i -irregular.

Towards this, notice that $t(H_i, W^{G_n}) = \int t_a(x, H_i, W^{G_n}) dx$, for all $1 \leq a \leq |V(H_i)|$, and hence,

$$\begin{aligned} & \mathbb{E} \left[\left| \hat{Z}'(H_i, G_n) - Y(H_i, W^{G_n}) \right|^2 \middle| G_n \right] \\ & \lesssim \int_0^1 \left(\hat{t}(x, H_i, G_n) - \frac{1}{|\text{Aut}(H_i)|} \sum_{a=1}^{|V(H_i)|} t_a(x, H_i, W^{G_n}) \right)^2 dx \\ & = \sum_{v=1}^n \int_{I_v} \left(\hat{t}(v, H_i, G_n) - \frac{1}{|\text{Aut}(H_i)|} \sum_{a=1}^{|V(H_i)|} t_a(x, H_i, W^{G_n}) \right)^2 dx, \end{aligned} \quad (\text{D.6})$$

almost surely, where $I_v = [\frac{v-1}{n}, \frac{v}{n})$, for $1 \leq v \leq n$. Note that the first inequality follows by the boundedness property of stochastic integrals (see Section G) and the second equality follows from the definition of $\hat{t}(\cdot, H_i, G_n)$ in (D.3). From (4.3) recall that,

$$\hat{t}(v, H_i, G_n) = \frac{1}{|\text{Aut}(H_i)|} \sum_{a=1}^{|V(H_i)|} \frac{X_a(v, H_i, G_n)}{n^{|V(H_i)|-1}},$$

for $1 \leq v \leq n$. Then from (D.6) we get,

$$\begin{aligned} & \mathbb{E} \left[\left| \hat{Z}'(H_i, G_n) - Y(H_i, W^{G_n}) \right|^2 \middle| G_n \right] \\ & \lesssim_{H_i} \sum_{v=1}^n \sum_{a=1}^{|V(H_i)|} \int_{I_v} \left(\frac{X_a(v, H_i, G_n)}{n^{|V(H_i)|-1}} - t_a(x, H_i, W^{G_n}) \right)^2 dx, \end{aligned} \quad (\text{D.7})$$

almost surely. Now, recall Definition 2.1 to write,

$$t_a(x, H_i, W^{G_n}) = \int_{[0,1]^{|V(H_i)|-1}} \prod_{u \in N_{H_i}(a)} W^{G_n}(x, x_u) \prod_{(u,v) \in E(H_i \setminus \{a\})} W^{G_n}(x_u, x_v) \prod_{v \in V(H_i \setminus \{a\})} dx_v.$$

Recalling the construction of empirical graphon W^{G_n} from (2.3), $t_a(x, H_i, W^{G_n})$ can be equivalently written as follows:

$$t_a(x, H_i, W^{G_n}) = \frac{1}{n^{|V(H_i)|-1}} \sum_{\tilde{\mathbf{s}}_{\{a\}^c}} \prod_{y \in N_{H_i}(a)} w_{vs_y}(G_n) \prod_{(x,y) \in E(H_i \setminus \{a\})} w_{sx sy}(G_n),$$

for all $x \in I_v$ and $1 \leq v \leq n$, where the sum is over tuples $\tilde{\mathbf{s}}_{\{a\}^c} = (s_x)_{x \in V(H_i) \setminus \{a\}} \in \{[n] \setminus \{v\}\}^{|V(H_i)|-1}$. From Definition 4.1 recall that $X_a(v, H_i, G_n)$ is sum over tuples $\mathbf{s}_{\{a\}^c} = (s_x)_{x \in V(H_i) \setminus \{a\}}$ where the elements are all distinct. Thus, for all $x \in I_v$ and $1 \leq v \leq n$,

$$\left| \frac{X_a(v, H_i, G_n)}{n^{|V(H_i)|-1}} - t_a(x, H_i, W^{G_n}) \right| \lesssim_{H_i} \frac{1}{n}, \quad (\text{D.8})$$

since the difference in the LHS above counts the number of non-injective homomorphisms $\phi : V(H_i) \rightarrow V(G_n)$ such that $\phi(a) = v$, up to a constant depending on H_i . Substituting the bound from (D.8) in (D.7) gives,

$$\mathbb{E} \left[\left| \hat{Z}'(H_i, G_n) - Y(H_i, W^{G_n}) \right|^2 \middle| G_n \right] \lesssim_{H_i} \frac{1}{n^2},$$

almost surely. This proves (D.5), for $1 \leq i \leq q$.

Next, consider $q+1 \leq i \leq r$, that is, W is H_i -regular. Note that

$$t(H_i, W^{G_n}) = \int_0^1 \int_0^1 W_{H_i}^{G_n}(x, y) dx dy.$$

Then once again by the boundedness property of stochastic integrals we get,

$$\begin{aligned} & \mathbb{E} \left[\left| \hat{Z}'(H_i, G_n) - Y(H_i, W^{G_n}) \right|^2 \middle| G_n \right] \\ & \lesssim \int_0^1 \int_0^1 \left(\hat{W}_{H_i}^{G_n}(x, y) - W_{H_i}^{G_n}(x, y) \right)^2 dx dy \\ & \lesssim_{H_i} \sum_{1 \leq u \neq v \leq n} \int_{I_u \times I_v} \left(\hat{W}_{H_i}^{G_n}(x, y) - W_{H_i}^{G_n}(x, y) \right)^2 + \frac{1}{n} \end{aligned} \quad (\text{D.9})$$

$$\lesssim_{H_i} \sum_{\substack{1 \leq u \neq v \leq n \\ 1 \leq a \neq b \leq |V(H_i)|}} \int_{I_u \times I_v} \left(\frac{X_{a,b}(u, v, H_i, G_n)}{n^{|V(H_i)|-2}} - W_{H_i}^{G_n}(x, y) \right)^2 + \frac{1}{n}, \quad (\text{D.10})$$

almost surely. Note that the inequality in (D.9) follows since $\hat{W}_{H_i}^{G_n} = 0$ on $I_u \times I_u$, for all $1 \leq u \leq n$, and $W_{H_i}^{G_n}$ is bounded. Furthermore, (D.10) follows from the definition of $\hat{W}_{H_i}^{G_n}$ in (4.4).

Now, recalling $X_{a,b}$ from Definition 4.2 and by counting arguments similar to (D.8) gives,

$$\mathbb{E} \left[\left| \hat{Z}'(H_i, G_n) - Y(H_i, W^{G_n}) \right|^2 \middle| G_n \right] \lesssim_{H_i} \frac{1}{n^2},$$

almost surely. This completes the proof of Lemma D.1. $\square$

Now, define

$$\mathbf{Y}(\mathcal{H}, W^{G_n}) := (Y(H_1, W^{G_n}), \dots, Y(H_r, W^{G_n}))^\top. \quad (\text{D.11})$$

Note that although $\mathbf{Y}(\mathcal{H}, W^{G_n})$ resembles the empirical analogue of $\mathbf{Z}(\mathcal{H}, W)$ (recall (2.12)), with W replaced with W^{G_n} , one important difference is that in the regular regime, that is, when $q+1 \leq i \leq r$, $Y(H_i, W^{G_n})$ does not have the Gaussian component G_i , unlike in its population counterpart $Z(H_i, W^{G_n})$.

Since $\hat{Z}'(\mathcal{H}, G_n)$ has the same distribution as $\hat{\mathbf{Z}}(\mathcal{H}, G_n)$ (recall (D.2)), Lemma D.1 implies that to prove Theorem 4.1 it suffices to show that $\mathbf{Y}(\mathcal{H}, W^{G_n})|_{G_n}$ converges in distribution (almost surely) to $\mathbf{Z}(\mathcal{H}, W)$. We will establish this by showing that the MGF of $\alpha^\top \mathbf{Y}(\mathcal{H}, W^{G_n})$ conditioned on the graph G_n will converge the MGF of $\alpha^\top \mathbf{Z}(\mathcal{H}, W)$ almost surely in a neighborhood of zero, for all $\alpha \in \mathbb{R}^r$. This is formalized in the following Proposition D.1, which is proved in Section D.1.

