Equivalence of the random intersection graph and $G(n, p)$

Katarzyna Rybarczyk*

*Faculty of Mathematics and Computer Science, Adam Mickiewicz University, 60–769 Poznań, Poland

Abstract

We solve the conjecture of Fill, Scheinerman and Singer-Cohen posed in [4] and show the equivalence of the sharp threshold functions of the random intersection graph $G(n, m, p)$ with $m \geq n^3$ and a graph $G(n, \hat{p})$ with independent edges. Moreover we prove sharper equivalence results under some additional assumptions.

keywords: random intersection graph, equivalence, graph properties

1 Introduction

In an intersection graph there is a set of vertices $V$ and an auxiliary set of objects $W$. Each vertex $v \in V$ is assigned a subset of objects $W(v) \subseteq W$ according to a given probability measure. Two vertices $v_1$, $v_2$ are adjacent in a random intersection graph if and only if $W(v_1) \cap W(v_2) \neq \emptyset$. A general model of the random intersection graph, in which each vertex is assigned a subset of objects $W(v) \subseteq W$ chosen uniformly from all $d$–element subsets, where the cardinality $d$ is determined according to the arbitrarily given probability distribution, was introduced in [5].

We will concentrate on analyzing the properties of the random intersection graph in which the cardinality $d$ is chosen according to the binomial distribution. Namely, we will investigate properties of the random intersection graph $G(n, m, p)$ introduced in [7]. $G(n, m, p)$ is a graph with number of vertices $|V| = n$, number of objects $|W| = m$, in which each feature $w$ is added to $W(v)$ with probability $p$ independently for all $v \in V$ and $w \in W$ (i.e. $\Pr \{w \in W(v)\} = p$). However, to some extent, the results obtained may be generalized to other random intersection graph models due to equivalence theorems proved in Section 4 in [2].

While introducing a new random graph model it is worth asking how it differs from those already studied. Therefore one of the first papers on the topic of the random intersection graphs was the one by Scheinerman, Fill and Singer–Cohen [4]. The aim of the paper was to compare $G(n, m, p)$ and random graph $G(n, \hat{p})$ in which each edge appears independently. The value of $\hat{p}$ was set to be approximately $\Pr \{(v_1, v_2) \in E(G(n, m, p))\}$. One of the conclusions of the article was that the differences between $G(n, \hat{p})$ and $G(n, m, p)$ are caused
by the dependency of edge appearance in the latter one and that the larger the \( m \) the less relevant the dependency.

The main theorem in \[4\] states that for \( \alpha > 6 \) the graphs \( G(n, \hat{p}) \) and \( G(n, m, p) \) have asymptotically the same properties. Moreover, it is pointed out that the theorem may be extended to smaller values of \( \alpha \) if we make additional assumptions about \( p \). The proof is based on the fact that for large \( \alpha \) and relevant values of \( p \), with probability tending to one as \( n \to \infty \), there are no properties assigned to more than two vertices and therefore the dependency between edges is asymptotically negligible. The authors of \[4\] suggest that the equivalence theorem is true for all properties for \( 3 \leq \alpha \leq 6 \), i.e. in the case where the number of vertices assigned to each property is still small.

The above mentioned result and conjecture are consistent with the simple observation that the number of vertices to which a given property \( w \) is assigned has essential impact on the dependency between edge appearance in \( G(n, m, p) \). The edge set of the random intersection graph \( G(n, m, p) \) is a union of cliques with the vertex sets \( V(w) := \{v \in V : w \in W(v)\} \), \( w \in W \), and we may divide the set of edges of \( G(n, m, p) \) according to the size of the clique in which it is contained. Let \( k \geq 2 \). We will denote by \( G_k(n, m, p) \) a graph with a vertex set \( V \) and an edge set \( \{(v_1, v_2) : \exists_w v_1, v_2 \in V(w) \text{ and } |V(w)| = k\} \). Alternatively we may define \( G_k(n, m, p) = G(H_k(n, m, p)) \), where \( H_k(n, m, p) \) is a hypergraph with a vertex set \( V \) and an edge set \( \{(v_1, v_2, \ldots, v_k) : \exists_w V(w) = \{v_1, v_2, \ldots, v_k\}\} \) and for a hypergraph \( H \) a graph \( GH \) is a graph with the same vertex set as the hypergraph \( H \) and an edge set consisting of those pairs of vertices which are contained in at least one edge of \( H \). Under this notation \( E(G(n, m, p)) = \bigcup_{k=2}^{m} E(GH_k(n, m, p)) = \bigcup_{k=2}^{m} E(G_k(n, m, p)) \). In \[4\] it is shown that for some \( m \) and \( p \) graphs \( G(n, m, p), G_2(n, m, p), G(n, \hat{p}) \) are asymptotically almost the same. To be precise \( G_k(n, m, p) \) are empty for \( k \geq 3 \) with probability tending to one as \( n \to \infty \) (we will say with high probability) and the edges in \( G_2(n, m, p) \) are almost independent.

