INSTITUTE OF MATHEMATICS

THE CZECH ACADEMY OF SCIENCES

A limit theorem for small cliques in inhomogeneous random graphs

Jan Hladký
Christos Pelekis
Matas Šileikis

Preprint No. 17-2019

PRAHA 2019

A limit theorem for small cliques in inhomogeneous random graphs

Jan Hladký * Christos Pelekis[†] Matas Šileikis[‡]

March 25, 2019

Abstract

The theory of graphons comes with a natural sampling procedure, which results in an inhomogeneous variant of the Erdős–Rényi random graph, called W-random graphs. We obtain a limit theorem for the number of r-cliques in such random graphs. We show that, whereas in the case of dense Erdős–Rényi random graphs the fluctuations are normal of order n^{r-1}, the fluctuations in the setting of W-random graphs may be of order $0, n^{r-1}$, or $n^{r-0.5}$. Furthermore, when the fluctuations are of order $n^{r-0.5}$ they are normal, while when the fluctuations are of order n^{r-1} they exhibit either normal or a particular type of chi-square behavior whose parameters relate to spectral properties of W.

Keywords: graphons; inhomogeneous random graphs; limit theorems; subgraph counts; quasirandomness

1 Introduction

The purpose of this work is to investigate the distribution of the number of fixed-size cliques in an inhomogeneous variant of the Erdős–Rényi random graph $\mathbb{G}(n,p)$. The study of Erdős–Rényi random graph (see [9]) is over a half-century old and a central part in the development of the theory concerns methods for understanding the distribution of subgraph counts. These "subgraphs" may be large-scale structures, like Hamilton cycles. However, here we are concerned with counting fixed-sized subgraphs. That is, we want to describe the (bulk of the) distribution of the random variable that counts the number of copies of a fixed subgraph H as n tends to infinity or, in probabilistic language, to obtain a limit theorem for the distribution of subgraph counts of H. This problem has many variants (all copies of H, induced copies of H, joint distribution

*Institute of Mathematics of the Czech Academy of Sciences, Žitná 25, 115 67, Praha 1, Czechia. Research supported by the Alexander von Humboldt Foundation, GAČR project 18-01472Y and RVO: 67985840. E-mail: hladky@math.cas.cz

[†]Institute of Mathematics of the Czech Academy of Sciences, Žitná 25, 115 67, Praha 1, Czechia. Research supported by GAČR project 18-01472Y and RVO: 67985840. E-mail: pelekis.chr@gmail.com

[‡]Institute of Computer Science of the Czech Academy of Sciences, Pod Vodárenskou věží 2, 182 07 Prague, Czechia. Research partially supported by the Czech Science Foundation, grants GJ16-07822Y and GA19-08740S. With institutional support RVO:67985807.

for several subgraph counts, ...), and a variety of tools have been applied to obtain these results: among others, in [1] using Stein's method, in [11] using ideas from U-statistics, in [14] using the method of moments. We refer the reader to [9] for an entire chapter devoted to the topic and for further references.

1.1 Inhomogeneous random graphs and the statement of our results

At the end of this section, in Theorem 1.1, we state our main result which is a limit theorem for clique counts in the so-called W-random graphs. Before doing so, let us introduce the background necessary to this end.

A *graphon* is a symmetric Lebesgue measurable function $W : [0,1]^2 \to [0,1]$. Graphons arise as limits of sequences of large finite undirected graphs with respect to the so-called cut metric (see [10, Part 3]) and, intuitively, may be thought of as graphs on the vertex set $[0,1]$ with infinitesimally small vertices and with a $W(x,y)$-proportion of all possible edges being present in the bipartite graph whose color classes are formed by a small neighbourhood of x and of y, respectively. Graphons come with a natural sampling procedure, which results in an inhomogeneous variant of the Erdős–Rényi random graph. More precisely, given a graphon W, a random graph $\mathbb{G}(n,W)$ is a finite simple graph on n vertices, labelled by the set $[n] := \{1,\ldots,n\}$, which is generated in two steps: in the first step we draw n numbers $U_1,\ldots,U_n$ independently from the interval $[0,1]$ according to the uniform distribution and we identify their index set with the labels of the vertex set of $\mathbb{G}(n,W)$; in the second step, each pair of vertices i and j in $\mathbb{G}(n,W)$ is connected independently with probability $W(U_{i,}U_j)$. Notice that if $W(x,y)$ is constant, say, $p \in [0,1]$, then $\mathbb{G}(n,W)$ is the same as the Erdős–Rényi random graph $\mathbb{G}(n,p)$. Inhomogeneous random graphs $\mathbb{G}(n,W)$ provide substantial additional challenges compared to $\mathbb{G}(n,p)$. For example, while a standard second moment argument shows that the clique number of $\mathbb{G}(n,p)$ satisfies $\omega(\mathbb{G}(n,p)) \sim \frac{2\log n}{\log(1/p)}$, extending this formula to $\mathbb{G}(n,W)$ required new techniques, [6]. Further work on inhomogeneous random graphs so far ([5, 8]) was done in a more general, possibly sparse, model which we mention in Section 5.

Given graphs H and F, let $N(H,F)$ denote the number of copies of H in F, i.e., the number of subgraphs of F that are isomorphic to H, and consider a random variable

$$X_n(H,W) := N(H,\mathbb{G}(n,W)) .$$

If H is a fixed multigraph and W is a graphon, the *density* of H in W is defined as

$$t(H,W) := \mathbb{E}\left[\prod_{\{i,j\}\in E(H)} W(U_i,U_j)\right] . \tag{1}$$

(Notice that if the edge $\{i,j\}$ has multiplicity m in H, then the corresponding contribution to the density equals $W(U_i,U_j)^m$.) When H is a simple graph on k vertices, then the constant $t(H,W) \in [0,1]$ is the probability that a particular copy of H is present in $\mathbb{G}(n,W)$, which implies

$$\mathbb{E}\,X_n(H,W) = \frac{(n)_k}{\mathrm{aut}(H)}t(H,W),$$

where $\mathrm{aut}(H)$ is the number of automorphisms of H, and $(n)_k = n \cdot (n-1) \cdot \ldots \cdot (n-k+1)$. Corollary 10.4 in [10] implies that $X_n(H,W)$ obeys the law of large numbers, that is, for every $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left\{ X_n(H,W) = (1\pm\epsilon)\frac{(n)_k}{\mathrm{aut}(H)} \cdot t(H,W) \right\} = 1. \tag{2}$$

This is one of the key results in the theory of limits of dense graph sequences because it shows that each graphon can be approximated by finite graphs with similar subgraph densities. In this article we aim to understand the nature of fluctuations of $X_n(H,W)$ around its expectation.

Before stating our main result (Theorem 1.1) we need to introduce some definitions about spectra of graphons and more advanced concepts related to subgraph densities.