Proposition D.1. *For any $\alpha \in \mathbb{R}^r$ and $|\theta| < \frac{1}{32\mathcal{C}}$, where $\mathcal{C}$ is defined in Proposition B.1, the following hold:*

$$\lim_{n \rightarrow \infty} \log \mathbb{E} \left[e^{\theta \alpha^\top \mathbf{Y}(\mathcal{H}, W^{G_n})} \middle| G_n \right] = \log \mathbb{E} \left[e^{\theta \alpha^\top \mathbf{Z}(\mathcal{H}, W)} \right],$$

on a set $\mathcal{A}$ (not depending on α) such that $\mathbb{P}(\mathcal{A}) = 1$.

Proposition D.1 implies that

$$\mathbf{Y}(\mathcal{H}, W^{G_n})|_{G_n} \xrightarrow{D} \mathbf{Z}(\mathcal{H}, W),$$

for all $\alpha \in \mathbb{R}^r$, on the set $\mathcal{A}$. Since the above convergence holds for all $\alpha \in \mathbb{R}^r$, the result in Theorem 4.1 follows from the Cramér-Wold device, Lemma D.1, and recalling that $\hat{Z}'(\mathcal{H}, G_n)$ has the same limiting distribution as $\hat{\mathbf{Z}}(\mathcal{H}, G_n)$.

D.1. Proof of Proposition D.1. Let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_r)^\top \in \mathbb{R}^r$. Similar to (B.2) we define,

$$V_{\alpha, G_n}(x) := \sum_{i=1}^q \alpha_i \left[\frac{1}{|\text{Aut}(H_i)|} \sum_{a=1}^{|V(H_i)|} t_a(x, H_i, W^{G_n}) - \frac{|V(H_i)|}{|\text{Aut}(H_i)|} t(H_i, W^{G_n}) \right], \quad (\text{D.12})$$

and, similar to (B.3), let us define,

$$U_{\alpha, G_n}(x, y) := \sum_{i=q+1}^r \alpha_i \left(W_{H_i}^{G_n}(x, y) - c(H_i, W^{G_n}) \right), \quad (\text{D.13})$$

where $c(H_i, W^{G_n}) := \frac{|V(H_i)|(|V(H_i)|-1)}{2|\text{Aut}(H_i)|} t(H_i, W^{G_n})$.

Lemma D.2. *There exists a set $\mathcal{A}$ such that $\mathbb{P}(\mathcal{A}) = 1$ on which the following hold:*

(1) For $L \geq 3$,

$$\lim_{n \rightarrow \infty} \int U_{\alpha, G_n}^{(L)}(x, x) dx = \int U_{\alpha}^{(L)}(x, x) dx,$$

where $U_{\alpha, G_n}^{(L)}$ and $U_{\alpha}^{(L)}$ are the L -th path composition of U_{α, G_n} and U_{α} , respectively.

(2) For $L \geq 1$,

$$\lim_{n \rightarrow \infty} \int V_{\alpha, G_n}(x) U_{\alpha, G_n}^{(L)}(x, y) V_{\alpha, G_n}(y) dx dy = \int V_{\alpha}(x) U_{\alpha}^{(L)}(x, y) V_{\alpha}(y) dx dy.$$

The proof of Lemma D.2 is given in Section D.2. Here, we apply this lemma to complete the proof of Proposition D.1.

To begin with note that the Proposition B.1 holds for any graphon, in particular, it holds for the empirical graphon W^{G_n} . Hence, using this expression, with W replaced by W^{G_n} and Σ replaced by zero (recalling the definition of $\mathbf{Y}(\mathcal{H}, W^{G_n})$ from (D.11)) we get,

$$\begin{aligned} & \log \mathbb{E} \left[e^{\theta \alpha^\top \mathbf{Y}(\mathcal{H}, W^{G_n})} | G_n \right] \\ &= \frac{\theta^2 c_{2,n}}{2} + \sum_{L=1}^{\infty} 2^{L-1} \theta^{L+2} \int_{[0,1]^2} V_{\alpha, G_n}(x) V_{\alpha, G_n}(y) U_{\alpha, G_n}^{(L)}(x, y) dx dy + \frac{1}{2} \sum_{L=3}^{\infty} \frac{(2\theta)^L}{L} \int_0^1 U_{\alpha, G_n}^{(L)}(x, x) dx, \end{aligned}$$

where

$$c_{2,n} = \|V_{\alpha, G_n}\|_2^2 + 2\|U_{\alpha, G_n}\|_2^2. \quad (\text{D.14})$$

Note that $|U_{\alpha, G_n}| \leq \mathcal{C}$, where $\mathcal{C}$ is defined in Proposition B.1. Then, for all $|\theta| < \frac{1}{32\mathcal{C}}$ and $L \geq 3$ we get,

$$\frac{2^L |\theta|^L}{L} \left| \int U_{\alpha, G_n}^{(L)}(x, x) dx \right| \leq \frac{2^L}{\mathcal{C}^L 32^L L} \mathcal{C}^L \leq \frac{1}{16^L L}.$$

Observe that $\sum_{L=3}^{\infty} \frac{1}{16^L L} < \infty$. Hence, using the Dominated Convergence Theorem and Lemma D.2 we conclude,

$$\lim_{n \rightarrow \infty} \frac{1}{2} \sum_{L=3}^{\infty} \frac{(2\theta)^L}{L} \int U_{\alpha, G_n}^{(L)}(x, x) dx = \frac{1}{2} \sum_{L=3}^{\infty} \frac{(2\theta)^L}{L} \int_0^1 U_{\alpha}^{(L)}(x, x) dx \quad (\text{D.15})$$

on the set $\mathcal{A}$. Next, note that

$$2^{L-1} |\theta|^{L+2} \left| \int V_{\alpha, G_n}(x) V_{\alpha, G_n}(y) U_{\alpha, G_n}^{(L)}(x, y) dx dy \right| \leq 2^{L-1} \frac{1}{32^{L+2} \mathcal{C}^{L+2}} \tilde{\mathcal{C}}^2 \mathcal{C}^L \leq \frac{1}{16^{L-1} \mathcal{C}^2} \tilde{\mathcal{C}}^2,$$

where $\tilde{\mathcal{C}} = 2 \sum_{i=1}^q |\alpha_i| \frac{|V(H_i)|}{|\text{Aut}(H_i)|}$. Now, observe that $\frac{\tilde{\mathcal{C}}^2}{\mathcal{C}^2} \sum_{L=1}^{\infty} \frac{1}{16^{L-1}} < \infty$. Hence, using DCT and Lemma D.2 we conclude,

$$\begin{aligned} & \lim_{n \rightarrow \infty} \sum_{L=1}^{\infty} 2^{L-1} s^{L+2} \int V_{\alpha, G_n}(x) V_{\alpha, G_n}(y) U_{\alpha, G_n}^{(L)}(x, y) dx dy \\ &= \sum_{L=1}^{\infty} 2^{L-1} s^{L+2} \int V_{\alpha}(x) V_{\alpha}(y) U_{\alpha}^{(L)}(x, y) dx dy, \end{aligned} \quad (\text{D.16})$$

on $\mathcal{A}$. Now, denote $\kappa_i = \frac{\alpha_i}{|\text{Aut}(H_i)|}$. Then from the proof of Lemma B.4 we get, as $n \rightarrow \infty$

$$\begin{aligned} \|V_{\alpha, G_n}\|_2^2 &= \sum_{1 \leq i, j \leq q} \kappa_i \kappa_j \sum_{\substack{a \in V(H_i) \\ b \in V(H_j)}} \left[t \left(H_i \bigoplus_{a,b} H_j, W^{G_n} \right) - t(H_i, W^{G_n}) t(H_j, W^{G_n}) \right] \\ &\rightarrow \sum_{1 \leq i, j \leq q} \kappa_i \kappa_j \sum_{\substack{a \in V(H_i) \\ b \in V(H_j)}} \left[t \left(H_i \bigoplus_{a,b} H_j, W \right) - t(H_i, W) t(H_j, W) \right] = \eta_{\alpha}, \end{aligned} \quad (\text{D.17})$$

on the set $\mathcal{A}$, since the vertex-join of 2 simple graphs produces another simple graph. Also, from the proof of Lemma B.4, with W^{G_n} in place of W (see in (B.26) and (B.28))

$$\|U_{\alpha, G_n}\|_2^2 = \sum_{q+1 \leq i, j \leq r} \frac{\kappa_i \kappa_j}{4} \sum_{\substack{a \neq b \in V(H_i) \\ c \neq d \in V(H_j)}} t \left(H_i \bigoplus_{(a,b),(c,d)} H_j, W^{G_n} \right) - \left(\sum_{i=1}^q \alpha_i c(H_i, W^{G_n}) \right)^2$$