The authors in \[4\] support the conjecture for \( 3 \leq \alpha \leq 6 \) by results concerning threshold functions for some properties of \( G(n, m, p) \). However, it should be pointed out that if there exists \( C > 0 \) such that

\[
(1) \quad p \geq C\left(\frac{1}{n^{\sqrt{m}}}\right),
\]

then the expected number of edges in \( G_3(n, m, p) \) tends to a constant or even to infinity. Therefore we may expect that the structure of \( G(n, m, p) \) and \( G(n, \hat{p}) \) differs. Namely, though the number of triangles in \( G_3(n, m, p) \) may make dominating contribution, the impact of triangles contained in \( G_3(n, m, p) \) on the structure of the random intersection graph cannot be omitted. As an example we may state the fact that for \( \alpha = 3 \) the number of triangles in \( G(n, m, p) \) and \( G(n, \hat{p}) \) on the threshold of appearance (i.e. for \( p = c/n^2 \) and \( \hat{p} \sim mp^2 = c^2/n \)) has Poisson distribution with parameters \((c^3 + c^5)/3! \) and \( c^5/3! \), respectively (see \[11\]). For larger values of \( \alpha \) the expected number of triangles in \( G(n, m, p) \) and \( G(n, \hat{p}) \) may also differ significantly. The same is true for cliques of size four contained in \( G_3(n, m, p) \). In fact \( G_k(n, m, p) \) should be rather compared with \( GH_k(n, \hat{p}_k) \), where \( \hat{p}_k \) is approximately the probability that for given \( \{v_1, \ldots, v_k\} \subseteq V \) there exists \( w \) such that \( V(w) = \{v_1, \ldots, v_k\} \), \( H_k(n, \hat{p}_k) \) is a \( k \)-uniform random hypergraph with each edge appearing independently with probability \( \hat{p}_k \) and \( GH_k(n, \hat{p}_k) \) is defined as above. The above observation leads us to the conclusion that the equivalence theorem may not be stated for \( 3 \leq \alpha \leq 6 \) in such a general
form as it was for \( \alpha > 6 \). Therefore we will draw our attention to the case of monotone properties. The concept of restriction of the equivalence theorems to the class of monotone properties has already been developed while examining the equivalence of \( G(n, \hat{p}) \) and \( G(n, M) \) (see [3, 6, 9]).

The article is organized as follows. In Section 2 we state and discuss the results. In Section 4 we outline the proof of the main theorems. In Sections 5 and 6 we prove results needed to complete the proof.

Throughout the article all limits are taken as \( n \to \infty \). We will also use standard Landau notation \( O(\cdot), \Theta(\cdot), \Omega(\cdot), o(\cdot), \sim \) (see for example [6]) and we will use the phrase 'with high probability' to say with probability tending to one as \( n \to \infty \).

### 2 Result

In our considerations we will draw our attention to \( G(n, m, p) \) for

\[
\Omega \left( \frac{1}{n^{3/2} \sqrt{m}} \right) = p = O \left( \sqrt{\frac{\ln n}{m}} \right).
\]

For the values of \( p \) significantly larger than \( \sqrt{\frac{\ln n}{m}} \) the graph \( G(n, m, p) \) is with high probability a complete graph on \( n \) vertices (see [4, 10]). And if

\[
p = o \left( \frac{1}{n^{3/2} m} \right),
\]

then with high probability \( G_k(n, m, p) \) are empty for all \( k \geq 3 \). Therefore a slight modification of the proof from [4] gives us the result that \( G(n, m, p) \) and \( G(n, \hat{p}) \) are asymptotically equivalent for all possible properties.

In fact we may state the following equivalence theorem.

**Theorem 1.** Let \( a \in [0; 1] \), \( \mathcal{A} \) be any graph property, \( p = o \left( \frac{1}{n^{3/2} m} \right) \) and

\[
\hat{p} = 1 - \exp \left( -mp^2 (1-p)^{n-2} \right).
\]

Then

\[
\Pr \{ G(n, \hat{p}) \in \mathcal{A} \} \to a
\]

if and only if

\[
\Pr \{ G(n, m, p) \in \mathcal{A} \} \to a.
\]

For the proof see [4] and Lemma [6].