1.1.1 Spectrum of a graphon

In this section, we follow [10, Section 7.5], where details and proofs can be found. We work with the real Hilbert space $L^2[0,1]$. Suppose that $W : [0,1]^2 \to [0,1]$ is a graphon. Then we can associate with W its kernel operator $T_W : L^2[0,1] \to L^2[0,1]$ by setting

$$(T_W f)(x) = \int_0^1 W(x,y) f(y)\ dy$$

for each $f \in L^2[0,1]$. T_W is a Hilbert-Schmidt operator that has a discrete spectrum. That is, there exists a countable multiset, denoted $\mathrm{Spec}(W)$, of non-zero real *eigenvalues associated with* W. Moreover, we have that

$$\sum_{\lambda\in\mathrm{Spec}(W)} \lambda^2 = \int_{[0,1]^2} W(x,y)^2\ dx\ dy \le 1. \tag{3}$$

The *degree function* of a graphon W is the function $\deg_W : [0,1] \to [0,1]$ defined as $\deg_W(x) = \int_y W(x,y)dy$. Observe that if W is regular (that is, $\deg_W(x) = d$ for almost every $x \in [0,1]$), then $f \equiv 1$ is an eigenfunction of T_W with eigenvalue d. In this case, let $\mathrm{Spec}^-(W)$ be $\mathrm{Spec}(W)$ with the multiplicity of d decreased by 1. (It can also be shown that all eigenvalues are at most d in absolute value, but this is not necessary for our proof.)

1.1.2 Conditional densities, K_r-regular graphons and $V_W^{(r)}$

For a natural number k, we write $[k] := \{1,2,\ldots,k\}$. Given an integer $\ell \le k$, let $\binom{[k]}{\ell}$ denote all ℓ-element subsets of $[k]$. Let $J \in \binom{[k]}{\ell}$ and suppose that H is a graph on the vertex set $[k]$ for which the vertices from the set J are considered as *marked*. Given a vector $\mathbf{x} = (x_j)_{j\in J} \in [0,1]^J$, we define

$$t_{\mathbf{x}}(H,W) = \mathbb{E}\left[\prod_{\{i,j\}\in E(H)} W(U_i,U_j)\ \middle|\ U_j = x_j : j \in J \right]. \tag{4}$$

That is, $t_{\mathbf{x}}(H,W)$ is the "conditional density" of H in W, given that $U_j = x_j$, for $j \in J$. Note that, when $H = K_r$ is the r-clique, the function $\mathbf{x} \mapsto t_{\mathbf{x}}(K_r,W)$ depends only on the cardinality of J

(and not on J itself). In this case, we write $K_r^{\bullet}$ and $K_r^{\bullet\bullet}$ for K_r with one, respectively two, marked vertices and denote the corresponding conditional densities by $t_x(K_r^{\bullet}, W)$ and $t_{x,y}(K_r^{\bullet\bullet}, W)$.

A graphon W is called K_r*-free* if $t(K_r, W) = 0$ and called *complete* if W equals 1 almost everywhere.

We say that W is K_r*-regular* if for almost every $x \in [0,1]$ we have

$$t_x\left(K_r^{\bullet}, W\right) = t(K_r, W)\ .$$

In the case $r = 2$, we have $t_x\left(K_r^{\bullet}, W\right) = \deg_W(x)$, hence K_2-regularity coincides with the usual concept of regularity.

Suppose that W is a graphon and $r \geq 2$. Then we define a graphon $V_W^{(r)} : [0,1]^2 \to [0,1]$ by setting

$$V_W^{(r)}(x,y) := t_{x,y}(K_r^{\bullet\bullet}, W). \tag{5}$$

So, $V_W^{(r)}(x,y)$ is intuitively the density of K_r's containing x and y.

Suppose that we have two numbers $r \in \mathbb{N}$ and $j \in \{0,\dots,r\}$. We write $K_r \oplus_j K_r$ for the (simple) graph on $2r-j$ vertices consisting of two copies of K_r sharing j vertices. In particular $K_2 \oplus_2 K_2 = K_2$.

Let

$$\sigma_{r,W}^2 := \frac{1}{2((r-2)!)^2}\left(t(K_r \oplus_2 K_r, W) - \int_{[0,1]^2} W(x,y)\cdot t_{x,y}(K_r \oplus_2 K_r, W)\; dx\; dy\right), \tag{6}$$

where the two marked vertices in $K_r \oplus_2 K_r$ are the vertices shared by the r-cliques. We have $\sigma_{r,W}^2 \geq 0$, since

$$t(K_r \oplus_2 K_r, W) = \int_{[0,1]^2} t_{x,y}(K_r \oplus_2 K_r, W)\; dx\; dy. \tag{7}$$

If W is K_r-regular, then $V_W^{(r)}$ is regular, with $\deg_{V_W^{(r)}}(x) = t_r := t(K_r, W)$ for almost every $x \in [0,1]$. Hence, by the remark we made in Section 1.1.1, one of the eigenvalues associated with $V_W^{(r)}$ is t_r. In this case, $\mathrm{Spec}^-(V_W^{(r)})$ is $\mathrm{Spec}(V_W^{(r)})$ with the multiplicity of t_r decreased by 1.

1.1.3 Statement of the main result

We are now ready to state our main result. Here and later, $Z_n \xrightarrow{\mathrm{d}} Z$ denotes the fact that the sequence of random variables $\{Z_n\}_n$ converges in distribution to the random variable Z.

Theorem 1.1. *Let W be a graphon. Fix $r \geq 2$ and set $t_r = t(K_r, W)$. Let $X_{n,r} = X_n(K_r, W)$ be the random variable counting r-cliques in $\mathbb{G}(n, W)$. Then we have the following.*

(a) If W is K_r-free or complete then almost surely $X_{n,r} = 0$ or $X_{n,r} = \binom{n}{r}$, respectively.

(b) If W is not K_r-regular, then

$$\frac{X_{n,r} - \binom{n}{r}t_r}{n^{r-\frac{1}{2}}} \xrightarrow{\mathrm{d}} \hat{\sigma}_{r,W} \cdot Z\ , \tag{8}$$

where Z is a standard normal random variable and $\hat{\sigma}_{r,W} = \frac{1}{(r-1)!}\left(t(K_r \oplus_1 K_r, W) - t_r^2\right)^{1/2} > 0$.

(c) If W is a K_r-regular graphon which is neither K_r-free nor complete, then

$$\frac{X_{n,r} - \binom{n}{r}t_r}{n^{r-1}} \xrightarrow{\mathrm{d}} \sigma_{r,W} \cdot Z + \frac{1}{2(r-2)!} \sum_{\lambda \in \mathrm{Spec}^-(V_W^{(r)})} \lambda \cdot (Z_\lambda^2 - 1)\ , \tag{9}$$

where Z and $(Z_\lambda)_{\lambda \in \mathrm{Spec}^-(V_W^{(r)})}$ are independent standard normal and $\sigma_{r,W}$ is defined in (6). *(The series on the right-hand side of* (9) *converges a.s. and in L_1 by Lemma 2.1.)*

Part (a) is immediate. Part (b) is obtained by Stein's method. Let us mention that Part (b) has been recently reported in [7], using a framework of the so-called mod-Gaussian convergence, developed in that paper. This concept actually gives much more: firstly, the authors establish normal behaviour under conditions analogous to those in Part (b) also for other graphs than K_r. Secondly, they also prove a moderate deviation principle and a local limit theorem in this setting. So the reason we provide a proof of Part (b) is that ours is much simpler (because we are proving a weaker statement). But the main emphasis of the paper is on Part (c), which is new and deals with a regime exhibiting a more exotic behaviour.

