

A limit theorem for small cliques in inhomogeneous random graphs

Jan Hladký^{*} Christos Pelekis[†] Matas Šileikis[‡]

January 22, 2021

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 prove, via the method of moments, 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 .

These results can also be deduced from a general setting [Janson and Nowicki, PTRF 1991], based on the projection method. In addition to providing alternative proofs, our approach makes direct links to the theory of graphons.

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

1 Introduction

1.1 Subgraph counts in random graphs

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 the Erdős–Rényi random graph (see [11]) is over a half-century old. A central part in the development of the random graph theory concerns methods for understanding the distribution of subgraph counts. Even though “subgraphs” may be large-scale structures, like Hamilton cycles, here we are concerned

^{*}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 GJ18-01472Y 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 GJ18-01472Y 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, GA19-08740S and GJ20-27757Y. With institutional support RVO:67985807. E-mail: matas.sileikis@gmail.com

with counting fixed-sized subgraphs. In particular, 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 for several subgraph counts, ...), and a variety of tools have been applied to tackle it, including Stein's method in [2], ideas from U-statistics in [16], and the method of moments in [19]. We refer the reader to [11] for an entire chapter devoted to the topic and for further references.

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

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

where $\mathbb{G}(n)$ is a model of random graphs on n vertices of interest. The asymptotic normality of X_n in the Erdős–Rényi random graph model $\mathbb{G}(n) = \mathbb{G}(n, p(n))$, has been fully described by Ruciński [19] already in 1988. When $p(n)$ is constant it first appeared as a special case of a result by Nowicki and Wierman [16] who showed that for any fixed graph H with at least one edge and any $p \in (0, 1)$,

$$\frac{X_n - \mathbb{E}[X_n]}{\sqrt{\text{Var}[X_n]}} \xrightarrow{d} Z \sim \mathcal{N}(0, 1).$$

(Here and below $Z_n \xrightarrow{d} Z$ denotes convergence in distribution.) Our result extends this result in the case of cliques $H = K_r$, $r \geq 2$ and fixed p by generalizing $\mathbb{G}(n, p)$ to the random graph model $\mathbb{G}(n, W)$ which plays a key role in the theory of dense graph limits introduced in [7, 15]. This model, which we formally define in Section 1.2, can be described as follows: each vertex is assigned a random *type* and then each pair of vertices is independently included as an edge in $\mathbb{G}(n, W)$ with a probability that depends only on the types of the two vertices. The random graph $\mathbb{G}(n, p)$ corresponds to a single type and hence the same probability p for every pair of vertices.

Henceforth, we shall write $X_n(H, W) := N(H, \mathbb{G}(n, W))$. In Theorem 1.2, we state our main result, a limit theorem for $X_n(H, W)$ for each $H = K_r$, $r \geq 2$. However, since our main result Theorem 1.2 requires quite a few additional definitions, for a preview we state its implicit version, Theorem 1.1. We start the next subsection by the minimum number of definitions.

1.2 Inhomogeneous random graphs and the simplified result statement

A *graphon* is a symmetric Lebesgue measurable function $W : [0, 1]^2 \rightarrow [0, 1]$. Graphons arise as limits of sequences of large finite undirected graphs with respect to the so-called cut metric (see [14, Part 3]). Intuitively, graphons 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 , the random graph $\mathbb{G}(n, W)$ is a finite simple graph on n vertices, labelled by the set $[n] := \{1, \dots, n\}$, which is generated in two

steps: in the first step we draw n numbers *types* U_1, \dots, 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, [8]. Further work on inhomogeneous random graphs so far ([6, 10]) was done in a more general, possibly sparse, model which we mention in Section 5.

Corollary 10.4 in [14] implies that $X_n(H, W)$ obeys the law of large numbers, that is, for every $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P} \{X_n(H, W) = (1 \pm \epsilon) \mathbb{E}[X_n(H, W)]\} = 1. \quad (1)$$

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.

Fix $r \geq 2$ and write $X_n = X_n(K_r, W)$ for the number of r -cliques in $\mathbb{G}(n, W)$. Since every r -set of vertices in $\mathbb{G}(n, W)$ induces a random graph distributed as $\mathbb{G}(r, W)$, every r -set of vertices induces a clique with probability

$$t(K_r, W) := \mathbb{P} \{\mathbb{G}(r, W) = K_r\}.$$

If we prescribe type x to one of the vertices in an r -set, then this r -set induces a clique with probability

$$t(x) = t_x(K_r, W) := \mathbb{P} \{\mathbb{G}(r, W) = K_r \mid U_1 = x\}.$$

Clearly, $t(K_r, W) = \int_0^1 t(x) dx$.

For a measurable function $f : [0, 1]^k \rightarrow \mathbb{R}$ we say f is *constant* and denote $f \equiv c$ whenever $f(x) = c$ for almost every $x \in [0, 1]^k$.

We start with a simplified version of our main result. The additional information provided by the complete version, that is, Theorem 1.2, is the description of constants σ and c_0, c_1, \dots in terms of the graphon W .

Theorem 1.1. *For $t(K_r, W)$, $t(x)$, and X_n defined above, the following holds.*

- (a) *If $t(K_r, W) = 0$ or $t(K_r, W) = 1$, then almost surely $X_n = 0$ or $X_n = \binom{n}{r}$, respectively.*
- (b) *If $t(x)$ is not constant, then there is a constant $\sigma > 0$ such that*

$$\frac{X_n - \mathbb{E}[X_n]}{n^{r-1/2}} \xrightarrow{d} \sigma Z, \quad (2)$$

where Z is a standard normal random variable.

- (c) *If $t(x)$ is constant (but other than 0 or 1), then there are real numbers c_0, c_1, \dots , such that $\sum_i c_i^2 \in (0, \infty)$ and*

$$\frac{X_n - \mathbb{E}[X_n]}{n^{r-1}} \xrightarrow{d} c_0 Z_0 + \sum_{i \geq 1} c_i (Z_i^2 - 1), \quad (3)$$

where Z_0, Z_1, \dots are independent standard normal random variables. The series on the right-hand side of (3) converges a.s. and in L^2 .

Proof. Theorem 1.2 implies all the claims of Theorem 1.1 except for the fact that $\sum_i c_i^2 > 0$, which follows from Remark 1.6. \square

We note that non-normal limit theorem occurs even in $\mathbb{G}(n, p)$ with p fixed, but for *induced* subgraph counts, see [11, Theorem 6.52].