Now, since W^{G_n} is the empirical graphon corresponding to the graph G_n , W^{G_n} is $\{0, 1\}$ -valued. Thus, $W^{G_n}(x, y)^2 = W^{G_n}(x, y)$, for all $x, y \in [0, 1]$, hence,

$$t \left(H_i \bigoplus_{(a,b),(c,d)} H_j, W^{G_n} \right) = t \left(H_i \bigoplus_{(a,b),(c,d)} H_j, W \right),$$

for $a \neq b \in V(H_i)$, $c \neq d \in V(H_j)$, and $q+1 \leq i, j \leq r$. Thus, on the set $\mathcal{A}$,

$$\begin{aligned} 2\|U_{\alpha, G_n}\|_2^2 &= \sum_{q+1 \leq i, j \leq r} \frac{\kappa_i \kappa_j}{2} \sum_{\substack{a \neq b \in V(H_i) \\ c \neq d \in V(H_j)}} t \left(H_i \bigoplus_{(a,b),(c,d)} H_j, W^{G_n} \right) - \left(\sum_{i=1}^q \alpha_i c(H_i, W^{G_n}) \right)^2 \\ &\rightarrow \sum_{q+1 \leq i, j \leq r} \frac{\kappa_i \kappa_j}{2} \sum_{\substack{a \neq b \in V(H_i) \\ c \neq d \in V(H_j)}} t \left(H_i \bigoplus_{(a,b),(c,d)} H_j, W \right) - \left(\sum_{i=1}^q \alpha_i c(H_i, W) \right)^2 = \tilde{\eta}_{\alpha}, \end{aligned} \quad (\text{D.18})$$

as the weak edge-join of 2 simple graphs produces a simple graph. Hence, recalling (D.14), (D.17), and (D.18) gives, $c_{2,n} \rightarrow \eta_{\alpha} + \tilde{\eta}_{\alpha}$. Combining this with (D.15), (D.16), and Proposition B.1, the result in Proposition D.1 follows.

D.2. Proof of Lemma D.2. We begin by recalling the definitions of *cut-distance* and *cut-metric*.

Definition D.1. [55, Chapter 8] The *cut-distance* between two bounded functions $W_1, W_2 : [0, 1]^2 \rightarrow \mathbb{R}$ is

$$\|W_1 - W_2\|_{\square} := \sup_{f, g: [0, 1] \rightarrow [0, 1]} \left| \int_{[0, 1]^2} (W_1(x, y) - W_2(x, y)) f(x) g(y) dx dy \right|.$$

The *cut-metric* between W_1, W_2 is defined as,

$$\delta_{\square}(W_1, W_2) := \inf_{\psi} \|W_1^{\psi} - W_2\|_{\square},$$

with the infimum taken over all measure-preserving bijections $\psi : [0, 1] \rightarrow [0, 1]$, and $W_1^{\psi}(x, y) := W_1(\psi(x), \psi(y))$, for $x, y \in [0, 1]$.

By [55, Lemma 10.16] we know that $\delta_{\square}(W^{G_n}, W) \rightarrow 0$ almost surely.

D.2.1. *Proof of Lemma D.2 (1).* Recall from (D.13),

$$U_{\alpha, G_n}(x, y) := \sum_{i=q+1}^r \alpha_i \left(W_{H_i}^{G_n}(x, y) - c(H_i, W^{G_n}) \right),$$

where $c(H_i, W) = \frac{|V(H_i)|(|V(H_i)|-1)}{2|\text{Aut}(H_i)|} t(H_i, W^{G_n})$. Denote $\nu := \max_{1 \leq i \leq r} |V(H_i)|(|V(H_i)|-1)$. Then, by the counting lemma (see [14, Theorem 3.7]), for $1 \leq i \leq r$,

$$|c(H_i, W^{G_n}) - c(H_i, W)| = \frac{|V(H_i)|(|V(H_i)|-1)}{2|\text{Aut}(H_i)|} |t(H_i, W^{G_n}) - t(H_i, W)| \lesssim_\nu \|W^{G_n} - W\|_\square.$$

Hence, given $f, g : [0, 1] \rightarrow [0, 1]$, by the triangle inequality,

$$\begin{aligned} & \left| \int f(x) (U_{\alpha, G_n}(x, y) - U_\alpha(x, y)) g(y) dx dy \right| \\ & \lesssim_\nu \sum_{i=q+1}^r |\alpha_i| \left(\left| \int f(x) (W_{H_i}^{G_n}(x, y) - W_{H_i}(x, y)) g(y) dx dy \right| + \|W^{G_n} - W\|_\square \right). \end{aligned} \quad (\text{D.19})$$

Now, by a telescoping argument, replacing W^{G_n} with W one at a time, as in the proof of the counting lemma [14, Theorem 3.7], it can be shown that

$$\left| \int f(x) (W_{H_i}^{G_n}(x, y) - W_{H_i}(x, y)) g(y) dx dy \right| \lesssim_\nu \|W^{G_n} - W\|_\square.$$

Hence, from (D.19),

$$\|U_{\alpha, G_n} - U_\alpha\|_\square \lesssim_\nu \|W^{G_n} - W\|_\square \quad (\text{D.20})$$

Again, by the counting lemma (adapted to general bounded functions) we have, for $L \geq 3$,

$$\left| \int U_{\alpha, G_n}^{(L)}(x, x) dx - \int U_\alpha^{(L)}(x, x) dx \right| \lesssim_L \|U_{\alpha, G_n} - U_\alpha\|_\square \lesssim \|W^{G_n} - W\|_\square, \quad (\text{D.21})$$

where the last inequality uses (D.20). Since $\int U_\alpha^{(L)}(x, x) dx$ is invariant to measure preserving transformations of W and $\delta_\square(W^{G_n}, W) \rightarrow 0$ almost surely, as $n \rightarrow \infty$, from (D.21) the result in Lemma D.2 (1) follows.

D.2.2. *Proof of Lemma D.2 (2).* Recall from (D.12),

$$V_{\alpha, G_n}(x) := \sum_{i=1}^q \alpha_i \left[\frac{1}{|\text{Aut}(H_i)|} \sum_{a=1}^{|V(H_i)|} t_a(x, H_i, W^{G_n}) - \frac{|V(H_i)|}{|\text{Aut}(H_i)|} t(H_i, W^{G_n}) \right].$$

Note that $t(H_i, W^{G_n}) \rightarrow t(H_i, W)$, for all $1 \leq i \leq r$, and

$$\lim_{n \rightarrow \infty} \int U_{\alpha, G_n}^{(L)}(x, y) dx dy = \int U_\alpha^{(L)}(x, y) dx dy,$$

by (D.20), almost surely. Hence, to establish Lemma D.2 (2) it suffices to show the following for $1 \leq i, j \leq q$,

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int t_a(x, H_i, W^{G_n}) U_{\alpha, G_n}^{(L)}(x, y) t_b(y, H_j, W^{G_n}) dx dy \\ & = \int t_a(x, H_i, W) U_\alpha^{(L)}(x, y) t_b(y, H_j, W) dx dy, \end{aligned} \quad (\text{D.22})$$

and

$$\lim_{n \rightarrow \infty} \int t_a(x, H_i, W^{G_n}) U_{\alpha, G_n}^{(L)}(x, y) dx dy = \int t_a(x, H_i, W) U_\alpha^{(L)}(x, y) dx dy. \quad (\text{D.23})$$

where $a \in V(H_i)$ and $b \in V(H_j)$.

We will prove (D.22). The proof of (D.23) follows similarly. For this, note that by a telescoping argument,

$$\begin{aligned} \int t_a(x, H_i, W^{G_n})(U_{\alpha, G_n}^{(L)}(x, y) - U_{\alpha}^{(L)}(x, y))t_b(y, H_j, W^{G_n})dxdy &\lesssim \|U_{\alpha, G_n} - U_{\alpha}\|_{\square} \\ &\lesssim \|W^{G_n} - W\|_{\square}, \end{aligned}$$

where the last step uses (D.20). Hence, to prove (D.22) it suffices to show that

$$\lim_{n \rightarrow \infty} \int t_a(x, H_i, W^{G_n})U_{\alpha}^{(L)}(x, y)t_b(y, H_j, W^{G_n})dxdy. \quad (\text{D.24})$$

Consider the functions (not necessarily symmetric) $B_n(x, y) := t_a(x, H_i, W^{G_n})t_b(y, H_j, W^{G_n})$ and $B(x, y) := t_a(x, H_i, W)t_b(y, H_j, W)$. By a telescoping argument it can be shown that

$$\|B_n - B\|_{\square} \lesssim \|W^{G_n} - W\|_{\square}.$$

The result in (D.24) then follows from [55, Lemma 8.22].