1.1.4 When the distribution in Theorem 1.1(c) is normal or normal-free

Recall that a chi-square distribution with k degrees of freedom is the distribution of a sum of the squares of k independent standard normal random variables. Therefore the series in (9) is a weighted infinite-dimensional variant of a chi-square distribution. Note that, by (3), this random variable has finite variance. Interestingly, very similar distributions appear in [3] and [4], also in connection with graph limits. That said, the particular setting of our paper seems to be substantially different from [3, 4].

In view of (9), the only way when in the setting of Theorem 1.1(c) we get a purely normal distribution is when $\mathrm{Spec}^-(V_W^{(r)}) = \emptyset$. Recall that in this case we assume that $V_W^{(r)}$ is regular with degrees t_r. We claim that then $V_W^{(r)} = t_r$ almost everywhere. While this can be viewed as a graphon version of the Chung–Graham–Wilson Theorem on quasirandom graph sequences, here we give a short self-contained proof. Indeed, we have that

$$t_r^2 = \left(\int_{[0,1]} \deg_{V_W^{(r)}}(y)\ dy\right)^2 = \left(\int_{[0,1]^2} V_W^{(r)}(x,y)\ dx\ dy\right)^2$$

$$\boxed{\text{Jensen's inequality}} \quad \le \int_{[0,1]^2} V_W^{(r)}(x,y)^2\ dx\ dy \overset{(3)}{=} \sum_{\lambda \in \mathrm{Spec}(V_W^{(r)})} \lambda^2 = t_r^2\,.$$

In order to have an equality in Jensen's inequality, we must have $V_W^{(r)}(\cdot,\cdot) = t_r$ almost everywhere. So, the question now is which graphons W lead to a constant graphon $V_W^{(r)}$. This is a triviality for

$r = 2$. For $r \geq 3$, we put forward the following conjecture, which was first hinted in concluding remarks of [12].

Conjecture 1.2. *Suppose that $r \geq 3$ and $V_W^{(r)}$ is a constant-d graphon for some $d \in [0,1]$, that is, for almost every $(x, y) \in [0,1]^2$ we have $t_{x,y}(K_r^{\bullet\bullet}, W) = d$. Then W is K_r-free (when $d = 0$), or W is the constant-$d^{1/\binom{r}{2}}$ graphon.*

In [12], the case $r = 3$ of the aforementioned conjecture was shown to be true. Therefore, we know that if W is a graphon which is K_3-regular and not K_3-free, then the only way we can get normal limit distribution in Theorem 1.1(c) is when W is a constant graphon.

Let us now comment on a complementary question: when is the normal term absent in (9)? Looking into (6) and using (7), we see that $\sigma_{r,W} = 0$ only if $W(x, y) = 1$ for almost every $(x, y) \in [0,1]^2$ for which $t_{x,y}(K_r \oplus_2 K_r, W) > 0$. Now, observe that $t_{x,y}(K_r \oplus_2 K_r, W) > 0$ if and only if $t_{x,y}(K_r^{\bullet\bullet}, W) > 0$. That is, our condition says that if a pair (x, y) is "included in K_r's", then we must have $W(x, y) = 1$. For $r = 2$ this is equivalent to a condition that the $W \in \{0,1\}$ almost everywhere. For $r \geq 3$ we have more freedom for constructions. For example, take $r = 3$, partition $[0,1]$ into 6 sets of measure $\frac{1}{6}$ each and put one copy of the complete 3-partite graphon on the first 3 sets and another copy on the last 3 sets. Make arbitrarily wild connections between the 1st and the 4th set, and set the rest of the connections between the first 3 and the last 3 sets to 0. Such a graphon W is K_3-regular but we have $\sigma_{r,W} = 0$.

Finally, we remark that the limit in (9) is never degenerate. Indeed, suppose the contrary. The absence of the non-normal term implies $V_W^{(r)} \equiv t_r \in (0,1]$. Note that $W(x,y)t_{x,y}(K_r \oplus_2 K_r, W) = V_W^{(r)}(x,y)^2$. We plug the other assumption $\sigma_{r,W}^2 = 0$ into (6),

$$0 = t(K_r \oplus_2 K_r, W) - \int_{[0,1]^2} W(x,y) \cdot t_{x,y}(K_r \oplus_2 K_r, W)\, dx\, dy$$

$$\boxed{\text{with convention } \infty \cdot 0 = 0} \quad = \int_{[0,1]^2} \left(\frac{1}{W(x,y)} - 1 \right) \cdot V_W^{(r)}(x,y)^2\, dx\, dy = t_r^2 \cdot \int_{[0,1]^2} \left(\frac{1}{W(x,y)} - 1 \right)\, dx\, dy\ ,$$

from which we immediately see that $t_r = 0$ or $W \equiv 1$, that is, we are actually in the setting of Theorem 1.1(a).

2 Preliminaries

Asymptotic notation like $a_n = O(b_n)$ and $a_n \sim b_n$ (equivalently, $a_n = (1 + o(1))b_n$) is stated with respect to $n \to \infty$.

2.1 Hypergraphs, associated graphs, and further spectral properties

Given $r \geq 2$, an *r-uniform hypergraph $\mathcal{H}$* on vertex set V is a family of r-element subsets (called *hyperedges*) of V. In this paper we assume that $\mathcal{H}$ is a multiset (even though a term *multihypergraph* would be a more standard term). We omit the words "r-uniform", when this is clear from the

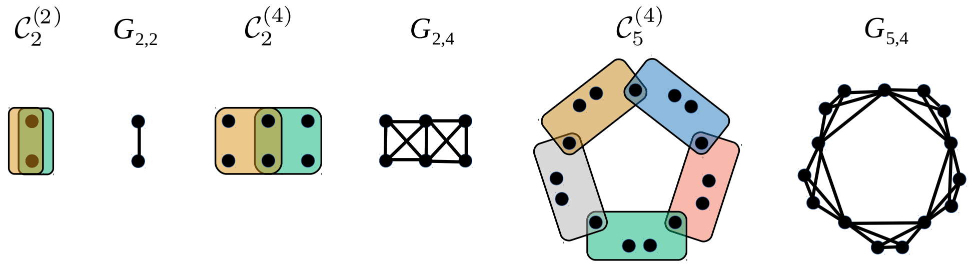

Figure 1: Examples of hypergraphs $\mathcal{C}_\ell^{(r)}$ and their associated graphs $G_{\ell,r}$.

context. By $|\mathcal{H}|$ we denote the number of hyperedges, counting multiplicities. The *degree* of a vertex $v \in V$, denoted by $\deg_{\mathcal{H}}(v)$, is the number of hyperedges (counting multiplicities) of $\mathcal{H}$ containing v. We say that $\mathcal{H}$ is *spanning* if $V = \cup_{e\in\mathcal{H}} e$. Given a hypergraph $\mathcal{H}$, the *graph associated with* $\mathcal{H}$, sometimes also called the clique graph of $\mathcal{H}$, is a graph on the same vertex set, where each hyperedge S of $\mathcal{H}$ is replaced by a clique on S, with multiple edges being replaced by single ones.