The result will concern monotone properties. The most important properties such as connectivity, having the largest component of size at least \( k \), containment of the perfect matching or containment of a given graph as a subgraph are included in the wide family of the monotone properties. Let \( \mathcal{G} \) be a family of graphs with a vertex set \( V \). We will call \( \mathcal{A} \subseteq \mathcal{G} \) an increasing (decreasing) property if \( \mathcal{A} \) is closed under isomorphism and \( G \in \mathcal{A} \) implies \( G' \in \mathcal{A} \) for all \( G' \) such that \( E(G) \subseteq E(G') \) (\( E(G') \subseteq E(G) \)).
Theorem 2. Let $a \in [0; 1]$, $m = n^\alpha$ for $\alpha \geq 3$ and $A$ be any monotone property.

(i) Let

$$\Omega \left(\frac{1}{n^{3/2}m}\right) = p = O\left(\frac{\ln n}{m}\right)$$

and $1/n^{3/2}m = o(p)$ for $\alpha = 3$.

If

$$\Pr \{ G(n, 1 - \exp(-mp^2(1 - p)^{(n-2)})) \in A \} \to a$$

and for all $\varepsilon = \varepsilon(n) \to 0$

$$\Pr \{ G(n, (1 + \varepsilon)(1 - \exp(-mp^2(1 - p)^{(n-2)}))) \in A \} \to a,$$

then

$$\Pr \{ G(n, m, p) \in A \} \to a.$$

(ii) Let $\hat{p} = \hat{p}_n \in [0; 1)$ be a sequence bounded away from one by a constant.

If for all $\varepsilon = \varepsilon(n) \to 0$

$$\Pr \left\{ G \left( n, \sqrt{\frac{-\ln(1 - \frac{\hat{p}}{1+\varepsilon})}{m}} \right) \in A \right\} \to a$$

and

$$\Pr \left\{ G \left( n, \sqrt{\frac{-\ln(1 - \hat{p})}{(1 - \varepsilon)m}} \right) \in A \right\} \to a,$$

then

$$\Pr \{ G(n, \hat{p}) \in A \} \to a.$$

In (i) we have to exclude the case $p = \Theta(1/n^{1/\sqrt{m}})$, since the thesis is not true on the threshold of the triangle appearance (see [11]).

The method of the proof is strong enough to show sharper results in many cases. For example, for $\alpha > 3$ the function $\varepsilon(n)$ may be replaced by $1/n^\delta$, where $\delta$ is a constant depending on $\alpha$. We state here two theorems as the example of how tight the results may be, if we make some additional assumptions.

Theorem 3. Let $a \in [0; 1]$, $A$ be any monotone property, $m = n^\alpha$ for $\alpha > 4$ and

$$\Omega \left(\frac{1}{n^{3/2}m}\right) = p = O\left(\frac{\ln n}{m}\right).$$

Let

$$p_- = 1 - \exp(-mp^2(1 - p)^{(n-2)});$$

$$p_+ = 1 - \exp(-mp^2(1 - p)^{(n-2)}) + 10^3\sqrt{mp^3}.$$
If
\[ \Pr\{G(n, p_-) \in A\} \to a \quad \text{and} \quad \Pr\{G(n, p_+) \in A\} \to a, \]
then
\[ \Pr\{G(n, m, p) \in A\} \to a. \]

**Theorem 4.** Let \( a \in [0; 1] \), \( c_3 = 30 \), \( c_4 = 157 \), \( A \) be any monotone property, \( m = n^\alpha \) for \( \alpha > 10/3 \). Let
\[
p_- = 1 - \exp(-mp^2(1-p)^{n-2});
\]
\[
p_+ = \begin{cases} 
1 - \exp(-mp^2(1-p)^{n-2}) + c_3 \sqrt{mp^3}, \\
\text{for } \Omega(n^{-1}m^{-1/3}) = p = o(n^{-1}m^{-1/4}); \\
1 - \exp(-mp^2(1-p)^{n-2}) + c_3 \sqrt{mp^3} + c_4 \sqrt{mp^3}, \\
\text{for } \Omega(n^{-1}m^{-1/4}) = p = O\left(m^{-1/2} \ln^{1/2} n\right). 
\end{cases}
\]
If
\[ \Pr\{G(n, p_-) \in A\} \to a \quad \text{and} \quad \Pr\{G(n, p_+) \in A\} \to a, \]
then
\[ \Pr\{G(n, m, p) \in A\} \to a. \]