APPENDIX E. PROOFS FROM SECTION 5

E.1. Proof of Proposition 5.1. To prove Proposition 5.1 we first replace with $R(H, G_n)$ by

$$R(H, W^{G_n}) = \sum_{1 \leq a, b \leq |V(H)|} t\left(H \bigoplus_{a, b} H, W^{G_n}\right) - |V(H)|^2 t(H, W^{G_n})^2,$$

where $t(\cdot, W^{G_n})$ is defined in (4.1). For a finite subgraph $F = (V(F), E(F))$, recalling (4.1) and (4.2) notice,

$$\begin{aligned} |\hat{t}(F, G_n) - t(F, W^{G_n})| &= \left| \frac{1}{n^{|V(F)|}} \sum_{\mathbf{s} \in [n]^{|V(F)|}} \prod_{(i, j) \in E(F)} w_{s_i s_j} - \frac{1}{(n)^{|V(F)|}} \sum_{\mathbf{s} \in ([n])^{|V(F)|}} \prod_{(i, j) \in E(F)} w_{s_i s_j} \right| \\ &\leq \frac{1}{n^{|V(F)|}} \left| \sum_{\mathbf{s} \in [n]^{|V(F)|} \setminus ([n])^{|V(F)|}} \prod_{(i, j) \in E(F)} w_{s_i s_j} \right| + O\left(\frac{1}{n}\right) = O\left(\frac{1}{n}\right). \end{aligned} \quad (\text{E.1})$$

This implies,

$$|R(H, G_n) - R(H, W^{G_n})| = O\left(\frac{1}{n}\right). \quad (\text{E.2})$$

Notice that, by definition, $R(H, W) \geq 0$ (see (5.2)) for any graphon W and in particular for the empirical graphon W^{G_n} . Now, we consider the following 2 cases:

- *W is H-regular:* Recalling that $R(H, W^{G_n}) \geq 0$ it is now enough to show that

$$\mathbb{E}R(H, W^{G_n}) = O(1/n). \quad (\text{E.3})$$

Towards that, recalling (5.3) note that,

$$\begin{aligned} &\mathbb{E}[R(H, W^{G_n})] \\ &= \sum_{1 \leq a, b \leq |V(H)|} \mathbb{E}\left[t\left(H \bigoplus_{a, b} H, W^{G_n}\right)\right] - |V(H)|^2 \mathbb{E}\left[t(H \sqcup H, W^{G_n})\right], \end{aligned} \quad (\text{E.4})$$

where $H \sqcup H$ is the disjoint union of two copies of H . Lemma 2.1 and Lemma 2.4 in [56] implies that

$$\begin{aligned} \mathbb{E} \left[t(H \sqcup H, W^{G_n}) \right] &\geq t(H \sqcup H, W) - \frac{1}{n} \binom{2|V(H)|}{2} \\ &= t(H, W)^2 - \frac{1}{n} \binom{2|V(H)|}{2} \end{aligned}$$

and

$$\begin{aligned} &\sum_{1 \leq a, b \leq |V(H)|} \mathbb{E} \left[t \left(H \oplus_{a,b} H, W^{G_n} \right) \right] \\ &\leq \sum_{1 \leq a, b \leq |V(H)|} \left[t \left(H \oplus_{a,b} H, W \right) + \frac{1}{n} \binom{2|V(H)| - 1}{2} \right]. \end{aligned}$$

Substituting these bounds in (E.4) give,

$$\begin{aligned} \mathbb{E} R(H, W^{G_n}) &\leq \sum_{a,b=1}^{|V(H)|} t \left(H \oplus_{a,b} H, W \right) - |V(H)|^2 t(H, W)^2 + O(1/n) \\ &= R(H, W) + O(1/n), \end{aligned}$$

where the last equality follows by recalling the definition of $R(H, W)$ from (5.2). The proof of Proposition 5.1 (1) is now complete by noticing that $R(H, W) = 0$ when W is H -regular.

- W is H -irregular: From Corollary 10.4 in [55] we know that $R(H, W^{G_n}) \xrightarrow{P} R(H, W)$. Since $R(H, W) > 0$ whenever W is H -irregular, this implies $\sqrt{n} R(H, W^{G_n}) \xrightarrow{P} \infty$. This completes the proof Proposition 5.1 (2).

Remark E.1. Note that (E.3) and (E.2) implies $R(H, G_n) = O_P(n)$, when W is H -regular. Hence, $a_n R(H, G_n) \xrightarrow{P} 0$, for any sequence $\{a_n\}_{n \geq 1}$ such that $a_n \rightarrow \infty$ and $a_n/n \rightarrow 0$. While choosing $a_n = \sqrt{n}$ shows Proposition 5.1 (1), the results in Section 5 would continue to hold whenever $a_n/n \rightarrow 0$.

E.2. Proof of Theorem 5.1. Without loss of generality assume that W is irregular with respect to $H_1, H_2, \dots, H_q$ and regular with respect to $H_{q+1}, H_{q+2}, \dots, H_r$. To proceed with the proof we first show that the distribution of $\tilde{\mathbf{Z}}(\mathcal{H}, G_n)$ (recall (5.7)) converges to $\mathbf{Z}(\mathcal{H}, W)$.

Lemma E.1. *Let $\tilde{\mathbf{Z}}(\mathcal{H}, G_n)$ and $\mathbf{Z}(\mathcal{H}, W)$ be as defined in (5.7) and (2.12), respectively. Then, as $n \rightarrow \infty$,*

$$\tilde{\mathbf{Z}}(\mathcal{H}, G_n) \xrightarrow{D} \mathbf{Z}(\mathcal{H}, W). \quad (\text{E.5})$$

Proof. Recall that $S(\mathcal{H}, G_n) = \{1 \leq i \leq r : \sqrt{n} R(H, G_n) > 1\}$ is the set of indices where the hypothesis of H_i -regularity is rejected. Define the event

$$\mathcal{D}_n := \{S(\mathcal{H}, G_n) = \{1, 2, \dots, q\}\}. \quad (\text{E.6})$$

Now, by Proposition 5.1 notice that,

$$\mathbb{P}(\mathcal{D}_n^c) \leq \sum_{i=1}^q \mathbb{P}(i \notin S(\mathcal{H}, G_n)) + \sum_{i=q+1}^r \mathbb{P}(i \in S(\mathcal{H}, G_n))$$

$$= \sum_{i=1}^q \mathbb{P}(\sqrt{n}R(H, G_n) \leq 1) + \sum_{i=q+1}^r \mathbb{P}(\sqrt{n}R(H, G_n) > 1) = o(1), \quad (\text{E.7})$$

since W is irregular with respect to $H_1, H_2, \dots, H_q$ and regular with respect to $H_{q+1}, H_{q+2}, \dots, H_r$. Hence, to prove (E.5) it is enough to show that the characteristic functions converge. For this, note that for any $\mathbf{t} \in \mathbb{R}^r$,

$$\mathbb{E} \left[e^{i\mathbf{t}^\top \tilde{\mathbf{Z}}(\mathcal{H}, G_n)} \right] = \mathbb{E} \left[e^{i\mathbf{t}^\top \tilde{\mathbf{Z}}(\mathcal{H}, G_n)} \mathbf{1}_{\{\mathcal{D}_n\}} \right] + o(1).$$

Note that on the event $\mathcal{D}_n$, $\tilde{\mathbf{Z}}_n(\mathcal{H}, G_n) = \mathbf{Z}(\mathcal{H}, G_n)$ (recall (2.10)). Therefore, (E.7) and Theorem 2.1 gives,

$$\begin{aligned} \mathbb{E} \left[e^{i\mathbf{t}^\top \tilde{\mathbf{Z}}(\mathcal{H}, G_n)} \right] &= \mathbb{E} \left[e^{i\mathbf{t}^\top \mathbf{Z}(\mathcal{H}, G_n)} \mathbf{1}_{\{\mathcal{D}_n\}} \right] + o(1) = \mathbb{E} \left[e^{i\mathbf{t}^\top \mathbf{Z}(\mathcal{H}, G_n)} \right] + o(1) \\ &\rightarrow \mathbb{E} \left[e^{i\mathbf{t}^\top \mathbf{Z}(\mathcal{H}, W)} \right]. \end{aligned}$$

This completes the proof of Lemma E.1. $\square$

Now, we show that conditional on G_n the distribution of $\mathbf{Q}(\mathcal{H}, G_n)$ converges to $\mathbf{Z}(\mathcal{H}, W)$ as well.