A toy instance of Theorem 1.1(c). The proof of Theorem 1.1(c) proceeds by the method of moments and thus in itself does not provide much intuition for the asserted non-normal limit fluctuations. This behavior is however suggested by the following simple example. Let $r \geq 2$ be arbitrary. Consider W which is a ‘disjoint union of two equally-sized cliques’, that is $W(x, y) = 1$ if $(x, y) \in [0, \frac{1}{2}]^2 \cup [\frac{1}{2}, 1]^2$ and $W(x, y) = 0$ otherwise. It is easy to verify that $t(x) = 2^{-r+1}$ for each $x \in [0, 1]$, in particular, $t(x)$ is constant. Now, the graph $\mathbb{G}(n, W)$ is determined already after the first step of the construction, that is, after sampling U_1, \dots, U_n . Indeed, $\mathbb{G}(n, W)$ will consist of two disconnected cliques, one consisting of the vertices $V_1 := \{i : U_i \in [0, \frac{1}{2}]\}$ and the other one consisting of the vertices $V_2 := \{i : U_i \in [\frac{1}{2}, 1]\}$. By the central limit theorem, we have $|V_1| \approx \frac{n}{2} + N \cdot \sqrt{n}$, where N is a normal random variable with mean 0 and variance $\frac{1}{4}$. Deterministically, $|V_2| = n - |V_1| \approx \frac{n}{2} - N \cdot \sqrt{n}$. Hence, the number of K_r ’s in $\mathbb{G}(n, W)$ is equal to

$$\binom{|V_1|}{r} + \binom{|V_2|}{r} \approx \binom{\frac{n}{2} + N \cdot \sqrt{n}}{r} + \binom{\frac{n}{2} - N \cdot \sqrt{n}}{r} \approx \frac{1}{r!} \cdot \left(\left(\frac{n}{2} + N \cdot \sqrt{n} \right)^r + \left(\frac{n}{2} - N \cdot \sqrt{n} \right)^r \right).$$

The first terms in the binomial expansions of these two summands add up to $\frac{2n^r}{2^r}$ which is non-random (and hence does not contribute to fluctuations). The second terms in the binomial expansions cancel out. Hence the main part of the fluctuations comes from the sum of the third terms,

$$\binom{r}{2} \left(\frac{n}{2} \right)^{r-2} (N \cdot \sqrt{n})^2 + \binom{r}{2} \left(\frac{n}{2} \right)^{r-2} (-N \cdot \sqrt{n})^2 = 2^{3-r} \binom{r}{2} \cdot N^2 \cdot n^{r-1}.$$

In particular, we see that these fluctuations are of order n^{r-1} and involve the square of a normal random variable.

Connection to generalized U-statistics. During the post-submission revision the authors became aware of a result by Janson and Nowicki [12] on orthogonal decomposition of generalized U-statistics of independent random variables indexed by both vertices and edges of a graph (see also [13]). In particular, Theorem 2 in [12] implies Theorem 1.1. To see this, one has to realize that the clique count can be expressed as a U-statistic by generating independent uniform random variables $Y_{uv} \in [0, 1]$ (independent of U_1, \dots, U_n) and defining $\mathbb{G}(n, W)$ as the graph with edge set

$$\{uv : Y_{uv} \leq W(U_u, U_v)\},$$

where $U_v, v \in [n]$ are the random types as defined above.

The contribution of this paper is thus (i) a direct proof of limit distribution using the method of moments; (ii) explicit description of the limit distribution in terms of the graphon W (which admittedly also can be read out relatively easily from the result of Janson and Nowicki); (iii) establishing

a link between asymptotic normality of clique counts and quasirandomness (see Conjecture 1.4 and the following comment).

Before stating the explicit version of Theorem 1.1 we need to introduce some definitions about spectra of graphons and more advanced concepts related to subgraph densities. On the way, we also recall facts that will be useful for the proof of the main theorem.

1.3 Spectrum of a graphon and cycle densities

Much of the spectral theory of graphs carries over to the dense graph limit setting, where graphons play the role that adjacency matrices of graphs play in algebraic graph theory. Spectral properties of graphons will be crucial in stating and proving the explicit version of Theorem 1.1(c).

In this section, we follow [14, Section 7.5], where details and proofs can be found. We work with the real Hilbert space $L^2[0, 1]$. Suppose that $W : [0, 1]^2 \rightarrow [0, 1]$ is a graphon. Then we can associate with W its kernel operator $T_W : L^2[0, 1] \rightarrow L^2[0, 1]$ by setting

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

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

$$\sum_{\lambda \in \text{Spec}(W)} \lambda^2 = \int_{[0,1]^2} W(x, y)^2 dx dy \leq 1. \quad (4)$$

The *degree function* of a graphon W is the function $\deg_W : [0, 1] \rightarrow [0, 1]$ defined as $\deg_W(x) = \int_y W(x, y) dy$. In order to gain some intuition about the degree function, the reader should note that if a vertex in $\mathbb{G}(n, W)$ is conditioned to have type x , it has expected degree $(n - 1) \cdot \deg_W(x)$. We say that W is *regular* if $\deg_W(x) \equiv d$ for some constant d . (Note that in such case $d = \int_{[0,1]^2} W(x, y) dx dy$.)

Observe that if W is regular, then $f \equiv 1$ is an eigenfunction of T_W with eigenvalue d . In this case, let $\text{Spec}^-(W)$ be $\text{Spec}(W)$ 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.)

One of the most useful properties of eigenvalues of a graphon is that they give a simple expression for cycle densities. Recall that C_ℓ is a cycle on ℓ vertices (with C_2 being a multigraph consisting of a double edge). We have (see [14, (7.22), (7.23)]) that

$$t(C_\ell, W) = \sum_{\lambda \in \text{Spec}(W)} \lambda^\ell, \quad \text{for } \ell \geq 2. \quad (5)$$

Note that for $\ell = 2$, the formula (5) is exactly (4).

The formula (5) plays a key part in the proof Theorem 1.2(c), and is the only fact that we use from the spectral theory of graphons.

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

If H is a fixed multigraph and W is a graphon, the *density* (usually called *homomorphism density*) of H in W is defined as

$$t(H, W) := \mathbb{E} \prod_{\{i,j\} \in E(H)} W(U_i, U_j). \quad (6)$$

(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

$$\mathbb{E}[X_n(H, W)] = \frac{\binom{n}{k}}{\text{aut}(H)} t(H, W),$$

where $\text{aut}(H)$ is the number of automorphisms of H , and $\binom{n}{k} = n \cdot (n-1) \cdot \dots \cdot (n-k+1)$. In particular we have that $\mathbb{E}X_n(K_r, W) = \binom{n}{r} t(K_r, W)$, a fact that will be used several times throughout the paper.

For a natural number k , we write $[k] := \{1, 2, \dots, k\}$. Given an integer $\ell \leq k$, 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 $\mathbf{x} = (x_j)_{j \in J} \in [0, 1]^\ell$, we define

$$t_{\mathbf{x}}(H, W) = \mathbb{E} \left[\prod_{\{i,j\} \in E(H)} W(U_i, U_j) \mid U_j = x_j : j \in J \right]. \quad (7)$$

Again, if H is simple with $|V(H)| = r$, then $t_{\mathbf{x}}(H, W)$ is the conditional probability that $\mathbb{G}(r, W) = H$ whenever vertex j is prescribed a type 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

$$t_x(K_r^\bullet, W) = t(K_r, W).$$

In the case $r = 2$, we have $t_x(K_2^\bullet, W) = \deg_W(x)$, hence K_2 -regularity coincides with the usual concept of regularity. For $r \geq 3$ condition of K_r -regularity means that in the random graph $\mathbb{G}(n, W)$ a vertex is expected to belong to the same number of copies of K_r , regardless of its type. Another important aspect of K_r -regularity of W is that any two particular copies of K_r that share exactly one vertex are uncorrelated, that is, in $\mathbb{G}(n, W)$ existence of one of the these two copies does not influence the probability of the existence of the other. As we will see in the proof, if W is *not* K_r -regular, then a pair of K_r -copies sharing one vertex are *positively* correlated (see (20)), intuitively causing larger variance of the K_r count.