### 3 Auxiliary definitions, inequalities and facts

#### 3.1 Coupling

We will frequently use a coupling argument. Let \( \prec \) be a countable partially ordered set. In our case \( \mathbb{P} \) may be a subset of \( \mathbb{N} \) with relation \( \leq \), a Cartesian product \( \mathbb{N}^t \) with relation \( (x_1, \ldots, x_t) \prec (y_1, \ldots, y_t) \Leftrightarrow \forall 1 \leq i \leq t \; x_i \leq y_i \) or a set of hypergraphs \( G \) on a given set of vertices with relation \( \subseteq \) of being a subhypergraph. In the article the set \( G \) will be either the set of all graphs or hypergraphs on \( n \) vertices or the set of \( k \)–partite graphs or hypergraphs with partitions with \( n \) vertices. To omit unnecessary formalities we will not say directly which partially ordered set we are considering, when it is obvious from the context. Let \( X \) and \( Y \) be two random variables with values in \( \mathbb{P} \). We will write \( X \prec_q Y \), if there exists a coupling \((X, Y)\) of the random variables such that \( X \prec Y \) with probability \( q \) (i.e. if there exists a probability space \( \Omega \) and two random variables \( X' \) and \( Y' \), such that \( X' \) and \( Y' \) are both defined on \( \Omega \), have probability distribution as \( X \) and \( Y \), respectively, and \( X' \prec Y' \) with probability \( q \)). We will use the fact that such coupling exists if and only if there exists a probability measure \( \mu : \mathbb{P} \times \mathbb{P} \to [0; 1] \) such that for any set \( A \subseteq \mathbb{P} \) we have \( \mu(A \times \mathbb{P}) = \Pr\{X \in A\} \) and \( \mu(\mathbb{P} \times A) = \Pr\{Y \in A\} \) and \( \mu(\{(x, y) \in \mathbb{P} \times \mathbb{P}; x < y\}) = q \).

Now we will state two useful facts. The simple proofs we add for completeness of considerations.
Fact 1. Let \( \mathbb{P} \) be a countable partially ordered set and \( X \) and \( Y \) be random variables with values in \( \mathbb{P} \). If

\[
X \preceq_{1-q_1} Y \quad \text{and} \quad Y \preceq_{1-q_2} Z,
\]

then for some \( q \leq q_1 + q_2 \)

\[
X \preceq_{1-q} Z.
\]

Proof. Let \( \mu_1, \mu_2 : \mathbb{P} \times \mathbb{P} \to [0; 1] \) be probability measures associated with couplings existing by (4). Let \( \mathbb{P}^* = \{ y \in \mathbb{P} : \Pr \{ Y = y \} \neq 0 \} \). Define

\[
\mu_3 : \mathbb{P} \times \mathbb{P}^* \times \mathbb{P} \to [0; 1], \quad \mu_3(x, y, z) = \frac{\mu_1(x, y)\mu_2(y, z)}{\Pr \{ Y = y \}}; \\
\mu_3(x, y, z) = \mu_3(\{ x \} \times \mathbb{P}^* \times \{ z \}).
\]

Then for \( A_1 = \{ (x, y, z) : x \prec y \} \) and \( A_2 = \{ (x, y, z) : y \prec z \} \) we have

\[
\mu(\{(x, z) : x \prec z\}) = \mu_3(\{(x, y, z) : x \prec z\}) \geq \mu_3(A_1 \cap A_2) \geq \mu_3(A_1) + \mu_3(A_2) - 1 = 1 - (q_1 + q_2).
\]

Fact 2. If \( (X_1, \ldots, X_t) \) and \( (Y_1, \ldots, Y_t) \) are vectors of independent random variables and

\[
X_i \preceq_{q_i} Y_i, \quad \text{for all } 1 \leq i \leq t,
\]

then

\[
(X_1, \ldots, X_t) \preceq_{q} (Y_1, \ldots, Y_t)
\]

and

\[
\sum_{i=1}^{t} X_i \preceq_{q'} \sum_{i=1}^{t} Y_i,
\]

where \( q, q' \geq \prod_{i=1}^{k} q_i \).