Lemma E.2. *Let $\mathbf{Q}(\mathcal{H}, G_n)$ and $\mathbf{Z}(\mathcal{H}, W)$ be as defined in (5.5) and (2.12), respectively. Then, as $n \rightarrow \infty$,*

$$\mathbf{Q}(\mathcal{H}, G_n) \mid G_n \xrightarrow{D} \mathbf{Z}(\mathcal{H}, W),$$

in probability.

Proof. Recall the event $\mathcal{D}_n$ from (E.6). It follows from (E.7) that,

$$\mathbf{1}_{\{\mathcal{D}_n^c\}} = o_P(1). \quad (\text{E.8})$$

Then for any $\mathbf{t} \in \mathbb{R}^r$,

$$\mathbb{E} \left[e^{i\mathbf{t}^\top \mathbf{Q}(\mathcal{H}, G_n)} \mid G_n \right] = \mathbb{E} \left[e^{i\mathbf{t}^\top \mathbf{Q}(\mathcal{H}, G_n)} \mid G_n \right] \mathbf{1}_{\{\mathcal{D}_n\}} + o_P(1).$$

As before, on the event $\mathcal{D}_n$, we have $\mathbf{Q}(\mathcal{H}, G_n) = \hat{\mathbf{Z}}(\mathcal{H}, G_n)$, where $\hat{\mathbf{Z}}(\mathcal{H}, G_n)$ is defined in (4.7). Hence, by (E.8) and Theorem 4.1,

$$\begin{aligned} \mathbb{E} \left[e^{i\mathbf{t}^\top \mathbf{Q}(\mathcal{H}, G_n)} \mid G_n \right] &= \mathbb{E} \left[e^{i\mathbf{t}^\top \hat{\mathbf{Z}}(\mathcal{H}, G_n)} \mid G_n \right] \mathbf{1}_{\{\mathcal{D}_n\}} + o_P(1) = \mathbb{E} \left[e^{i\mathbf{t}^\top \hat{\mathbf{Z}}(\mathcal{H}, G_n)} \mid G_n \right] + o_P(1) \\ &\xrightarrow{P} \mathbb{E} \left[e^{i\mathbf{t}^\top \mathbf{Z}(\mathcal{H}, W)} \right]. \end{aligned}$$

This completes the proof of Lemma E.2. $\square$

We now proceed to complete the proof of Theorem 5.1. To this end, using the continuous mapping theorem along with Lemma E.1 and Lemma E.2 gives,

$$\left\| \tilde{\mathbf{Z}}(\mathcal{H}, G_n) \right\|_2 \xrightarrow{D} \left\| \mathbf{Z}(\mathcal{H}, W) \right\|_2 \text{ and } \left\| \mathbf{Q}(\mathcal{H}, G_n) \right\|_2 \mid G_n \xrightarrow{D} \left\| \mathbf{Z}(\mathcal{H}, W) \right\|_2 \text{ in probability.}$$

Hence, recalling (5.8) and by Polyá's Theorem,

$$\mathbb{P}(C(\mathcal{H}, G_n)) = \mathbb{P} \left(\left\| \tilde{\mathbf{Z}}(\mathcal{H}, G_n) \right\|_2 \leq \hat{q}_{1-\alpha, \mathcal{H}, G_n} \right) \rightarrow 1 - \alpha,$$

where $\hat{q}_{1-\alpha, \mathcal{H}, G_n}$ is the $(1 - \alpha)$ -quantile of the distribution of $\left\| \mathbf{Q}(\mathcal{H}, G_n) \right\|_2 \mid G_n$. $\square$

APPENDIX F. PROOFS FROM SECTION 6

This section is organized as follows: In Section F.1 we prove Proposition 6.1. Proposition 6.2 is proved in Section F.2. In Section F.3 we derive the distribution of $\hat{f}(G_n)$ under the alternative.

F.1. Proof of Proposition 6.1. For notational convenience define

$$T_n^{K_2} := \frac{1}{2}\hat{t}(K_2, G_n) \quad \text{and} \quad T_n^{C_4} := \frac{1}{8}\hat{t}(C_4, G_n),$$

where $\hat{t}(H, G_n)$ is defined in (1.4). Recalling (6.2), note that

$$\hat{f}(G_n) = h(T_n^{K_2}, T_n^{C_4}), \quad (\text{F.1})$$

where $h(x, y) = 16x^4 - 8y$.

We begin by deriving the joint distribution of $(T_n^{K_2}, T_n^{C_4})$ under H_0 as in (6.1). Note that under H_0 , the G_n is distributed as an Erdős-Rényi random graph $G(n, p)$, for $p = t(K_2, W) \in (0, 1)$. Hence, recalling Example 3.2, in particular from (3.3) and (3.4) we get,

$$n \begin{pmatrix} T_n^{K_2} - \frac{1}{2}p \\ T_n^{C_4} - \frac{1}{8}p^4 \end{pmatrix} \xrightarrow{D} N_2(0, \Sigma), \quad \text{where } \Sigma = \frac{1}{2}p(1-p) \begin{pmatrix} 1 & p^3 \\ p^3 & p^6 \end{pmatrix}. \quad (\text{F.2})$$

Now, a Taylor expansion of the function h around the point $(\frac{1}{2}p, \frac{1}{8}p^4)$ gives,

$$h(T_n^{K_2}, T_n^{C_4}) - h(\frac{1}{2}p, \frac{1}{8}p^4) = \nabla h(\frac{1}{2}p, \frac{1}{8}p^4)^\top \mathbf{T}_n + \mathbf{T}_n^\top \nabla^2 h(\frac{1}{2}p, \frac{1}{8}p^4) \mathbf{T}_n + o_P(\|\mathbf{T}_n\|^2),$$

where

$$\mathbf{T}_n := (T_n^{K_2} - \frac{1}{2}p, T_n^{C_4} - \frac{1}{8}p^4)^\top, \quad (\text{F.3})$$

and $\nabla h, \nabla^2 h$ denote the gradient and hessian of h evaluated at the corresponding points. By definition, $h(\frac{1}{2}p, \frac{1}{8}p^4) = 0$. Hence, a direct computation of $\nabla^2 h(\frac{1}{2}p, \frac{1}{8}p^4)$ along with the convergence in (F.2) gives,

$$h(T_n^{K_2}, T_n^{C_4}) = \nabla h(\frac{1}{2}p, \frac{1}{8}p^4)^\top \mathbf{T}_n + O_P\left(\frac{1}{n^2}\right). \quad (\text{F.4})$$

Since by definition $\hat{f}(G_n) = h(T_n^{K_2}, T_n^{C_4})$ (recall (F.1)), the result in Proposition 6.1 follows from (F.4) and the following lemma:

Lemma F.1. *Under H_0 as in (6.1),*

$$n^{\frac{3}{2}} \nabla h(\frac{1}{2}p, \frac{1}{8}p^4)^\top \mathbf{T}_n \xrightarrow{D} N(0, 32p^6(1-p)^2),$$

where $\mathbf{T}_n$ is defined in (F.3).

F.1.1. Proof of Lemma F.1. Notice that $\nabla h(\frac{1}{2}p, \frac{1}{8}p^4)^\top = (8p^3, -8)$. Hence, recalling the definitions of $\hat{t}(K_2, G_n)$ and $\hat{t}(C_4, G_n)$ from (1.4) gives,

$$\nabla h(\frac{1}{2}p, \frac{1}{8}p^4)^\top \mathbf{T}_n = \frac{8p^3}{(n)_2} \left(X(K_2, G_n) - \frac{(n)_2}{2}p \right) - \frac{8}{(n)_4} \left(X(C_4, G_n) - \frac{(n)_4}{8}p^4 \right), \quad (\text{F.5})$$

where $(n)_2 = n(n-1)$ and $(n)_4 = n(n-1)(n-2)(n-3)$. We will now compute the orthogonal decomposition of (F.5) using the framework described in Section A.1.