Recall that graph C_ℓ is a cycle on ℓ vertices (with C_2 being the multigraph consisting of a double edge). We use a particular hypergraph version of cycles, known as loose cycles. For $\ell \geq 2$, let $\mathcal{C}_\ell^{(r)}$ be a hypergraph with ℓ edges, each of size r, created from a graph C_ℓ by inserting into each edge additional $r-2$ vertices, all $\ell(r-2)$ new vertices being distinct (hence $\mathcal{C}_2^{(r)}$ is a pair of r-sets sharing exactly 2 vertices). Finally, let $G_{\ell,r}$ be the graph associated with $\mathcal{C}_\ell^{(r)}$. Note that $G_{2,r} = K_r \oplus_2 K_r$. See Figure 1 for examples.

We can express the densities of cycles in the graphon $V_W^{(r)}$, defined in (5), in terms of the graphon W as follows.

$$t\left(C_2, V_W^{(r)}\right) = \int_{[0,1]^2} W(x,y) \cdot t_{x,y}\left(K_r \oplus_2 K_r, W\right) dx\, dy \qquad \text{for } r \geq 2\,, \tag{10}$$

and

$$t\left(C_\ell, V_W^{(r)}\right) = t\left(G_{\ell,r}, W\right), \qquad \text{for } \ell \geq 3 \text{ and } r \geq 2\,. \tag{11}$$

Finally, we use the following relationship between the cycle densities and eigenvalues which holds for any graphon U (see [10, (7.22), (7.23)]).

$$t(C_\ell, U) = \sum_{\lambda \in \mathrm{Spec}(U)} \lambda^\ell\,. \tag{12}$$

2.2 Moment generating functions

As we noted above, the main result of [3] entails a sum of squares of normal random variables, which is very similar to the one appearing in our Theorem 1.1(c). The following lemma asserts that such distributions are well-defined and characterizes their moment generating functions.

Lemma 2.1 (see [3], Proposition 7.1[1])**.** *Let $\{\lambda_j\}_j$ be a finite or countable sequence of real numbers such that $\sum_j \lambda_j^2 < +\infty$ and let $\{Z_j\}_j$ be independent standard normal random variables. Define a (possibly infinite) sum $S = \sum_j \lambda_j (Z_j^2 - 1)$. Then S converges almost surely and in L^1. Furthermore, the moment generating function $M_S(t) := \mathbb{E}\left[\exp(tS)\right]$ is finite for $|t| < \frac{1}{8}\left(\sum_j \lambda_j^2\right)^{-1/2}$ and equals*

$$M_S(t) = \prod_j \frac{\exp(-\lambda_j t)}{\sqrt{1 - 2\lambda_j t}} \,.$$

2.3 Stein's method and the Wasserstein distance

Stein's method is one of the most powerful tools for obtaining limit theorems. Here, we follow a survey article by Ross [13]. Recall that a collection of random variables $\{Z_i\}_{i=1}^n$ is said to have a *dependency graph* $\mathcal{G}$ (on a vertex set $[n]$) if, for all $i \in [m]$, Z_i is independent of the random variables $\{Z_j\}_{j \notin N_i}$, where N_i is the neighborhood of i in $\mathcal{G}$ (including i itself). So, a dependency graph is not uniquely determined. But in many scenarios, there exists *the* dependency graph, which naturally arises by capturing the obvious dependencies, and which is also minimal among all dependency graphs.

We shall work with the Wasserstein distance between two random variables, say X and Y, which we denote $d_{\mathrm{Wass}}(X, Y)$. We do not need an exact definition — which the reader can find on page 214 of [13] — since the only time we shall employ it, we simply use an upper bound on $d_{\mathrm{Wass}}(X, Y)$ in terms of other parameters in order to prove that $d_{\mathrm{Wass}}(X_n, Z) \to 0$. The important property of the Wasserstein distance, however, is that for $Z \sim \mathcal{N}(0, 1)$ and a sequence X_n of random variables $d_{\mathrm{Wass}}(X_n, Z) \to 0$ implies that X_n converges to Z in distribution (see Section 3 of [13]).

3 Proof of Theorem 1.1(b)

We establish asymptotic normality in the setting of Theorem 1.1(b) by using a Stein-method-based off-the-shelf bound for the Wasserstein distance.

To simplify notation, set $d_j = t(K_r \oplus_j K_r, W)$, for $j \in [r]$. By Jensen's inequality we have that

$$t_r^2 = \left(\int_0^1 t_x(K_r^\bullet, W)\, dx\right)^2 < \int_0^1 t_x(K_r^\bullet, W)^2\, dx = d_1, \tag{13}$$

where the strict inequality follows from the fact that W is not K_r-regular.

Given a set $R \in \binom{[n]}{r}$, let I_R be the indicator of the event that R induces a clique in $\mathbb{G}(n, W)$ and note that $\mathbb{E}\, I_R = t_r$. Let $\mathcal{G}$ be the natural dependency graph of the random variables $I_R, R \in \binom{[n]}{r}$ (with

[1]Note that there is an error in the arXiv version of Proposition 7.1 in [3].

edges corresponding to pairs R_1R_2 such that $R_1 \cap R_2 \neq \emptyset$). Notice that in $\mathcal{G}$ every neighbourhood N_R has the same size D, defined as

$$D = \sum_{\ell=1}^{r} \binom{r}{\ell}\binom{n-r}{r-\ell} = O(n^{r-1}). \tag{14}$$

Writing $Y_R := I_R - t_r$, we have$X_{n,r} - \binom{n}{r}t_r = \sum_{R\in\binom{[n]}{r}} Y_i$. For disjoint R_1, R_2, variables Y_{R_1}, Y_{R_2} are independent, while $|R_1 \cap R_2| = \ell \geq 1$ implies $\mathbb{E}(Y_{R_1}Y_{R_2}) = d_\ell - t_r^2$, and (13) implies $\hat{\sigma}_{r,W}^2 = (d_1 - t_r^2)/((r-1)!)^2 > 0$. Hence

$$\begin{aligned}\sigma_n^2 := \mathrm{Var}\left(\sum_R Y_R\right) &= \sum_{\ell=1}^{r} \binom{n}{\ell}\binom{n-\ell}{r-\ell}\binom{n-r}{r-\ell}\left(d_\ell - t_r^2\right) \\ &= \binom{n}{1}\binom{n-1}{r-1}\binom{n-r}{r-1}\left(d_1 - t_r^2\right) + \sum_{\ell=2}^{r} O(n^{2r-\ell}) \\ &\sim \hat{\sigma}_{r,W}^2 n^{2r-1}\ . \end{aligned} \tag{15}$$