Proof. For all \( 1 \leq i \leq t \), let \( \mu_i : \mathbb{P} \times \mathbb{P} \to [0; 1] \) be a probability measure associated with the coupling existing by \( X_i \preceq_{q_i} Y_i \). Simple calculation shows that \( \mu : \mathbb{P}^t \times \mathbb{P}^t \to [0; 1] \) such that

\[
\mu(x_1, \ldots, x_t, y_1, \ldots, y_t) = \prod_{i=1}^{t} \mu_i(x_i, y_i)
\]

implies the thesis. 

\( \square \)
3.2 Total variation distance

Let $X$ and $Y$ be the random variables with values in the same countable set $\mathbb{P}$. We define the total variation distance between $X$ and $Y$ by

$$d_{TV}(X, Y) = \max_{A \subseteq \mathbb{P}} |\Pr\{X \in A\} - \Pr\{Y \in A\}| = \frac{1}{2} \sum_{x \in \mathbb{P}} |\Pr\{X = x\} - \Pr\{Y = x\}|.$$  

Now let $\mathbb{P} = \mathcal{G}$ be the set of hypergraphs (graphs) with a given vertex set. Let $G_1$ and $G_2$ be two random variables with values in $\mathcal{G}$. Since $d_{TV}(G_1, G_2) = 1 - 2d_{TV}(G_1, G_2)$, it is simple to construct the probability measure $\mu$ on $\mathcal{G} \times \mathcal{G}$ such that $\mu\{(G, G) : G \in \mathcal{G}\} = 1 - 2d_{TV}(G_1, G_2)$ with marginal distributions as distributions of $G_1$ and $G_2$. This implies:

**Fact 3.**

$$G_1 \preceq_q G_2 \quad \text{and} \quad G_2 \preceq_{q'} G_1,$$

where $q, q' \geq 1 - 2d_{TV}(G_1, G_2)$.

In our considerations we will also use the following facts concerning total variation distance, which are Facts 3 and 4 in [4].

**Fact 4.** Let $A$ and $A'$ be random variables with values in the same set. If there exist random variables $B$ and $B'$ such that for all possible $b$ the distribution of $A$ under condition $B = b$ and the distribution of $A'$ under condition $B' = b$ are the same, then

$$d_{TV}(A, A') \leq 2d_{TV}(B, B').$$

**Fact 5.** Let $A$ and $A'$ be two random variables. If there exists a probability space on which random variables $B$ and $B'$ are both defined and have probability distribution as $A$ and $A'$, respectively, then

$$d_{TV}(A, A') \leq \Pr\{B \neq B'\}.$$

We will also use a standard result (see for example [1] equation (1.23)).

**Fact 6.** Let $A$ be a random variable with binomial distribution $\text{Bin}(\hat{n}, \hat{p})$ and let $A'$ be a random variable with Poisson distribution $\text{Po}(\hat{n}\hat{p})$. Then

$$d_{TV}(A, A') \leq \hat{p}.$$

3.3 Coupon collector model

We will define two auxiliary random variables, which are generalized versions of random variables defined in [4]. Let $K \geq 2$ be a given constant integer, $M$ be any random variable with values in $\mathbb{N}$, $\overline{\pi} = (n_2, \ldots, n_K)$ be a vector of positive integers and $\overline{P} = (P_2, \ldots, P_K)$ be the vector of nonnegative reals such that $\sum_{k=2}^{K} n_k P_k \leq 1$. Assume now that we have
\[ \sum_{k=2}^{K} n_k \text{ coupons } \bigcup_{k=2}^{K} \{c_{1}^{(k)}, \ldots, c_{n_k}^{(k)} \} \text{ and one blank coupon } d_0. \] We make \( M \) independent draws, with replacement, such that in each draw

\[ \Pr\{c_i^{(k)} \text{ is chosen}\} = P_k, \quad \text{for } 2 \leq k \leq K, 1 \leq i \leq n_k; \]
\[ \Pr\{d_0 \text{ is chosen}\} = 1 - \sum_{k=2}^{K} n_k P_k. \]

In this scheme we define \( R_i^{(k)}(M) \) to be a random variable denoting the number of times that a coupon \( c_i^{(k)} \) was chosen and

\[ X_i^{(k)}(M) = \begin{cases} 1 & \text{if } R_i^{(k)}(M) \geq 1; \\ 0 & \text{otherwise.} \end{cases} \]

The first auxiliary random variable is

\[ X(M) = X(\pi, \overline{P}, M) = (X^{(2)}(M), \ldots, X^{(K)}(M)), \]
where

\[ X^{(k)}(M) = \sum_{i=1}^{n_k} X_i^{(k)}(M). \]

The second random variable is

\[ Y = Y(\pi, \overline{P'}) = (Y^{(2)}, \ldots, Y^{(K)}) \]
where \( \overline{P'} = (P'_2, \ldots, P'_K) \) is a vector such that \( P'_k \leq 1 \) for all \( 2 \leq k \leq K \) and \( Y^{(k)}, 2 \leq k \leq K, \) are independent random variables with binomial distribution \( \text{Bin}(n_k, 1 - \exp(-\lambda P_k)) \).