To this end, recall the definition of $S_{n,\cdot}(\cdot)$ from (A.10). Then using (A.16) we can rewrite (F.5) as,

$$\nabla h(\frac{1}{2}p, \frac{1}{8}p^4)^\top \mathbf{T}_n = \frac{8p^3}{(n)_2} S_{n,2}(f) - \frac{8}{(n)_4} S_{n,4}(g), \quad (\text{F.6})$$

where the functions f and g are defined as follows:

- $f(Y_{12}) = \mathbf{1}\{Y_{12} \leq p\} - p$.

- $g(\{Y_{ij} : (i, j) \in E(K_4)\}) = \sum_{i=1}^3 \prod_{e \in E(G_i)} \mathbf{1}\{Y_e \leq p\} - 3p^4$, where G_1, G_2, G_3 are cycles of length four with edges $E(G_1) = \{(1, 2), (2, 3), (3, 4), (4, 1)\}$, $E(G_2) = \{(1, 2), (2, 4), (4, 3), (3, 1)\}$ and $E(G_3) = \{(1, 3), (3, 2), (2, 4), (4, 1)\}$, respectively, and $\{Y_{ij} = Y_{ji} : 1 \leq i < j \leq 4\}$ are independently generated from $U[0, 1]$.

Now, recalling (A.13) and following [45, Example 2] we get,

$$S_{n,2}(f) = \frac{1}{2} \tilde{S}_{n,2}(f_{K_2}), \quad (\text{F.7})$$

and

$$S_{n,4}(g) = \frac{1}{4} \tilde{S}_{n,4}(g_{K_2}) + \frac{1}{6} \tilde{S}_{n,4}(g_{K_3}) + \frac{1}{2} \tilde{S}_{n,4}(g_{K_{1,2}}) + \sum_{G \in \Gamma_4} \frac{\tilde{S}_{n,4}(g_G)}{|\text{Aut}(G)|}. \quad (\text{F.8})$$

where f_G and g_G are the projections of f and g on the subspace M_G (recall (A.3)). A direct counting argument gives, $\tilde{S}_{n,2}(f_{K_2}) = \frac{(n-4)!}{(n-2)!} \tilde{S}_{n,4}(f_{K_2})$. Hence, from (F.7) and (F.8), the RHS of (F.6) can be rewritten as,

$$\begin{aligned} & \nabla h(\tfrac{1}{2}p, \tfrac{1}{8}p^4)^\top \mathbf{T}_n \\ &= \frac{8}{(n)_4} \left(\tilde{S}_{n,4} \left(\tfrac{1}{2}p^3 f_{K_2} - \tfrac{1}{4}g_{K_2} \right) - \frac{1}{6} \tilde{S}_{n,4}(g_{K_3}) - \frac{1}{2} \tilde{S}_{n,4}(g_{K_{1,2}}) - \sum_{G \in \Gamma_4} \frac{\tilde{S}_{n,4}(g_G)}{|\text{Aut}(G)|} \right). \end{aligned} \quad (\text{F.9})$$

Using (A.9) we now compute the different projections. To begin with, note that

$$f_{K_2} = (\mathbf{1}\{Y_{12} \leq p\} - p) \text{ and } g_{K_2} = 2p^3(\mathbf{1}\{Y_{12} \leq p\} - p).$$

This shows that $\frac{1}{2}p^3 f_{K_2} - \frac{1}{4}g_{K_2} = 0$. Also, by computations similar to those in (A.26) and (A.27), it can shown that $g_{K_3} = 0$. Further, by [45, Lemma 4],

$$\frac{8}{(n)_4} \sum_{G \in \Gamma_4} \frac{\tilde{S}_{n,4}(g_G)}{|\text{Aut}(G)|} = o_P(n^{-\frac{3}{2}}).$$

Hence, (F.9) can be simplified as follows,

$$\nabla h(\tfrac{1}{2}p, \tfrac{1}{8}p^4)^\top \mathbf{T}_n = \frac{4}{(n)_4} \tilde{S}_{n,4}(g_{K_{1,2}}) + o_P(n^{-\frac{3}{2}}). \quad (\text{F.10})$$

Once again, computing the projection of g on $M_{K_{1,2}}$ using (A.9) we get,

$$g_{K_{1,2}} = p^2(\mathbf{1}\{Y_{12} \leq p\} - p)(\mathbf{1}\{Y_{13} \leq p\} - p). \quad (\text{F.11})$$

Using (F.10) and (F.11) together with the distributional convergence result in [45, Lemma 7], we have

$$n^{\frac{3}{2}} \nabla h(\tfrac{1}{2}p, \tfrac{1}{8}p^4)^\top \mathbf{T}_n = n^{\frac{3}{2}} \frac{4}{(n)_4} \tilde{S}_{n,4}(g_{K_{1,2}}) + o_P(1) \xrightarrow{D} N(0, 32p^6(1-p)^2).$$

This completes the proof of Lemma F.1. $\square$

F.2. Proof of Proposition 6.2. Recall the definition of $\hat{f}(G_n)$ from (6.2). Then, since $\hat{t}(H, G_n) \xrightarrow{P} t(H, W)$ for every fixed graph H , it follows that

$$\hat{f}(G_n) = \hat{t}(K_2, G_n)^4 - \hat{t}(C_4, G_n) \xrightarrow{P} f(W) = t(K_2, W)^4 - t(C_4, W).$$

The quasi-randomness result of Chung, Graham, and Wilson [20, Theorem 1], formulated in terms of graphons (see [55, Section 11.8]), implies that $f(W) \neq 0$, for any graphon W that is not constant almost everywhere. Hence, under H_1 ,

$$\frac{\hat{f}(G_n)}{4\sqrt{2} \hat{t}(K_2, G_n)^3 (1 - \hat{t}(K_2, G_n))} \xrightarrow{P} \frac{f(W)}{4\sqrt{2} t(K_2, W)^3 (1 - t(K_2, W))} > 0.$$

Hence, recalling (6.4), $T_n \xrightarrow{P} \infty$, which now completes the proof.

F.3. Distribution of $\hat{f}(G_n)$ under the Alternative. In this section we apply Theorem 2.1 to derive the limiting of $\hat{f}(G_n)$ under the alternative.

Proposition F.1. *Suppose $|t(K_2, W)^4 - t(C_4, W)| > 0$. Then the following hold:*

- *If W is irregular with respect to K_2 and C_4 , then*

$$\sqrt{n}(\hat{f}(G_n) - f(W)) \xrightarrow{D} N(0, \tau_1^2),$$

where τ_1^2 is defined in (F.12).

- *If W is irregular with respect to K_2 and regular with respect to C_4 , then*

$$\sqrt{n}(\hat{f}(G_n) - f(W)) \xrightarrow{D} N(0, \tau_2^2),$$

where τ_2^2 is defined in (F.13).

- *If W is regular with respect to K_2 and irregular with respect to C_4 , then*

$$\sqrt{n}(\hat{f}(G_n) - f(W)) \xrightarrow{D} N(0, \tau_3^2),$$

where τ_3^2 is defined in (F.14).

- *If W is regular with respect to K_2 and C_4 , then*

$$n(\hat{f}(G_n) - f(W)) \xrightarrow{D} Z,$$

where the random variable Z is defined in (F.16).

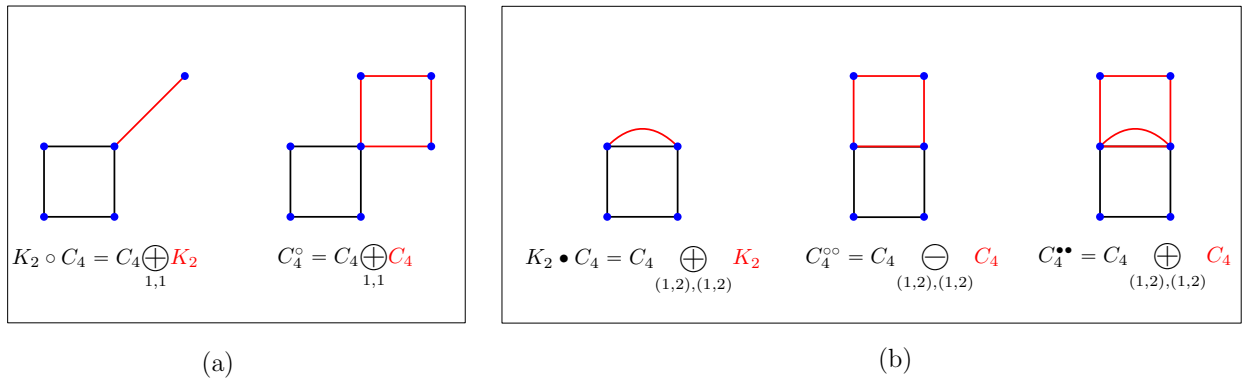


FIGURE 9. Graphs obtained from (a) vertex join operations and (b) edge join operations, between a copy of K_2 and a copy of C_4 .