Theorem 3.6 in [13] tells us that for $Q := \frac{1}{\sigma_n}\sum_R Y_R$ and $Z \sim \mathcal{N}(0,1)$ we have

$$d_{\mathrm{Wass}}(Q,Z) \leq \frac{D^2}{\sigma_n^3} \cdot \sum_R \mathbb{E}\left[|Y_R|^3\right] + \frac{\sqrt{28}D^{3/2}}{\sqrt{\pi}\cdot\sigma_n^2} \cdot \sqrt{\sum_R \mathbb{E}\left[Y_R^4\right]}.$$

Crudely bounding each of the sums on the right-hand side by $\binom{n}{r} \leq n^r$, and using (14) and (15), we obtain that $d_{\mathrm{Wass}}(Q,Z) = O(n^{-1/2})$. By the remark we made in Section 2.3, we conclude that $Q \xrightarrow{\mathrm{d}} Z$. In view of (15) and Slutsky's theorem,

$$\frac{X_{n,r} - \binom{n}{r}t_r}{n^{r-1/2}} = \frac{\sigma_n}{n^{r-1/2}} \cdot Q \xrightarrow{\mathrm{d}} \hat{\sigma}_{r,W} Z,$$

which completes the proof.

4 Proof of Theorem 1.1(c)

We employ the method of moments, in the way it is described in Section 6.1 of [9]. For this it is enough to show that the central moments of $X_{n,r}/n^r$ converge to the corresponding moments of the random variable

$$Y = \sigma_{r,W} \cdot Z + \frac{1}{2(r-2)!} \sum_{\lambda\in\mathrm{Spec}^-(V_W^{(r)})} \lambda \cdot (Z_\lambda^2 - 1) \tag{16}$$

from (9), and to verify that the moment generating function $M_Y(t) = \mathbb{E}\, e^{tY}$ is finite in some neighbourhood of zero (so that the distribution of Y is determined by its moments).

Recall that for a standard normal random variable Z we have $M_Z(x) = \exp\left(x^2/2\right)$ and hence $M_{\sigma_{r,W}\cdot Z}(x) = \exp\left(\frac{\sigma_{r,W}^2 x^2}{2}\right)$. On the other hand, Lemma 2.1 tells us that the moment generating function of the second summand in (16) is $\prod_{\lambda\in\mathrm{Spec}^-(V_W^{(r)})} \exp\left(-\frac{\lambda x}{2(r-2)!}\right) / \sqrt{1-\frac{\lambda x}{(r-2)!}}$. Since the moment generating function of a sum of independent random variables equals the product of the moment generating functions of individual generating functions, it follows that

$$M_Y(x) = \exp\left(\frac{\sigma_{r,W}^2 x^2}{2}\right) \prod_{\lambda\in\mathrm{Spec}^-(V_W^{(r)})} \frac{\exp\left(-\frac{\lambda x}{2(r-2)!}\right)}{\sqrt{1-\frac{\lambda x}{(r-2)!}}}\,. \tag{17}$$

As in Section 3, let I_R be the indicator of the event that the set of vertices R induces a clique in $\mathbb{G}(n, W)$. Given an m-tuple $(R_1, \ldots, R_m)$ of (not necessary distinct) elements of $\binom{[n]}{r}$, let

$$\Delta(R_1, \ldots, R_m) := \mathbb{E}\left[\prod_{i=1}^{m}(I_{R_i} - t_r)\right]\,. \tag{18}$$

Then

$$\mathbb{E}\left[\left(X_{n,r} - \binom{n}{r}t_r\right)^m\right] = \sum_{(R_1,\ldots,R_m)\in\binom{[n]}{r}^m} \Delta(R_1, \ldots, R_m)\,. \tag{19}$$

Let $\mathfrak{X}(n, r, m) \subset \binom{[n]}{r}^m$ be those m-tuples $(R_1, \ldots, R_m)$ for which we have $|R_i \cap \cup_{j\neq i} R_j| \leq 1$ for some $i \in [m]$. Suppose now that $(R_1, \ldots, R_m) \in \mathfrak{X}(n, r, m)$. Without loss of generality, suppose that $|R_m \cap \cup_{j=1}^{m-1} R_j| \leq 1$. Assume first that $|R_m \cap \cup_{j=1}^{m-1} R_j| = 1$, say $\{v\} = R_m \cap \cup_{j=1}^{m-1} R_j$. If we condition on $U_v = x$, the indicator I_{R_m} becomes independent of $\{I_{R_i} : i \in [m-1]\}$, and hence

$$\Delta(R_1, \ldots, R_m) = \int_0^1 \mathbb{E}\left[I_{R_m} - t_r \,\middle|\, U_v = x\right] \cdot \mathbb{E}\left[\prod_{i=1}^{m-1}(I_{R_i} - t_r) \,\middle|\, U_v = x\right]\, dx. \tag{20}$$

Since W is K_r-regular, we have $\mathbb{E}\left[I_{R_m} - t_r \,\middle|\, U_v = x\right] = t_x(K_r^\bullet, W) = t_r$ for almost every x. Therefore $\Delta(R_1, \ldots, R_m) = 0$. An even simpler calculation yields the same conclusion when $|R_m \cap \cup_{j=1}^{m-1} R_j| = 0$. Hence, we can rewrite (19) as

$$\mathbb{E}\left[\left(X_{n,r} - t_r\binom{n}{r}\right)^m\right] = \sum_{(R_1,\ldots,R_m)\in\binom{[n]}{r}^m\setminus\mathfrak{X}(n,r,m)} \Delta(R_1, \ldots, R_m)\,. \tag{21}$$

Every m-tuple $(R_1, \ldots, R_m)$ can be identified with a spanning hypergraph henceforth denoted $\mathcal{H}(R_1, \ldots, R_m)$, with vertex set $\cup_i R_i$ and hyperedge multiset $\{R_1, \ldots, R_m\}$.