As we will see in Theorem 1.2, the limit distribution of edges (i.e. case $r = 2$) can be described in terms of the function \deg_W and the spectrum of W . For $r \geq 3$, however, we need to consider

an auxiliary graphon $V_W^{(r)}$ defined below which encodes the information about the local clique densities in W .

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

$$V_W^{(r)}(x, y) := t_{x, y}(K_r^{\bullet\bullet}, W). \quad (8)$$

So, $V_W^{(r)}(x, y)$ is intuitively the density of K_r 's containing x and y . Note that $V_W^{(2)} = W$.

Suppose that we have two numbers $r \in \mathbb{N}$ and $j \in \{0, \dots, r\}$. We write $K_r \ominus_j K_r$ for the (simple) graph on $2r - j$ vertices consisting of two copies of K_r sharing j vertices. For $j = 2$, we also denote by $K_r \oplus_2 K_r$ the multigraph obtained from $K_r \ominus_2 K_r$ by doubling the shared edge. In particular $K_2 \ominus_2 K_2 = K_2$ and $K_2 \oplus_2 K_2 = C_2$.

Denote by K_r' and K_r'' two copies of K_r that share exactly two vertices, which have labels 1 and 2. We have

$$\prod_{ij \in K_r \oplus K_r} W(U_i, U_j) = W(U_1, U_2) \prod_{ij \in K_r \ominus_2 K_r} W(U_i, U_j) = \prod_{ij \in K_r'} W(U_i, U_j) \prod_{ij \in K_r''} W(U_i, U_j)$$

Taking conditional expectation with respect to the event $U_1 = x, U_2 = y$ and noting that products $\prod_{ij \in K_r'} W(U_i, U_j)$ and $\prod_{ij \in K_r''} W(U_i, U_j)$ are conditionally independent, from definition (7) we obtain

$$t_{x, y}(K_r \oplus_2 K_r, W) = W(x, y) t_{x, y}(K_r \ominus_2 K_r, W) = (t_{x, y}(K_r^{\bullet\bullet}, W))^2 = \left(V_W^{(r)}(x, y)\right)^2. \quad (9)$$

Let

$$\sigma_{r, W}^2 := \frac{1}{2((r-2)!)^2} (t(K_r \ominus_2 K_r, W) - t(K_r \oplus_2 K_r, W)). \quad (10)$$

We have $\sigma_{r, W}^2 \geq 0$, since, assuming that the two shared vertices have labels 1 and 2,

$$\begin{aligned} t(K_r \ominus_2 K_r, W) - t(K_r \oplus_2 K_r, W) &= \int (t_{x, y}(K_r \ominus_2 K_r) - t_{x, y}(K_r \oplus_2 K_r)) dx dy \\ &\stackrel{(9)}{=} \int [1 - W(x, y)] t_{x, y}(K_r \ominus_2 K_r) dx dy \\ &\geq \int 0 dx dy = 0. \end{aligned} \quad (11)$$

Observe that

$$\deg_{V_W^{(r)}}(x) = \int_0^1 V_W^{(r)}(x, y) dy = \int_0^1 t_{x, y}(K_r^{\bullet\bullet}, W) dy = t_x(K_r^\bullet, W).$$

So W is K_r -regular if and only if $V_W^{(r)}$ is regular, with $\deg_{V_W^{(r)}} \equiv t_r := t(K_r, W)$. Hence, by the remark we made in Section 1.3, one of the eigenvalues associated with $V_W^{(r)}$ is t_r . In this case, $\text{Spec}^-(V_W^{(r)})$ is $\text{Spec}(V_W^{(r)})$ with the multiplicity of t_r decreased by 1.

1.5 Statement of the main result

We are now ready to state our main result.

Theorem 1.2. *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 number of 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*

$$\frac{X_{n,r} - \binom{n}{r} t_r}{n^{r-\frac{1}{2}}} \xrightarrow{d} \hat{\sigma}_{r,W} \cdot Z, \quad (12)$$

where Z is a standard normal random variable and $\hat{\sigma}_{r,W} = \frac{1}{(r-1)!} (t(K_r \ominus_1 K_r, W) - t_r^2)^{1/2} > 0$.

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

$$\frac{X_{n,r} - \binom{n}{r} t_r}{n^{r-1}} \xrightarrow{d} \sigma_{r,W} \cdot Z + \frac{1}{2(r-2)!} \sum_{\lambda \in \text{Spec}^-(V_W^{(r)})} \lambda \cdot (Z_\lambda^2 - 1), \quad (13)$$

where Z and $(Z_\lambda)_{\lambda \in \text{Spec}^-(V_W^{(r)})}$ are independent standard normal and $\sigma_{r,W}$ is defined in (10). (The series on the right-hand side of (13) converges a.s. and in L^2 by Lemma 2.1.)

Let us comment briefly on Theorem 1.2. Part (a) is immediate. Part (b) tells us that in this setting we have a behaviour as in the central limit theorem; indeed we could have stated Part (b) as $(X_{n,r} - \mathbb{E}[X_{n,r}]) / \sqrt{\text{Var}[X_{n,r}]} \xrightarrow{d} Z$. Finally, we will see in Remark 1.6 below that the limit in (13) is non-degenerate, implying the Theorem 1.2 gives a complete description of limit distributions of small cliques in $\mathbb{G}(n, W)$.

Let us mention that Part (b) has been recently reported in [9], 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 — based on Stein's method — 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.

1.6 When the distribution in Theorem 1.2(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 (13) is a weighted infinite-dimensional variant of a chi-square distribution. By (4) and (17) below, this random variable has finite variance. Interestingly, very similar distributions appear in [4] and [5], also in connection with graph limits. That said, the particular setting of our paper seems to be substantially different from [4, 5].

Proposition 1.3. *In Theorem 1.2(c) the limit distribution is normal if and only if $V_W^{(r)} \equiv t_r$.*

Proof. In view of (13), a normally distributed limit implies that $\text{Spec}^-(V_W^{(r)}) = \emptyset$ and therefore $\text{Spec}(V_W^{(r)}) = \{t_r\}$. Recall that $V_W^{(r)}$ is regular with degree function being constant t_r . We claim that then $V_W^{(r)} \equiv t_r$. This claim can be viewed as a graphon version of a consequence of Chung–Graham–Wilson Theorem on quasirandom graph sequences¹; here we give a short self-contained proof. Indeed, regularity of $V_W^{(r)}$ implies $\deg_{V_W^{(r)}} \equiv t_r$, so we have that

$$\begin{aligned} t_r^2 &= \left(\int_{[0,1]} \deg_{V_W^{(r)}}(y) dy \right)^2 = \left(\int_{[0,1]^2} V_W^{(r)}(x, y) dx dy \right)^2 \\ &\stackrel{\text{Jensen's inequality}}{\leq} \int_{[0,1]^2} \left(V_W^{(r)}(x, y) \right)^2 dx dy \stackrel{(4)}{=} \sum_{\lambda \in \text{Spec}(V_W^{(r)})} \lambda^2 = t_r^2. \end{aligned}$$

Since the quadratic function is strictly convex, equality in Jensen's inequality is attained if and only if $V_W^{(r)}$ is constant, which implies $V_W^{(r)} \equiv t_r$. \square

So, the question now is which graphons W lead to a constant graphon $V_W^{(r)}$. Since $V_W^{(2)} = W$, for $r = 2$ the answer is clearly given by constant graphons W . For $r \geq 3$, we put forward the following conjecture, which was first hinted in concluding remarks of [17].