Proof. Fix a graphon W such that $|t(K_2, W)^4 - t(C_4, W)| > 0$. We consider the 4 cases separately, depending on whether or not W is K_2 or C_4 -regular.

Case 1: W is irregular with respect to K_2 and C_4 . In this case, Corollary 2.2 gives,

$$\mathbf{Z}_n = \sqrt{n} \begin{pmatrix} \hat{t}(K_2, G_n) - t(K_2, W) \\ \hat{t}(C_4, G_n) - t(C_4, W) \end{pmatrix} \xrightarrow{D} N_2 \left(\mathbf{0}, \begin{pmatrix} \tau_{11} & \tau_{12} \\ \tau_{21} & \tau_{22} \end{pmatrix} \right),$$

where

- $\tau_{11} = t(K_{1,2}, W) - t(K_2, W)^2$,
- $\tau_{22} = \frac{1}{4} [t(C_4^\circ, W) - t(C_4, W)^2]$, where C_4° is the graph obtained by the vertex join of 2 copies of C_4 (as shown in Figure 9(a)).
- $\tau_{12} = \tau_{21} = \frac{1}{2} (t(C_4 \circ K_2) - t(C_4, W)t(K_2, W))$, where $C_4 \circ K_2$ is the graph obtained the vertex join of C_4 and K_2 (as shown in Figure 9(a)).

Note that $\sqrt{n}(\hat{f}(G_n) - f(W)) = \mathbf{1}^\top \mathbf{Z}_n$. Hence, by the continuous mapping theorem,

$$\sqrt{n}(\hat{f}(G_n) - f(W)) \xrightarrow{D} N(0, \sigma_1^2),$$

with

$$\tau_1^2 := \tau_{11} + \tau_{22} - 2\tau_{12}. \quad (\text{F.12})$$

Case 2: W is irregular with respect to K_2 and regular with respect to C_4 . In this case, $\sqrt{n}(\hat{t}(K_2, W^{G_n}) - t(K_2, W))$ has a non-degenerate Gaussian limit, but $\sqrt{n}(\hat{t}(C_4, G_n) - t(C_4, W))$ has a degenerate limit. In particular, from Theorem 2.1 we know that

$$\mathbf{Z}_n = \sqrt{n} \begin{pmatrix} \hat{t}(K_2, G_n) - t(K_2, W) \\ \hat{t}(C_4, G_n) - t(C_4, W) \end{pmatrix} \xrightarrow{D} \begin{pmatrix} G_1 \\ 0 \end{pmatrix},$$

where $G_1 \sim N(0, t(K_{1,2}, W) - t(K_2, W)^2)$. Since $\sqrt{n}(\hat{f}(G_n) - f(W)) = \mathbf{1}^\top \mathbf{Z}_n$, this implies,

$$\sqrt{n}(\hat{f}(G_n) - f(W)) \xrightarrow{D} N(0, \tau_2^2),$$

where

$$\tau_2^2 := t(K_{1,2}, W) - t(K_2, W)^2. \quad (\text{F.13})$$

Case 3: W is regular with respect to K_2 and irregular with respect to C_4 . In this case, $\sqrt{n}(\hat{t}(K_2, G_n) - t(K_2, W))$ has a degenerate limit, but $\sqrt{n}(\hat{t}(C_4, G_n) - t(C_4, W))$ has a non-degenerate Gaussian limit. Hence, applying Theorem 2.1 we have,

$$\mathbf{Z}_n = \sqrt{n} \begin{pmatrix} \hat{t}(K_2, G_n) - t(K_2, W) \\ \hat{t}(C_4, G_n) - t(C_4, W) \end{pmatrix} \xrightarrow{D} \begin{pmatrix} 0 \\ G_2 \end{pmatrix},$$

where $G_2 \sim N(0, \frac{1}{4} (t(C_4^\circ, W) - t(C_4, W)^2))$. Taking inner product of $\mathbf{Z}_n$ with $\mathbf{1}$, then gives

$$\sqrt{n}(\hat{f}(G_n) - f(W)) \xrightarrow{D} N(0, \tau_3^2),$$

where

$$\tau_3^2 := \frac{1}{4} (t(C_4^\circ, W) - t(C_4, W)^2). \quad (\text{F.14})$$

Case 4: W is regular with respect to both K_2 and C_4 . Then from Theorem 2.1,

$$\mathbf{Z}_n = n \begin{pmatrix} \hat{t}(K_2, G_n) - t(K_2, W) \\ \hat{t}(C_4, G_n) - t(C_4, W) \end{pmatrix} \xrightarrow{D} \mathbf{Z} := \begin{pmatrix} G_1 \\ G_2 \end{pmatrix} + \begin{pmatrix} \iint (W(x, y) - \frac{1}{2}t(K_2, W)) \, dB_x dB_y \\ \iint (W_{C_4}(x, y) - \frac{3}{4}t(C_4, W)) \, dB_x dB_y \end{pmatrix}, \quad (\text{F.15})$$

where the Brownian motion $\{B_t\}_{t \in [0,1]}$ and $(G_1, G_2)^\top \sim N_2(\mathbf{0}, \Sigma)$ are independent, and the matrix Σ is given by:

$$\Sigma = \frac{1}{2} \begin{pmatrix} t(K_2, W) - t(C_2, W) & t(C_4, W) - t(K_2 \bullet C_4, W) \\ t(C_4, W) - t(K_2 \bullet C_4, W) & t(C_4^{\circ\circ}, W) - t(C_4^{\bullet\bullet}, W) \end{pmatrix}$$

In this case, as shown in Figure 9(b), $K_2 \bullet C_4$ is the graph obtained by the strong edge join of K_2 and C_4 , $C_4^{\circ\circ}$ is the graph obtained by the weak edge join of 2 copies of C_4 , and $C_4^{\bullet\bullet}$ is the graph obtained by the strong edge join of 2 copies of C_4 .

Therefore,

$$\sqrt{n}(\hat{f}(G_n) - f(W)) = \mathbf{1}^\top \mathbf{Z}_n \xrightarrow{D} \mathbf{1}^\top \mathbf{Z} := Z, \quad (\text{F.16})$$

where $\mathbf{Z}$ as defined in (F.15). □

APPENDIX G. MULTIPLE WEINER-ITÔ STOCHASTIC INTEGRALS

In this section we recall the basic properties of multiple Wiener-Itô stochastic integrals as presented in [41]. To begin with, let $\{B_t\}_{t \in [0,1]}$ be the standard Brownian motion in $[0, 1]$. We interpret the Brownian motion as a stochastic measure on $([0, 1], \mathcal{B}([0, 1]))$, where $\mathcal{B}([0, 1])$ is the sigma-algebra generated by open sets of $[0, 1]$. Specifically, suppose $\{B(A) : A \in \mathcal{B}([0, 1])\}$ is a collection of random variables defined on a common probability space $(\Omega, \mathcal{F}, \mu)$ such that

- $B(A) \sim N(0, \lambda(A))$, for all $A \in \mathcal{B}([0, 1])$, where $\lambda(A)$ is the Lebesgue measure of A .
- For any finite collection of disjoint sets $A_1, \dots, A_t \in \mathcal{B}(\mathcal{X})$, the random variables $\{B(A_1), B(A_2), \dots, B(A_t)\}$ are independent and

$$B\left(\bigcup_{s=1}^t A_s\right) = \sum_{s=1}^t B(A_s).$$

For $d \geq 1$, denote by $L^2([0, 1]^d)$ the space of measurable functions $f : \mathcal{X}^d \rightarrow \mathbb{R}$ such that

$$\|f\|_2^2 := \int_{[0,1]^d} |f(x_1, x_2, \dots, x_d)|^2 dx_1 dx_2 \dots, dx_d < \infty.$$