Claim 4.1. *Suppose that $(R_1, \ldots, R_m) \in \binom{[n]}{r}^m \setminus \mathfrak{X}(n, r, m)$. The number v of vertices in the hypergraph $\mathcal{H} = \mathcal{H}(R_1, \ldots, R_m)$ satisfies $v \leq (r-1)m$. The equality is attained if and only if each hyperedge in $\mathcal{H}$ contains exactly 2 vertices of degree 2 and all other vertices have degree 1.*

Proof of Claim 4.1. Let k be the number of vertices in $\mathcal{H}$ of degree 1. Since $(R_1,\ldots,R_m) \notin \mathfrak{X}(n,r,m)$ we have that

$$k \le (r-2)m \ . \tag{22}$$

Since $\mathcal{H}$ is spanning, $v-k$ vertices have degree at least 2, and it follows

$$k + 2(v-k) \le \sum_{v\in V(\mathcal{H})} \deg(v) = rm \ . \tag{23}$$

Therefore

$$v \overset{(23)}{\le} \frac{rm-k}{2} + k = \frac{rm+k}{2} \overset{(22)}{\le} \frac{rm+(r-2)m}{2} = (r-1)m,$$

as required. The second statement of the claim is immediate. □

Let $\mathfrak{F}(n,r,m)$ be those $(R_1,\ldots,R_m) \in \binom{[n]}{r}^m \setminus \mathfrak{X}(n,r,m)$ for which the corresponding hypergraph $\mathcal{H}(R_1,\ldots,R_m)$ has $(r-1)m$ vertices. Since for each $(R_1,\ldots,R_m) \in \binom{[n]}{r}^m \setminus (\mathfrak{X}(n,r,m) \cup \mathfrak{F}(n,r,m))$ we have $|\cup_i R_i| \le (r-1)m-1$, we can record each element of $\binom{[n]}{r}^m \setminus (\mathfrak{X}(n,r,m) \cup \mathfrak{F}(n,r,m))$ by an $((r-1)m-1)$-tuple of $[n]$, and then by specifying to which of the sets $R_1,\ldots,R_m$ each element of that $((r-1)m-1)$-tuple is an element of. Thus,

$$\left| \binom{[n]}{r}^m \setminus (\mathfrak{X}(n,r,m) \cup \mathfrak{F}(n,r,m)) \right| \le \binom{n}{(r-1)m-1} \cdot (2^m)^{(r-1)m-1} = O(n^{(r-1)m-1}) \ . \tag{24}$$

Since $|\Delta(R_1,\ldots,R_m)| \le 1$, from (21) and (24) we infer

$$\mathbb{E}\left[\left(X_{n,r} - t_r\binom{n}{r}\right)^m\right] = \sum_{(R_1,\ldots,R_m)\in\mathfrak{F}(n,r,m)} \Delta(R_1,\ldots,R_m) + O(n^{(r-1)m-1}) \ . \tag{25}$$

Now, fix $(R_1,\ldots,R_m) \in \mathfrak{F}(n,r,m)$ and consider the hypergraph $\mathcal{H} = \mathcal{H}(R_1,\ldots,R_m)$. Notice that when $r=2$ then some edges in $\mathcal{H}$ may have multiplicities, but this is not the case when $r \ge 3$. Now, replace every r-edge, say R, of $\mathcal{H}$ by 2-edge that consists of the vertices of R having degree 2 and notice that this results in a 2-regular multigraph, that is, a union of vertex-disjoint cycles and double edges. In particular, this implies that the hypergraph $\mathcal{H}$ is a union of vertex-disjoint loose cycles.

The next claim deals with proper subhypergraphs of a loose cycle $\mathcal{C}_i^{(r)}$.

Claim 4.2. *For each $i, r \ge 2$, for any proper subhypergraph $\mathcal{C} \subset \mathcal{C}_i^{(r)}$, we have $\mathbb{E}\left[\prod_{Q\in\mathcal{C}} I_Q\right] = t_r^{|\mathcal{C}|}$.*

Proof of Claim 4.2. We proceed by induction on the number of hyperedges of $\mathcal{C}$. The case when $\mathcal{C} = \emptyset$ is trivial. So suppose that $\mathcal{C} \ne \emptyset$.

Since $\mathcal{C}$ is a proper subhypergraph of $\mathcal{C}_i^{(r)}$, it contains a hyperedge $S \in \mathcal{C}$ such that for $\mathcal{C}^- := \mathcal{C} \setminus \{S\}$ we have $|S \cap \bigcup \mathcal{C}^-| \le 1$ (here and below $\bigcup \mathcal{H}$ stands for the union of hyperedges of $\mathcal{H}$). Let us deal first with the case $|S \cap \bigcup \mathcal{C}^-| = 1$, and let v be the vertex shared by S and $\bigcup \mathcal{C}^-$. By the same argument as in (20), we have

$$\mathbb{E}\left[\prod_{Q\in\mathcal{C}} I_Q\right] = \int_0^1 \mathbb{E}\left[I_S \mid U_v = x\right] \cdot \mathbb{E}\left[\prod_{Q\in\mathcal{C}^-} I_Q \mid U_v = x\right] \, dx,$$

By the K_r-regularity, we have $\mathbb{E}[I_S \mid U_v = \cdot] = t_r$ almost everywhere. Thus, using the induction hypothesis on $\mathcal{C}^-$, we conclude that

$$\mathbb{E}\left[\prod_{Q\in\mathcal{C}} I_Q\right] = \int_0^1 t_r \, \mathbb{E}\left[\prod_{Q\in\mathcal{C}^-} I_Q \mid U_v = x\right] \, dx = t_r \, \mathbb{E}\left[\prod_{Q\in\mathcal{C}^-} I_Q\right] = t_r \cdot t_r^{|\mathcal{C}^-|} \, ,$$

as was needed. The case $|S \cap \bigcup \mathcal{C}^-| = 0$ is even simpler:

$$\mathbb{E}\left[\prod_{Q\in\mathcal{C}} I_Q\right] = \mathbb{E}\left[I_S\right] \cdot \mathbb{E}\left[\prod_{Q\in\mathcal{C}^-} I_Q\right] = t_r \cdot t_r^{|\mathcal{C}^-|} \, .$$

□

Let us consider the following set of $(m-1)$-dimensional vectors,

$$\mathcal{V}_m := \left\{ \mathbf{k} = (k_2, \ldots, k_m) \in \mathbb{N}_0^{m-1} : \sum_{i=2}^{m} i k_i = m \right\}.$$

Suppose that $\mathbf{k} \in \mathcal{V}_m$. Let $\mathcal{H}_{\mathbf{k}}^{(r)}$ denote the hypergraph formed by k_i copies of $\mathcal{C}_i^{(r)}$ for each $i = 2, \ldots, m$. These are precisely the hypergraphs which can be obtained from m-tuples $(R_1, \ldots, R_m) \in \mathfrak{F}(n, r, m)$.

Claim 4.3. *Suppose that $(R_1, \ldots, R_m) \in \mathfrak{F}(n, r, m)$ is an m-tuple for which $\mathcal{H}(R_1, \ldots, R_m)$ is isomorphic to $\mathcal{H}_{\mathbf{k}}^{(r)}$, for some $\mathbf{k} \in \mathcal{V}_m$. Then*

$$\Delta(R_1, \ldots, R_m) = \prod_{\ell=2}^{m} \left(t\left(G_{\ell,r}, W\right) - t_r^{\ell} \right)^{k_\ell} , \tag{26}$$

where $G_{\ell,r}$ is the multigraph associated to $\mathcal{C}_\ell^{(r)}$, as defined in Section 2.1.