Conjecture 1.4. *Suppose that $r \geq 3$ and $V_W^{(r)}$ is a constant- d graphon for some $d \in [0, 1]$, that is, $t_{\cdot, \cdot}(K_r^{\bullet\bullet}, W) \equiv d$. Then W is K_r -free (when $d = 0$), or $W \equiv d^{1/\binom{r}{2}}$.*

In [17], 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 1.2(c) is when W is a constant graphon. Conjecture 1.4 can also be rephrased as follows: among random graphs $\mathbb{G}(n, W)$, where W is a K_r -regular graphon, only $\mathbb{G}(n, p)$, $p \in (0, 1)$, has asymptotically normal count of K_r .

Let us now comment on a complementary question: when is the normal term absent in (13)?

Proposition 1.5. *We have $\sigma_{r, W} = 0$ if and only if $W(x, y) = 1$ for almost every $(x, y) \in [0, 1]^2$ for which $t_{x, y}(K_r^{\bullet\bullet}, W) > 0$.*

Proof. We first observe that $t_{x, y}(K_r^{\bullet\bullet}, W) > 0$ if and only if $t_{x, y}(K_r \ominus_2 K_r, W) > 0$. For $W(x, y) > 0$ this is immediate from the second equality in (9), while for $W(x, y) = 0$ from definition of $t_{x, y}(H, W)$ it is clear that both $t_{x, y}(K_r \ominus_2 K_r, W) = 0$ and $t_{x, y}(K_r^{\bullet\bullet}, W) = 0$. From (10) and (11), we have $\sigma_{r, W} = 0$ if and only if $W(x, y) = 1$ for almost every $(x, y) \in [0, 1]^2$ for which $t_{x, y}(K_r \ominus_2 K_r, W) > 0$. Together with the first observation this proves the proposition. \square

For $r = 2$ the condition of Proposition 1.5 is equivalent to a condition that $W \in \{0, 1\}$ 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

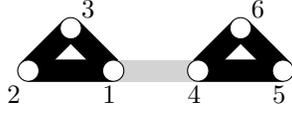


Figure 1: Example of K_3 -regular graphon (depicted here as a graph) with no linear term in the non-normal limit.

4th set, and set the rest of the connections between the first 3 and the last 3 sets to 0 (see Figure 1). Such a graphon W is K_3 -regular but we have $\sigma_{r,W} = 0$.

Remark 1.6. *The limit in (13) is never degenerate. Indeed, if the non-normal part vanishes, then Proposition 1.3 implies $V_W^{(r)} \equiv t_r \in (0, 1)$. Further, from (9) we obtain that almost everywhere*

$$W(x, y)t_{x,y}(K_r \ominus_2 K_r, W) = \left(V_W^{(r)}(x, y) \right)^2 = t_r^2 > 0.$$

In particular this implies $t_{x,y}(K_r \ominus_2 K_r, W) > 0$ almost everywhere. Moreover, we have $W < 1$ on a set of positive measure (otherwise $W \equiv 1$ and we would be in the case (a)), which implies that the inequality in (11) is strict. Since the left-hand side of (11) is a multiple of $\sigma_{r,W}^2$, this implies that $\sigma_{r,W}^2 > 0$.

1.7 Acknowledgment

We thank the anonymous referee for a very detailed report which helped us to improve the exposition.

2 Preliminaries

In this section we state definitions and facts needed in the proof of Theorem 1.2. It can be skipped at the first reading.

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 $n \rightarrow \infty$.

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 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.

¹more precisely, the part “each spot has the same density \iff non-principal eigenvalues are small”, denoted as P_4 and P_3 in [1, page 158]

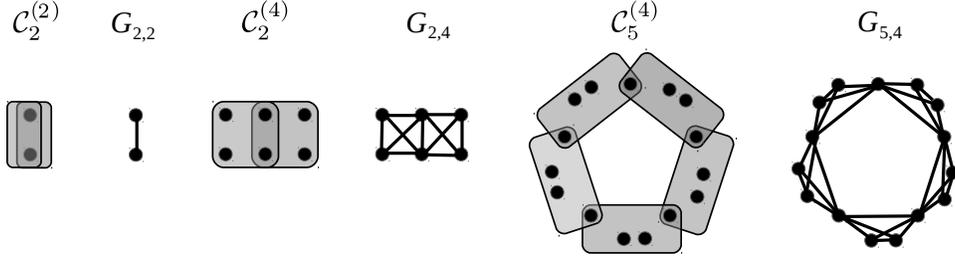


Figure 2: Examples of hypergraphs $C_\ell^{(r)}$ and their associated graphs $G_{\ell,r}$.

We use a particular hypergraph version of cycles, known as loose cycles. For $\ell \geq 2$, let $C_\ell^{(r)}$ 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 $C_2^{(r)}$ is a pair of r -sets sharing exactly 2 vertices). Finally, let $G_{\ell,r}$ be the graph associated with $C_\ell^{(r)}$. Note that $G_{2,r} = K_r \oplus_2 K_r$. See Figure 2 for examples.

We can express the densities of cycles in the graphon $V_W^{(r)}$, defined in (8), in terms of the graphon W as follows. Assume that the vertices shared by consecutive r -cliques in $G_{\ell,r}$ have labels $1, \dots, \ell$. Denoting $F(x_1, \dots, x_r) = \prod_{1 \leq i < j \leq r} W(x_i, x_j)$, we write

$$V_W^{(r)}(x_i, x_{i+1}) = \int F(x_i, x_{i+1}, y_{i,3}, \dots, y_{i,r}) \prod_{j=3}^r dy_{i,r},$$

whence, writing indices of x_i 's modulo r and using Fubini's theorem,

$$\begin{aligned} t(C_\ell, V_W^{(r)}) &= \int \prod_{i=1}^r \left(\int F(x_i, x_{i+1}, y_{i,3}, \dots, y_{i,r}) \prod_{j=3}^r dy_{i,j} \right) \prod_{i=1}^r dx_i \\ &= \int \prod_{i=1}^r F(x_i, x_{i+1}, y_{i,3}, \dots, y_{i,r}) dx_1 \dots dx_r dy_{1,3} \dots dy_{1,r} \dots dy_{r,3} \dots dy_{r,r}. \end{aligned} \quad (14)$$

For $\ell \geq 3$, a moment of thought reveals that the last integral is exactly $t(G_{\ell,r}, W)$, implying

$$t(C_\ell, V_W^{(r)}) = t(G_{\ell,r}, W), \quad \text{for } \ell \geq 3 \text{ and } r \geq 2. \quad (15)$$

On the other hand, for $\ell = 2$ we have to be a bit more careful, noting that in the product (14) factor $W(x_1, x_2)$ appears twice, implying

$$t(C_2, V_W^{(r)}) = t(K_r \oplus_2 K_r, W), \quad \text{for } r \geq 2, \quad (16)$$

where, recall, $K_r \oplus_2 K_r$ is a multigraph obtained by gluing two cliques along two vertices (with a double edge between these vertices).

2.2 Moment generating functions

As we noted above, the main result of [4] expresses a particular random variable as a sum of squares of independent normal random variables, which is very similar to the expression appearing in our Theorem 1.2(c). The following lemma asserts that such distributions are well-defined and gives a formula of their moment generating functions. Note that in [4] only the convergence of the series in L^1 is asserted, but it clearly converges in L^2 as well, since $S_n := \sum_{j=1}^n \lambda_j (Z_j^2 - 1)$ is easily seen to be a Cauchy sequence in the (complete) space L^2 . In particular this implies that $\text{Var}[S_n] \rightarrow \text{Var}[S]$.