Define $\mathcal{E}_d \subseteq L^2([0, 1]^d)$ as the set of all elementary functions having the form

$$f(t_1, t_2, \dots, t_d) = \sum_{1 \leq i_1, i_2, \dots, i_d \leq m} a_{i_1, i_2, \dots, i_d} \mathbf{1}\{(t_1, t_2, \dots, t_d) \in A_{i_1} \times \dots \times A_{i_d}\}, \quad (\text{G.1})$$

where $A_1, A_2, \dots, A_m \in [0, 1]$ are measurable sets which are pairwise disjoint and $a_{i_1, i_2, \dots, i_d}$ is zero if two indices are equal. The multiple Wiener-Itô integral for functions in $\mathcal{E}_d$ is defined as follows:

Definition G.1. (Multiple Wiener-Itô integral for elementary functions) The d -dimensional Wiener-Itô stochastic integral, with respect to the standard Brownian motion $\{B_t\}_{t \in [0,1]}$, for the function $f \in \mathcal{E}_d$ in (G.1) is defined as

$$I_d(f) := \int_{[0,1]^d} f(x_1, x_2, \dots, x_d) \prod_{a=1}^d dB(x_a) := \sum_{1 \leq i_1, i_2, \dots, i_d \leq m} a_{i_1, i_2, \dots, i_d} B(A_{i_1}) \times \dots \times B(A_{i_d}).$$

The multiple Wiener-Itô integral for elementary functions satisfies the following two properties [41]:

- (Boundedness) For $f \in \mathcal{E}_d$, $\mathbb{E}[I_d(f)^2] \leq d! \|f\|_2^2 < \infty$.
- (Linearity) For $f, g \in \mathcal{E}_d$, $I_d(f + g) \stackrel{a.s.}{=} I_d(f) + I_d(g)$.

This shows that I_d is a bounded linear operator from $\mathcal{E}_d$ to $L^2(\Omega, \mathcal{F}, \mu)$, the collection of square-integrable random variables defined on $(\Omega, \mathcal{F}, \mu)$. Since $\mathcal{E}_d$ is dense in $L^2([0, 1]^d, \mathcal{B}([0, 1]^d), \lambda^d)$ (by [41, Theorem 2.1]), using the BLT theorem (see [71, Theorem I.7]) I_d can be uniquely extended to $L^2([0, 1]^d, \mathcal{B}([0, 1]^d), \lambda^d)$ by taking limits. (Here, λ^d denotes the Lebesgue measure on $[0, 1]^d$.) This leads to the following definition:

Definition G.2. (Multiple Wiener-Itô integral for general L^2 -functions) The d -dimensional Wiener-Itô stochastic integral, with respect to the standard Brownian motion $\{B_t\}_{t \in [0, 1]}$, for a function $f \in L^2([0, 1]^d)$ is defined as the L^2 limit of the sequence $\{I_d(f_n)\}_{n \geq 1}$, where $\{f_n\}_{n \geq 1}$ is a sequence such that $f_n \in \mathcal{E}_d$ with $\lim_{n \rightarrow \infty} \|f_n - f\|_2 = 0$. This is denoted by:

$$I_d(f) := \int_{[0, 1]^d} f(x_1, x_2, \dots, x_d) \prod_{a=1}^d dB(x_a). \quad (\text{G.2})$$

As in the case of elementary functions, it can be easily checked that $I_d(f)$ satisfies the following properties:

- (Boundedness) For $f \in L^2([0, 1]^d)$, $\mathbb{E}[I_d(f)^2] \leq d! \|f\|_2^2 < \infty$.
- (Linearity) For $f, g \in L^2([0, 1]^d)$, $I_d(f + g) \stackrel{\text{a.s.}}{=} I_d(f) + I_d(g)$.

It is also important to note that multiple Wiener-Itô integrals do not behave like classical (non-stochastic) integrals with respect to product measures, since by definition diagonal sets do not contribute to the stochastic integral. Nevertheless, one can express the multiple Wiener-Itô integral for a product function in terms of univariate stochastic integrals using the Wick product (cf. [44, Theorem 7.26]). In the bivariate case, with 2 functions $f, g \in L^2([0, 1]^2)$, this simplifies to

$$\int_{[0, 1]} \int_{[0, 1]} f(x)g(y)dB(x)dB(y) = \int_{[0, 1]} f(x)dB(x) \int_{[0, 1]} g(y)dB(y) - \int_{[0, 1]} f(x)g(x)dx. \quad (\text{G.3})$$

Another important property is that one can interchange stochastic integrals with infinite sums over an orthonormal and symmetric set of functions, as shown in the following result:

Proposition G.1. Let $f \in L^2([0, 1]^d)$ and $\{\varphi_s\}_{s \geq 1}$ is an orthonormal and symmetric set of functions in $L^2([0, 1]^d)$. Suppose there exists constants $\{\alpha_s\}_{s \geq 1}$ such that

$$f = \sum_{s \geq 1} \alpha_s \varphi_s,$$

is well-defined. Then

$$I_d(f) \stackrel{\text{a.s.}}{=} \sum_{s \geq 1} \alpha_s I_d(\varphi_s). \quad (\text{G.4})$$

Proof. For $N \geq 1$, define the truncated the truncated version of f :

$$f_N := \sum_{s=1}^N \alpha_s \varphi_s.$$

By Bessel's inequality, $\sum_{s \geq 1} \alpha_s^2 \leq \|f\|_2^2 < \infty$. Then by [49, Lemma 6.8] along with the orthonormality of $\{\varphi_s\}_{s \geq 1}$ gives,

$$\|f - f_N\|_2^2 = o(1).$$

By the linearity of stochastic integrals we know that $I_d(f_N) = \sum_{s=1}^N \alpha_s I_d(\varphi_s)$. To complete the proof it is now enough to show that both RHS and LHS of (G.4) are L^2 limits of $I_d(f_N)$, as

$N \rightarrow \infty$. For this, using the boundedness property of stochastic integrals note that,

$$\mathbb{E} \left[(I_d(f) - I_d(f_N))^2 \right] = \mathbb{E} [I_d(f - f_N)^2] \leq d! \|f - f_N\|_2^2 = o(1).$$

This shows that the LHS of (G.4) is the L^2 limit of $I_d(f_N)$, as $N \rightarrow \infty$. For the RHS note that by definition the functions $\{\varphi_s\}_{s \geq 1}$ are symmetric in their arguments. Hence, by [44, Theorem 7.29], for all $s \geq 1$, $I_d(\varphi_s) = d! J_d(\varphi_s)$, where the operator $J_d(\cdot)$ is defined as,

$$J_d(g) := \int_{D_d} g(x_1, x_2, \dots, x_d) dB_{x_1} \cdots dB_{x_d},$$

for $g \in L^2(D_d)$, with

$$D_d := \left\{ (x_1, \dots, x_d) \in [0, 1]^d : 0 < x_1 < x_2 < \cdots < x_d < 1 \right\}.$$

By Theorem 7.6 and Theorem 7.3 from [44] we know that J_d is an isometry. Hence,

$$\mathbb{E} \left[\left(\sum_{s \geq 1} \alpha_s I_d(\varphi_s) - \sum_{s \geq 1}^N \alpha_s I_d(\varphi_s) \right)^2 \right] = d!^2 \mathbb{E} \left[\left(\sum_{s \geq 1} \alpha_s J_d(\varphi_s) - \sum_{s \geq 1}^N \alpha_s J_d(\varphi_s) \right)^2 \right] = o(1),$$

where the last equality follows by noticing that $\{\varphi_s\}_{s \geq 1}$ are orthonormal and J_d is an isometry. This shows, recalling the linearity of stochastic integrals, that the RHS of (G.4) is the L^2 limit of $I_d(f_N)$, which completes the proof. $\square$

DEPARTMENT OF STATISTICS, UNIVERSITY OF PENNSYLVANIA, PHILADELPHIA, PA 19104, UNITED STATES
Email address: anirbanc@wharton.upenn.edu

IBM RESEARCH, YORKTOWN HEIGHTS, NY 10598, UNITED STATES
Email address: Soham.Dan@ibm.com

DEPARTMENT OF STATISTICS, UNIVERSITY OF PENNSYLVANIA, PHILADELPHIA, PA 19104, UNITED STATES
Email address: bhaswar@wharton.upenn.edu