Proof of Claim 4.3. Suppose first that we are given an ℓ-tuple $(Q_1, \ldots, Q_\ell) \in \mathfrak{F}(n, r, \ell)$ for which $\mathcal{H}(Q_1, \ldots, Q_\ell)$ is isomorphic to $\mathcal{C}_\ell^{(r)}$. We have

$$\begin{aligned}
\Delta(Q_1, \ldots, Q_\ell) &= \sum_{A \subseteq [\ell]} (-t_r)^{\ell - |A|} \cdot \mathbb{E}\left[\prod_{j\in A} I_{Q_j}\right] \\
\boxed{\text{Claim 4.2}} \quad &= \mathbb{E}\left[\prod_{j\in[\ell]} I_{Q_j}\right] + \sum_{A \subsetneq [\ell]} (-t_r)^{\ell-|A|} \, t_r^{|A|} \\
&= \mathbb{E}\left[\prod_{j\in[\ell]} I_{Q_j}\right] + (t_r - t_r)^{\ell} - t_r^{\ell} \\
&= t\left(G_{\ell,r}, W\right) - t_r^{\ell} \, .
\end{aligned}$$

To conclude the proof, observe that $\Delta(R_1,\dots,R_m)$ can be written as a product of $\Delta(R_{s_1},\dots,R_{s_\ell})$ over tuples $(R_{s_1},\dots,R_{s_\ell})$ that form copies of some $\mathcal{C}_\ell^{(r)}$. □

Claim 4.4. *Fix $\mathbf{k}\in\mathcal{V}_m$. Then the number of m-tuples $(R_1,\dots,R_m)\in\binom{[n]}{r}^m$ for which $\mathcal{H}(R_1,\dots,R_m)$ is isomorphic to $\mathcal{H}_{\mathbf{k}}^{(r)}$ is equal to*

$$A(n,r,\mathbf{k}) := \frac{m!\cdot(n)_{(r-1)m}}{\prod_{i=2}^m (2i((r-2)!)^i)^{k_i}\ \cdot\ k_i!}. \tag{27}$$

Proof of Claim 4.4. Suppose first that $r\geq 3$. Notice that the number of automorphisms of $\mathcal{C}_i^{(r)}$ equals $2i\cdot((r-2)!)^i$, and therefore the number of automorphisms of $\mathcal{H}_{\mathbf{k}}^{(r)}$ satisfies

$$\mathrm{aut}(\mathcal{H}_{\mathbf{k}}^{(r)}) = \prod_{i=2}^m (2i((r-2)!)^i)^{k_i}\ \cdot\ k_i!\ .$$

As there are $\frac{(n)_{(r-1)m}}{\mathrm{aut}(\mathcal{H}_{\mathbf{k}}^{(r)})}$ copies of $\mathcal{H}_{\mathbf{k}}^{(r)}$ on n vertices and each copy corresponds to $m!$ many m-tuples $(R_1,\dots,R_m)$, the proof of the case $r\geq 3$ is complete.

The case $r=2$ is similar; the only difference being that the number of automorphisms of $\mathcal{C}_2^{(2)}$ equals 2 and that each copy of $\mathcal{H}_{\mathbf{k}}^{(2)}$ corresponds to $\frac{m!}{2^{k_2}}$ many m-tuples. The details are left to the reader. □

We now resume to expressing $\mathbb{E}\left[\left(X_{n,r}-t_r\binom{n}{r}\right)^m\right]$, which we abandoned at (25). Recall that $G_{i,r}$ is the graph associated with $\mathcal{C}_i^{(r)}$. Adding (26) and (27), we get

$$\begin{aligned}\mathbb{E}\left[\left(X_{n,r}-t_r\binom{n}{r}\right)^m\right] &= \sum_{\mathbf{k}\in\mathcal{V}_m} A(n,r,\mathbf{k})\prod_{\ell=2}^m (t(G_{\ell,r},W)-t_r^\ell)^{k_\ell} + O(n^{(r-1)m-1}) \\ &= n^{(r-1)m}m!\sum_{\mathbf{k}\in\mathcal{V}_m}\prod_{\ell=2}^m\left(\frac{t(G_{\ell,r},W)-t_r^\ell}{2\ell((r-2)!)^\ell}\right)^{k_\ell}\frac{1}{k_\ell!}+O(n^{(r-1)m-1})\,.\end{aligned} \tag{28}$$

For $\ell=2,3,\dots$, let us write

$$d_\ell := \frac{t(G_{\ell,r},W)-t_r^\ell}{2\ell((r-2)!)^\ell}\ . \tag{29}$$

Let us consider a formal power series

$$f(x) := \prod_{\ell=2}^\infty \exp(d_\ell x^\ell) = \exp\left(\sum_{\ell=2}^\infty d_\ell x^\ell\right)\ . \tag{30}$$

Claim 4.5. *For each $m\in\mathbb{N}$, as $n\to\infty$, the quantity $\frac{1}{m!}\cdot\mathbb{E}[\left(X_{n,r}-t_r\binom{n}{r}\right)^m/(n^{(r-1)m})]$ converges to the coefficient of x^m of $f(x)$.*

Proof. We have

$$
\begin{aligned}
[\![x^m]\!] f(x) = [\![x^m]\!]\left(\exp\left(\sum_{\ell=2}^{\infty} d_\ell x^\ell\right)\right) &= [\![x^m]\!]\left(\sum_{j=1}^{\infty}\frac{1}{j!}\cdot\left(\sum_{\ell=2}^{\infty} d_\ell x^\ell\right)^j\right)\\
\boxed{\text{multinomial theorem}} \qquad &= \sum_{j=1}^{m}\frac{1}{j!}\sum_{\mathbf{k}\in\mathcal{V}_m : k_2+\ldots+k_m=j}\binom{j}{k_2,\cdots,k_m}\cdot\prod_{\ell=1}^{m} d_\ell^{k_\ell}\\
\boxed{\text{definition of the multinomial coefficient}} \qquad &= \sum_{j=1}^{m}\sum_{\mathbf{k}\in\mathcal{V}_m : k_2+\ldots+k_m=j}\prod_{\ell=1}^{m}\frac{d_\ell^{k_\ell}}{k_\ell!}\\
&= \sum_{\mathbf{k}\in\mathcal{V}_m}\prod_{\ell=2}^{m} d_\ell^{k_\ell}\frac{1}{k_\ell!}\\
\boxed{\text{by (28)}} \qquad &= \lim_{n\to\infty}\frac{1}{m!}\cdot\frac{\mathbb{E}\left[\left(X_{n,r}-t_r\binom{n}{r}\right)^m\right]}{n^{(r-1)m}}.
\end{aligned}
$$

$\square$

Since $|d_i| \leq ((r-2)!)^{-i}$, the series $\sum_i d_i x^i$ has positive radius of convergence and $f(x)$ can be expanded as power series around zero. In the next claim, we show that the function f equals the moment-generating function M_Y defined in (17).