Lemma 2.1 (see [4], Proposition 7.1²). *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^2 and*

$$\text{Var}[S] = 2 \sum_j \lambda_j^2. \quad (17)$$

Furthermore, the moment generating function $M_S(t) := \mathbb{E}[\exp(tS)]$ 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 (in fact it gives much more, by providing bounds on approximation errors). It is particularly efficient for sums of dependent random variables, when each variable depends on a relatively small (but not necessarily bounded) number of other variables. Here, we follow a survey article by Ross [18, Section 3].

The dependency structure of a collection of random variables $(Y_i : i \in I)$ is encoded by a *dependency graph* \mathcal{G} (on a vertex set I) if, for all $i \in I$, Y_i is independent of the random variables $\{Y_j\}_{j \notin N_i}$, where N_i is the neighborhood of i in \mathcal{G} (including i itself). In general, a dependency graph is not uniquely determined, but in many scenarios, there exists a 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 by $d_{\text{Wass}}(X, Y)$. We do not need an exact definition — which the reader can find on page 214 of [18] — since the only property of the Wasserstein distance that we will use is that for $Z \sim \mathcal{N}(0, 1)$ and a sequence X_n of random variables $d_{\text{Wass}}(X_n, Z) \rightarrow 0$ implies that $X_n \xrightarrow{d} Z$ (see [18, Section 3]). We will use the following off-the-shelf bound for the Wasserstein distance based on Stein's method.

Theorem 2.2 (see Theorem 3.6 in [18]). *Let $(Y_i : i \in I)$ be a finite collection of random variables such that for every $i \in I$ we have $\mathbb{E}[Y_i] = 0$ and $\mathbb{E}[Y_i^4] < \infty$. Writing $\sigma^2 := \text{Var}[\sum_{i \in I} Y_i]$, define $Q = \sum_{i \in I} Y_i / \sigma$.*

²Note that there is an error in the arXiv version of Proposition 7.1 in [4].

Let \mathcal{G} be a dependency graph of $(Y_i : i \in I)$. Writing $D = \max_{i \in I} |N_i|$, we have

$$d_{\text{Wass}}(Q, Z) \leq \frac{D^2}{\sigma^3} \cdot \sum_i \mathbb{E}[|Y_i|^3] + \frac{\sqrt{28}D^{3/2}}{\sqrt{\pi} \cdot \sigma^2} \cdot \sqrt{\sum_i \mathbb{E}[Y_i^4]}. \quad (18)$$

3 Proof of Theorem 1.2(b)

We will apply Theorem 2.2 to a collection $(Y_R : R \in \binom{[n]}{r})$ of random variables indexed by r -sets R of vertices. Writing I_R for the indicator that R induces a clique in $\mathbb{G}(n, W)$, we define $Y_R = I_R - \mathbb{E}[I_R] = I_R - t_r$. In this notation we have $X_{n,r} - \binom{[n]}{r} t_r = \sum_{R \in \binom{[n]}{r}} Y_R$.

We let \mathcal{G} be the natural dependency graph of the collection $(Y_R : R \in \binom{[n]}{r})$ with edges corresponding to pairs $R_1 R_2$ such that $R_1 \cap R_2 \neq \emptyset$.

The proof consists of two steps. In the first one we bound the maximum degree of \mathcal{G} . This step is fairly straightforward. In the second and less elementary step we compute the asymptotics of the variance of $\sum_R Y_R$. As we will see, the leading term in the asymptotics comes from pairs of cliques that share exactly one vertex.

Let us proceed with the details of the first step. Notice that in \mathcal{G} every neighbourhood N_R has the same size D , namely,

$$D = \sum_{\ell=1}^r \binom{r}{\ell} \binom{n-r}{r-\ell} = O(n^{r-1}). \quad (19)$$

We now turn to the second step. Recall that $K_r \ominus_j K_r$ is a simple graph consisting of two r -cliques sharing j vertices. Set $d_j = t(K_r \ominus_j K_r, W)$, for $j \in [r]$. By Jensen's inequality we have that

$$t_r^2 = \left(\int_0^1 t_x(K_r^\bullet, W) dx \right)^2 < \int_0^1 t_x(K_r^\bullet, W)^2 dx = d_1, \quad (20)$$

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

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 (20) implies $\hat{\sigma}_{r,W}^2 = (d_1 - t_r^2)/((r-1)!)^2 > 0$. Since $\mathbb{E}[Y_R] = 0$ for every R , we have $\sigma_n^2 := \text{Var}[\sum_R Y_R] = \sum_{R_1, R_2} \mathbb{E}[Y_{R_1} Y_{R_2}]$. Since the number of ordered pairs R_1, R_2 such that $|R_1 \cap R_2| = \ell$ is

$$\binom{n}{\ell} \binom{n-\ell}{r-\ell} \binom{n-r}{r-\ell} \sim \frac{n^{2r-\ell}}{\ell!(r-\ell)!^2},$$

we obtain

$$\begin{aligned} \sigma_n^2 &= \sum_{\ell=1}^r \frac{n^{2r-\ell}(1+o(1))}{\ell!(r-\ell)!^2} (d_\ell - t_r^2) \\ &= \frac{n^{2r-1}}{(r-1)!^2} (d_1 - t_r^2) + \sum_{\ell=2}^r O(n^{2r-\ell}) \\ &\sim \hat{\sigma}_{r,W}^2 n^{2r-1}. \end{aligned} \quad (21)$$

Writing $Q_n = \sum_{R \in \binom{[n]}{r}} Y_R / \sigma_n$, we are ready to apply Theorem 2.2. Crudely bounding each of the sums of moments on the right-hand side of (18) by $\binom{n}{r} \leq n^r$, and using (19) and (21), we obtain that $d_{\text{Wass}}(Q_n, Z) = O(n^{-1/2}) \rightarrow 0$. By the remark we made just before Theorem 2.2, we conclude that $Q_n \xrightarrow{d} Z$. In view of (21) and Slutsky's theorem,

$$\frac{\sum_{R \in \binom{[n]}{r}} Y_R}{n^{r-1/2}} = \frac{\sigma_n}{n^{r-1/2}} \cdot Q_n \xrightarrow{d} \hat{\sigma}_{r,W} Z,$$

which completes the proof, in view of $\sum_{R \in \binom{[n]}{r}} Y_R = X_{n,r} - \binom{n}{r} t_r$.

4 Proof of Theorem 1.2(c)

Define a random variable as on the right-hand side of (13),

$$Y = \sigma_{r,W} \cdot Z + \frac{1}{2(r-2)!} \sum_{\lambda \in \text{Spec}^-(V_W^{(r)})} \lambda \cdot (Z_\lambda^2 - 1). \quad (22)$$

We employ the method of moments, in the way it is described in Section 6.1 of [11]. For this it is enough to show that the moments of the random variable $(X_{n,r} - \binom{n}{r} t_r) / n^{r-1}$ converge to the corresponding moments of the random variable Y , 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 $M_Z(x) = \exp(x^2/2)$ 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 (22) is $\prod_{\lambda \in \text{Spec}^-(V_W^{(r)})} \exp\left(-\frac{\lambda x}{2(r-2)!}\right) / \sqrt{1 - \frac{\lambda x}{(r-2)!}}$. 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

$$M_Y(x) = \exp\left(\frac{\sigma_{r,W}^2 x^2}{2}\right) \prod_{\lambda \in \text{Spec}^-(V_W^{(r)})} \frac{\exp\left(-\frac{\lambda x}{2(r-2)!}\right)}{\sqrt{1 - \frac{\lambda x}{(r-2)!}}}. \quad (23)$$

On the other hand, to compute the moments of $(X_{n,r} - \binom{n}{r} t_r) / n^{r-1}$, we write

$$X_{n,r} - \binom{n}{r} t_r = \sum_{R \in \binom{[n]}{r}} (I_R - t_r),$$

where, as in Section 3, I_R is 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, \dots, R_m) of (not necessary distinct) elements of $\binom{[n]}{r}$, let

$$\Delta(R_1, \dots, R_m) := \mathbb{E} \left[\prod_{i=1}^m (I_{R_i} - t_r) \right]. \quad (24)$$

Then

$$\mathbb{E} \left[\left(X_{n,r} - \binom{n}{r} t_r \right)^m \right] = \sum_{(R_1, \dots, R_m) \in \binom{[n]}{r}^m} \Delta(R_1, \dots, R_m). \quad (25)$$

It is hence natural to analyze $\Delta(R_1, \dots, R_m)$ depending on the structure of the tuple (R_1, \dots, R_m) . Our plan is as follows. First, we introduce a certain family $\mathfrak{X}(n, r, m) \subset \binom{[n]}{r}^m$ and show that $\Delta(R_1, \dots, R_m) = 0$ for each $(R_1, \dots, R_m) \in \mathfrak{X}(n, r, m)$. Next, we define another family $\mathfrak{F}(n, r, m) \subset \binom{[n]}{r}^m \setminus \mathfrak{X}(n, r, m)$. In (30) we will show that the set $\binom{[n]}{r}^m \setminus (\mathfrak{X}(n, r, m) \cup \mathfrak{F}(n, r, m))$ has size $O(n^{(r-1)m-1})$, and hence the corresponding tuples have a contribution which is negligible with respect to the renormalization of (25) by $n^{-(r-1)m}$ (as, for example, in the statement of Claim 4.5). We then classify the tuples in $\mathfrak{F}(n, r, m)$ according to the pattern in which they overlap so that in each class every tuple the contribution $\Delta(R_1, \dots, R_m)$ is the same. By obtaining an explicit expression for this contribution (in Claim 4.3) and counting the number of tuples in each class (in Claim 4.4) we will arrive at a rather complicated asymptotic formula, which we interpret as a coefficient of some reasonably simple power series (see Claim 4.5). Finally, with some luck we discover that this power series is exactly the moment generating function M_Y (see Claim 4.6). Let us give details now.

Let $\mathfrak{X}(n, r, m) \subset \binom{[n]}{r}^m$ be those m -tuples (R_1, \dots, 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 $(R_1, \dots, R_m) \in \mathfrak{X}(n, r, m)$. 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

$$\Delta(R_1, \dots, R_m) = \int_0^1 \mathbb{E} [I_{R_m} - t_r \mid U_v = x] \cdot \mathbb{E} \left[\prod_{i=1}^{m-1} (I_{R_i} - t_r) \mid U_v = x \right] dx. \quad (26)$$

Since W is K_r -regular, we have $\mathbb{E} [I_{R_m} - t_r \mid U_v = x] = t_x(K_r^\bullet, W) - t_r = 0$ for almost every x . Therefore $\Delta(R_1, \dots, R_m) = 0$, as claimed. An even simpler calculation yields the same conclusion when $|R_m \cap (\cup_{j=1}^{m-1} R_j)| = 0$. Hence, we can rewrite (25) as

$$\mathbb{E} \left[\left(X_{n,r} - t_r \binom{n}{r} \right)^m \right] = \sum_{(R_1, \dots, R_m) \in \binom{[n]}{r}^m \setminus \mathfrak{X}(n, r, m)} \Delta(R_1, \dots, R_m). \quad (27)$$

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

The following claim provides a sharp upper bound on the number of vertices in $\mathcal{H}(R_1, \dots, R_m)$, for $(R_1, \dots, R_m) \in \binom{[n]}{r}^m \setminus \mathfrak{X}(n, r, m)$.

Claim 4.1. *Suppose that $(R_1, \dots, R_m) \in \binom{[n]}{r}^m \setminus \mathfrak{X}(n, r, m)$. The number v of vertices in the hypergraph $\mathcal{H} = \mathcal{H}(R_1, \dots, R_m)$ 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 $(R_1, \dots, R_m) \notin \mathfrak{X}(n, r, m)$ we have that

$$k \leq (r-2)m. \quad (28)$$

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

$$k + 2(v - k) \leq \sum_{v \in V(\mathcal{H})} \deg(v) = rm. \quad (29)$$

Therefore

$$v \stackrel{(29)}{\leq} \frac{rm - k}{2} + k = \frac{rm + k}{2} \stackrel{(28)}{\leq} \frac{rm + (r - 2)m}{2} = (r - 1)m,$$

and the first statement follows. To prove the second statement, suppose that the number of vertices in $\mathcal{H} = \mathcal{H}(R_1, \dots, R_m)$ satisfies $v = (r - 1)m$. Then (28) and (29) are both equalities and so \mathcal{H} consists of $(r - 2)m$ vertices of degree 1, and m vertices of degree 2. Since, by assumption, every R_i contains at least two vertices of degree 2 it readily follows that it contains exactly two such vertices. The other implication is immediate. \square

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

$$\left| \binom{[n]}{r}^m \setminus (\mathfrak{X}(n, r, m) \cup \mathfrak{F}(n, r, m)) \right| \leq \binom{n}{(r - 1)m - 1} \cdot (2^m)^{(r - 1)m - 1} = O(n^{(r - 1)m - 1}). \quad (30)$$

Since $|\Delta(R_1, \dots, R_m)| \leq 1$, from (27) and (30) we infer

$$\mathbb{E} \left[\left(X_{n,r} - t_r \binom{n}{r} \right)^m \right] = \sum_{(R_1, \dots, R_m) \in \mathfrak{F}(n, r, m)} \Delta(R_1, \dots, R_m) + O(n^{(r - 1)m - 1}). \quad (31)$$

Now, fix $(R_1, \dots, R_m) \in \mathfrak{F}(n, r, m)$ and consider the hypergraph $\mathcal{H} = \mathcal{H}(R_1, \dots, R_m)$. Notice that when $r = 2$ then some edges in \mathcal{H} may be double edges, but all hyperedges are simple when $r \geq 3$. Now, replace every r -edge, say R , of \mathcal{H} by a 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.

We now need to deal with the right-hand side of (24) for tuples in $\mathfrak{F}(n, r, m)$, that is for tuples corresponding to unions of loose cycles. We first note that we can factor it over cycles, namely if there is a partition $[m] = V_1 \cup \dots \cup V_k$ such that each $\mathcal{H}(R_i : i \in V_j)$ is a loose cycle, then the random variables $\prod_{i \in V_j} (I_{R_i} - t_r)$, $j = 1, \dots, k$ are independent and hence

$$\Delta(R_1, \dots, R_m) = \prod_j \Delta(R_i : i \in V_j). \quad (32)$$

In order to calculate the individual factors in (32), the following claim will be useful.