Claim 4.6. *In some neighbourhood of zero we have $M_Y(x) = f(x)$.*

Proof of Claim 4.6. Using (12) and (10) we infer that

$$
\begin{aligned}
d_2 &= \frac{t(K_r \oplus_2 K_r, W) - t_r^2}{4((r-2)!)^2}\\
&= \frac{t(K_r \oplus_2 K_r, W) - t(C_2, V_W^{(r)}) + t(C_2, V_W^{(r)}) - t_r^2}{4((r-2)!)^2}\\
&= \frac{t(K_r \oplus_2 K_r, W) - t(C_2, V_W^{(r)})}{4((r-2)!)^2} + \frac{1}{4((r-2)!)^2}\sum_{\lambda\in\mathrm{Spec}^-(V_W^{(r)})}\lambda^2\\
&= \frac{\sigma_{r,W}^2}{2} + \frac{1}{4}\sum_{\lambda\in\mathrm{Spec}^-(V_W^{(r)})}\left(\frac{\lambda}{(r-2)!}\right)^2
\end{aligned}
$$

and for $\ell \geq 3$, using (12) and (11) that

$$
\begin{aligned}
d_\ell &= \frac{t(G_{\ell,r}, W) - t_r^\ell}{2\ell((r-2)!)^\ell} =\\
&= \frac{t(C_\ell, V_W^{(r)}) - t_r^\ell}{2\ell((r-2)!)^\ell} = \frac{1}{2\ell}\sum_{\lambda\in\mathrm{Spec}^-(V_W^{(r)})}\left(\frac{\lambda}{(r-2)!}\right)^\ell.
\end{aligned}
$$

Plugging this into $\sum_{\ell=2}^{\infty} d_\ell x^\ell$ and interchanging the order of summation (using Fubini's theorem and assuming $|x|$ is small enough), we obtain

$$\begin{aligned}\log f(x) = \sum_{\ell=2}^{\infty} d_\ell x^\ell &= \frac{\sigma_{r,W}^2 x^2}{2} + \sum_{\lambda \in \mathrm{Spec}^-(V_W^{(r)})} \sum_{\ell=2}^{\infty} \frac{1}{2\ell} \left(\frac{\lambda x}{(r-2)!} \right)^\ell \\ &= \frac{\sigma_{r,W}^2 x^2}{2} - \frac{1}{2} \sum_{\lambda \in \mathrm{Spec}^-(V_W^{(r)})} \left(\ln\left(1 - \frac{\lambda x}{(r-2)!}\right) + \frac{\lambda x}{(r-2)!} \right) .\end{aligned}$$

By exponentiating the above expression we easily obtain (17), thus completing the proof. □

We are now finished with the proof of Theorem 1.1(c). Indeed, Claims 4.5 and 4.6 imply that the mth moment of $(X_{n,r} - \binom{n}{r} t_r)/n^{r-1}$ converges to $m! [\![x^m]\!] f(x) = m! [\![x^m]\!] M_Y(x) = \mathbb{E}\, Y^m$ for every m. Hence the method of moments (see Theorem 6.1 and the preceding comments in [9]) implies that $\frac{X_{n,r} - \binom{n}{r} t_r}{n^{r-1}} \xrightarrow{\mathrm{d}} Y$.

5 Concluding remarks

In this paper, we initiated the study of limit theorems for subgraph counts in $\mathbb{G}(n, W)$. However, the results in this paper should be considered just first steps, and the area offers several obvious open problems.

- Extend Theorem 1.1 to other graphs than K_r. Recall the counterpart to Part (b) was worked out in [7], so the only challenge left is Part (c). The calculations seem to be substantially more involved in this case.
- Extend Theorem 1.1 to sparser regimes. Recall that the Central limit theorem for the count of K_r in $\mathbb{G}(n, p)$ holds for $p = p(n)$ as small as $p(n) \gg n^{-2/r-1}$, that is, as long as the expected number of K_r's tends to infinity.
- To model a sparse inhomogeneous random graph, fix a scaling factor $p = p(n) \to 0$. Then $\mathbb{G}(n, p \cdot W)$ is a sparse inhomogeneous random graph model. Note that then the assumption that W is bounded from above by 1 can be relaxed somewhat. For example, the giant component of $\mathbb{G}(n, W/n)$ is studied in [5]. So, we suggest to obtain limit theorems for the count of K_r (or other graphs) in $\mathbb{G}(n, p \cdot W)$.
- To strengthen the limit theorem obtained here to a local limit theorem. Even in the case of $\mathbb{G}(n, p)$ this is a very difficult problem which was resolved only recently, [2]. Note that such a local limit theorem would have additional restrictions. For example, if W is a graphon consisting of two constant-1 components of measure $\frac{1}{2}$ each, then $X_{n,2}$ is almost surely of the form $\binom{k}{2} + n - \binom{k}{2}$, $k \in \mathbb{N}$, that is, not all integer values can be achieved.

References

[1] A.D. Barbour, M. Karoński, A. Ruciński, *A central limit theorem for decomposable random variables with applications to random graphs*, Journal of Combinatorial Theory, Series B **47** (1989) 125–145.

[2] R. Berkowitz, *A local limit theorem for cliques in* $G(n,p)$, (preprint) arXiv:1811.03527.

[3] B.B. Bhattacharya, P. Diaconis, S. Mukherjee, *Universal limit theorems in graph coloring problems with connections to extremal combinatorics*, The Annals of Applied Probability **27** (1) (2017) 337–394.

[4] B.B. Bhattacharya, S. Mukherjee, *Monochromatic subgraphs in randomly colored graphons*, (preprint) arXiv:1707.05889.

[5] B. Bollobás, S. Janson, O. Riordan, *The phase transition in inhomogeneous random graphs*, Random Structures & Algorithms, **31** (1) (2007) 3–122.

[6] M. Doležal, J. Hladký, A. Máthé, *Cliques in dense inhomogeneous random graphs*, Random Structures & Algorithms, **51** (2) (2017) 275–314.

[7] V. Féray, P.-L. Méliot, A. Nikeghbali, *Graphons, permutons and the Thoma simplex: three mod-Gaussian moduli spaces*, (preprint) arXiv:1712.06841.

[8] N. Fraiman, D. Mitsche, *The diameter of inhomogeneous random graphs*, Random Structures & Algorithms, **53** (2) (2018) 308–326.

[9] S. Janson, T. Łuczak, A. Ruciński, *Random Graphs*, John Wiley and Sons, 2011.

[10] L. Lovász, *Large networks and graph limits*, Vol. 60 of American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, 2012.

[11] K. Nowicki, J.C. Wierman, *Subgraph Counts in Random Graphs Using Incomplete U-statistics methods*, Annals of Discrete Mathematics, **38** (1988) 299–310.

[12] C. Reiher, M. Schacht, *Forcing quasirandomness with triangles*, (preprint), arXiv:1711.04754.

[13] N. Ross, *Fundamentals of Stein's method*, Probability Surveys **8** (2011), 210–293.

[14] A. Ruciński, *When Are Small Subgraphs of a Random Graph Normally Distributed?*, Probability Theory and Related Fields **78** (1988) 1–10.