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

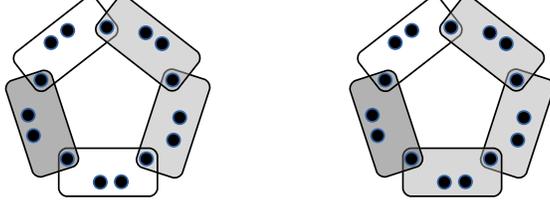


Figure 3: Two examples of a hypergraph \mathcal{C}^- (light grey) and an edge S (dark grey) in a hypergraph $\mathcal{C}_5^{(4)}$ in the proof of Claim 4.2.

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 $\mathcal{C} \neq \emptyset$.

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 $|S \cap \bigcup \mathcal{C}^-| \leq 1$ (here and below $\bigcup \mathcal{H}$ stands for the union of the hyperedges of \mathcal{H}). See Figure 3. 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 (26), we have

$$\mathbb{E} [\prod_{R \in \mathcal{C}} I_R] = \int_0^1 \mathbb{E} [I_S | U_v = x] \cdot \mathbb{E} [\prod_{R \in \mathcal{C}^-} I_R | U_v = x] dx,$$

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

$$\mathbb{E} [\prod_{R \in \mathcal{C}} I_R] = \int_0^1 t_r \mathbb{E} [\prod_{R \in \mathcal{C}^-} I_R | U_v = x] dx = t_r \mathbb{E} [\prod_{R \in \mathcal{C}^-} I_R] = t_r \cdot t_r^{|\mathcal{C}^-|},$$

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

$$\mathbb{E} [\prod_{R \in \mathcal{C}} I_R] = \mathbb{E} [I_S] \cdot \mathbb{E} [\prod_{R \in \mathcal{C}^-} I_R] = t_r \cdot t_r^{|\mathcal{C}^-|}.$$

□

Recall that for each $(R_1, \dots, R_m) \in \mathfrak{F}(n, r, m)$, the hypergraph $\mathcal{H}(R_1, \dots, R_m)$ is a union of loose cycles. Isomorphism classes of such hypergraphs can be encoded by a vector whose entry at position i is the number of loose cycles of length i . More precisely, let us consider the following set of $(m-1)$ -dimensional vectors,

$$\mathcal{V}_m := \left\{ \mathbf{k} = (k_2, \dots, k_m) \in \mathbb{N}_0^{m-1} : \sum_{i=2}^m i k_i = m \right\}. \quad (33)$$

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 $i = 2, \dots, m$.

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

$$\Delta(R_1, \dots, R_m) = \prod_{\ell=2}^m (t(G_{\ell,r}, W) - t_r^\ell)^{k_\ell}, \quad (34)$$

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

Proof of Claim 4.3. In view of (32), it is enough to prove (34) for an ℓ -tuple $(Q_1, \dots, Q_\ell) \in \mathfrak{F}(n, r, \ell)$ for which $\mathcal{H}(Q_1, \dots, Q_\ell)$ is isomorphic to $\mathcal{C}_\ell^{(r)}$. We have

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

□

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

$$A(n, r, \mathbf{k}) := \frac{m! \cdot \binom{n}{(r-1)m}}{\prod_{\ell=2}^m (2\ell((r-2)!)^\ell)^{k_\ell} \cdot k_\ell!}. \quad (35)$$

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

$$\text{aut}(\mathcal{H}_{\mathbf{k}}^{(r)}) = \prod_{\ell=2}^m (2\ell((r-2)!)^\ell)^{k_\ell} \cdot k_\ell!.$$

As there are $\frac{\binom{n}{(r-1)m}}{\text{aut}(\mathcal{H}_{\mathbf{k}}^{(r)})}$ 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 expressing $\mathbb{E}[(X_{n,r} - t_r \binom{n}{r})^m]$, which we abandoned at (31). Recall that $G_{\ell,r}$ is the graph associated with $\mathcal{C}_\ell^{(r)}$. Adding (34) and (35), we get

$$\begin{aligned} \mathbb{E}[(X_{n,r} - t_r \binom{n}{r})^m] &= \sum_{\mathbf{k} \in \mathcal{V}_m} A(n, r, \mathbf{k}) \prod_{\ell=2}^m (t(G_{\ell,r}, W) - t_r^\ell)^{k_\ell} + O(n^{(r-1)m-1}) \\ &= n^{(r-1)m} m! \sum_{\mathbf{k} \in \mathcal{V}_m} \prod_{\ell=2}^m \left(\frac{t(G_{\ell,r}, W) - t_r^\ell}{2\ell((r-2)!)^\ell} \right)^{k_\ell} \frac{1}{k_\ell!} + O(n^{(r-1)m-1}). \end{aligned} \quad (36)$$

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

$$d_\ell := \frac{t(G_{\ell,r}, W) - t_r^\ell}{2\ell((r-2)!)^\ell}. \quad (37)$$

Treating $g(x) := \sum_{\ell=2}^{\infty} d_\ell x^\ell$ as a formal power series, and substituting $y = g(x)$ in $e^y = \sum_{j=0}^{\infty} y^j / j!$ we obtain another power series (since the free coefficient of g is zero),

$$f(x) := \exp(g(x)) = \exp\left(\sum_{\ell=2}^{\infty} d_\ell x^\ell\right). \quad (38)$$

The following two claims are needed to show that $f(x)$ is the moment generating function of the limit of $(X_{n,r} - \binom{n}{r} t_r) / n^{r-1}$.

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

Proof. We have

$$\begin{aligned} \llbracket x^m \rrbracket f(x) &= \llbracket x^m \rrbracket \left(\exp\left(\sum_{\ell=2}^{\infty} d_\ell x^\ell\right) \right) = \llbracket x^m \rrbracket \left(\sum_{j=1}^{\infty} \frac{1}{j!} \cdot \left(\sum_{\ell=2}^{\infty} d_\ell x^\ell\right)^j \right) \\ &\stackrel{\text{multinomial theorem}}{=} \sum_{j=1}^m \frac{1}{j!} \sum_{\mathbf{k} \in \mathcal{V}_m: k_2 + \dots + k_m = j} \binom{j}{k_2, \dots, k_m} \cdot \prod_{\ell=2}^m d_\ell^{k_\ell} \\ &\stackrel{\text{definition of the multinomial coefficient}}{=} \sum_{j=1}^m \sum_{\mathbf{k} \in \mathcal{V}_m: k_2 + \dots + k_m = j} \prod_{\ell=2}^m \frac{d_\ell^{k_\ell}}{k_\ell!} \\ &= \sum_{\mathbf{k} \in \mathcal{V}_m} \prod_{\ell=2}^m d_\ell^{k_\ell} \frac{1}{k_\ell!} \\ &\stackrel{\text{by (36)}}{=} \lim_{n \rightarrow \infty} \frac{1}{m!} \cdot \frac{\mathbb{E}[(X_{n,r} - t_r \binom{n}{r})^m]}{n^{(r-1)m}}. \end{aligned}$$

□

Since $|d_i| \leq ((r-2)!)^{-i}$, in particular, the sequence $|d_i|, i \geq 2$ is bounded. Therefore the series $\sum_i d_i x^i$ has positive radius of convergence, and $f(x)$ can be expanded as its Taylor series around zero. In the next claim, we show that the function f equals the moment-generating function M_Y defined in (23).

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

Proof of Claim 4.6. Recall the definition (37) of d_ℓ . For $\ell = 2$ we have

$$\begin{aligned}
4((r-2)!)^2 \cdot d_2 &= t(K_r \ominus_2 K_r, W) - t_r^2 \\
&= t(K_r \ominus_2 K_r, W) - t(C_2, V_W^{(r)}) + t(C_2, V_W^{(r)}) - t_r^2 \\
\boxed{(16)} \quad &= t(K_r \ominus_2 K_r, W) - t(K_r \oplus_2 K_r, W) + t(C_2, V_W^{(r)}) - t_r^2 \\
\boxed{(5)} \quad &= t(K_r \ominus_2 K_r, W) - t(K_r \oplus_2 K_r, W) + \sum_{\lambda \in \text{Spec}^-(V_W^{(r)})} \lambda^2,
\end{aligned}$$

which, in view of the definition (10) of $\sigma_{r,W}^2$, implies

$$d_2 = \frac{\sigma_{r,W}^2}{2} + \frac{1}{4} \sum_{\lambda \in \text{Spec}^-(V_W^{(r)})} \left(\frac{\lambda}{(r-2)!} \right)^2. \quad (39)$$

For $\ell \geq 3$, from (15) and (5) it follows

$$\begin{aligned}
d_\ell &= \frac{t(G_{\ell,r}, W) - t_r^\ell}{2\ell((r-2)!)^\ell} \\
&= \frac{t(C_\ell, V_W^{(r)}) - t_r^\ell}{2\ell((r-2)!)^\ell} = \frac{1}{2\ell} \sum_{\lambda \in \text{Spec}^-(V_W^{(r)})} \left(\frac{\lambda}{(r-2)!} \right)^\ell.
\end{aligned} \quad (40)$$

We are now ready to relate f and M_Y . In view of (4), we have

$$\sum_{\lambda \in \text{Spec}^-(V_W^{(r)})} \lambda^2 < \infty \quad (41)$$

and in particular

$$\sup_{\lambda \in \text{Spec}^-(V_W^{(r)})} |\lambda| < \infty. \quad (42)$$

Substituting (39) and (40) into (38), we obtain

$$\log f(x) = \sum_{\ell=2}^{\infty} d_\ell x^\ell = \frac{\sigma_{r,W}^2 x^2}{2} + \sum_{\ell=2}^{\infty} \sum_{\lambda \in \text{Spec}^-(V_W^{(r)})} \frac{1}{2\ell} \left(\frac{\lambda x}{(r-2)!} \right)^\ell. \quad (43)$$

In order to interchange the order of summation in (43), we check a condition that allows applying Fubini's theorem, namely that (43) remains finite, if we replace all summands by their absolute values. By (42), for $|x|$ small enough, the sequence $a_\lambda := \frac{\lambda x}{(r-2)!}$ satisfies $c := \sup_\lambda |a_\lambda| < 1$. Therefore, for each $\ell \geq 2$,

$$\sum_\lambda |a_\lambda|^\ell \leq \sum_\lambda c^{\ell-2} a_\lambda^2 \stackrel{(41)}{=} O(c^\ell),$$

which implies $\sum_{\ell} \sum_{\lambda} |a_{\lambda}|^{\ell} / (2\ell) = \sum_{\ell} O(c^{\ell}) < \infty$, verifying the desired condition. So, changing the order of summation in (43), we obtain

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

By exponentiating the above expression we easily obtain (23), thus completing the proof. \square

We are now finished with the proof of Theorem 1.2(c). Indeed, Claims 4.5 and 4.6 imply that the m th moment of $(X_{n,r} - \binom{n}{r} t_r) / n^{r-1}$ converges to $m! \llbracket x^m \rrbracket f(x) = m! \llbracket x^m \rrbracket M_Y(x) = \mathbb{E}[Y^m]$ for every m . As argued at the beginning of Section 4, this implies that $\frac{X_{n,r} - \binom{n}{r} t_r}{n^{r-1}} \xrightarrow{d} Y$.

5 Concluding remarks

In this paper, we initiated the study of limit theorems for complete 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 1.2 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 $p = p(n) \rightarrow 0$. 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 [6]. 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 for cliques [3] and even more recently for general connected graphs [20]. 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 of the form $\binom{k}{2} + \binom{n-k}{2}$, $k \in \mathbb{N}$, that is, not all integer values can be achieved (including those in the bulk of the distribution).

References

- [1] N. Alon, J. H. Spencer, *The Probabilistic Method*, Fourth edition, John Wiley and Sons, 2015.
- [2] 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.
- [3] R. Berkowitz, *A local limit theorem for cliques in $G(n, p)$* , (preprint) arXiv:1811.03527.

- [4] 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.
- [5] B.B. Bhattacharya, S. Mukherjee, *Monochromatic subgraphs in randomly colored graphons*, *European Journal of Combinatorics*, **81**, 328–353.
- [6] B. Bollobás, S. Janson, O. Riordan, *The phase transition in inhomogeneous random graphs*, *Random Structures & Algorithms*, **31** (1) (2007) 3–122.
- [7] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, K. Vesztegombi, K., *Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing*, *Advances in Mathematics*, **219** (6) (2008) 1801–1851.
- [8] M. Doležal, J. Hladký, A. Máthé, *Cliques in dense inhomogeneous random graphs*, *Random Structures & Algorithms*, **51** (2) (2017) 275–314.
- [9] V. Féray, P.-L. Méliot, A. Nikeghbali, *Graphons, permutons and the Thoma simplex: three mod-Gaussian moduli spaces*, *Proc. Lond. Math. Soc.* **121** (2020), no. 4, 876–926.
- [10] N. Fraiman, D. Mitsche, *The diameter of inhomogeneous random graphs*, *Random Structures & Algorithms*, **53** (2) (2018) 308–326.
- [11] S. Janson, T. Łuczak, A. Ruciński, *Random Graphs*, John Wiley and Sons, 2011.
- [12] S. Janson, K. Nowicki, *The asymptotic distributions of generalized U-statistics with applications to random graphs*, *Probability theory and related fields*, **90** (3) (1991), 341–375.
- [13] G. Kaur, A. Röllin. *Higher-order fluctuations in dense random graph models*, (preprint) arXiv:2006.15805.
- [14] L. Lovász, *Large networks and graph limits*, Vol. 60 of American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, 2012.
- [15] L. Lovász, B. Szegedy, *Limits of dense graph sequences*, *Journal of Combinatorial Theory, Series B*, **96** (6) (2006), 933–957.
- [16] K. Nowicki, J.C. Wierman, *Subgraph Counts in Random Graphs Using Incomplete U-statistics methods*, *Annals of Discrete Mathematics*, **38** (1988) 299–310.
- [17] C. Reiher, M. Schacht, *Forcing quasirandomness with triangles*, *Forum of Mathematics, Sigma* **7** (2019) E9. doi:10.1017/fms.2019.7
- [18] N. Ross, *Fundamentals of Stein’s method*, *Probability Surveys* **8** (2011), 210–293.
- [19] A. Ruciński, *When Are Small Subgraphs of a Random Graph Normally Distributed?*, *Probability Theory and Related Fields* **78** (1988) 1–10.
- [20] A. Sah, M. Sawhney, *Local limit theorems for subgraph counts*, (preprint) arXiv:2006.11369.