The simple observation stated below is a generalization of the part of the proof of Claim 1 in [4] and may be shown by careful calculation.

**Fact 7.** Let \( M \) be a random variable with Poisson distribution \( \text{Po}(\lambda) \), then \( R_i^{(k)}(M), 2 \leq k \leq K \text{ and } 1 \leq i \leq n_k, \) are independent random variables with Poisson distribution \( \text{Po}(\lambda P_k) \). Moreover \( X^{(k)}(M), 2 \leq k \leq K, \) are independent random variables with binomial distribution \( \text{Bin}(n_k, 1 - \exp(-\lambda P_k)) \). Therefore \( X(M) \) and \( Y \) have the same distribution for \( P'_k = 1 - \exp(-\lambda P_k) \).

It is also simple to show the following fact.

**Fact 8.** Let \( M \) and \( M' \) be random variables with values in \( \mathbb{N} \). If

\[ M \preceq_{1-o(1)} M', \]
then

\[ X(M) \preceq_{1-o(1)} X(M'). \]
3.4 Chernoff bound

We will use the Chernoff bound (see Theorem 2.1 in [6]).

**Lemma 1.** Let $X$ be a random variable with binomial distribution and $\lambda = \mathbb{E}X$. Let $a \geq \lambda$, then

$$\Pr \{ X \geq a \} \leq \exp \left( -\frac{\lambda - a \ln a}{\lambda} + a \right)$$

After careful calculation we obtain the following lemma.

**Lemma 2.** Let $t \geq 1$ be an integer and $X_n$ be a sequence of random variables with binomial distribution, such that $\mathbb{E}X_n = \lambda_n$. Let $\varepsilon > 0$ and $\omega(n)$ be any function tending to infinity. If

$$a_n = a_n(\lambda_n, t, \varepsilon) = \begin{cases} (t + \varepsilon) \ln n / (\ln n - \ln \lambda_n), & \text{for } \lambda_n = o(\ln n); \\ \omega(n) \lambda_n, & \text{for } \lambda_n = \Theta(\ln n); \\ (1 + \varepsilon) \lambda_n, & \text{for } \ln n = o(\lambda_n), \end{cases}$$

then

$$\Pr \{ X_n \geq a_n \} \leq o(n^{-t}).$$

**Lemma 3.** Let $X_n$ be a sequence of random variables with binomial distribution. Then

$$\begin{align*}
\Pr \{ X_n \leq \mathbb{E}X_n - t_n \} &\leq \exp \left( -\frac{t_n^2}{2\mathbb{E}X_n} \right), & \text{for } t_n \geq 0; \\
\Pr \{ X_n \geq \mathbb{E}X_n + t_n \} &\leq \exp \left( -\frac{3t_n^2}{2(3\mathbb{E}X_n + t_n)} \right), & \text{for } t_n \geq 0.
\end{align*}$$

It is also possible to formulate the version of the Chernoff bound for random variables with Poisson distribution.

**Lemma 4.** Let $X_n$ be a sequence of random variables with Poisson distribution $\text{Po}(\lambda)$ and $i > 0$ be any constant, then

$$\begin{align*}
\Pr \{ X_n \leq \mathbb{E}X_n - t_n \} &\leq \exp \left( -\frac{t_n^2}{2\mathbb{E}X_n} \right) + o \left( \frac{1}{n^i} \right), & \text{for } t_n \geq 0; \\
\Pr \{ X_n \geq \mathbb{E}X_n + t_n \} &\leq \exp \left( -\frac{3t_n^2}{2(3\mathbb{E}X_n + t_n)} \right) + o \left( \frac{1}{n^i} \right), & \text{for } t_n \geq 0.
\end{align*}$$

**Proof.** It follows by (9) applied to random variable with binomial distribution $\text{Bin}(\lambda n^{i+1}, 1/n^{i+1})$, definition of the total variation distance and Fact 6.

4 Outline of the proof of theorems

Let $p_-$ and $p_+$ be such that

$$G(n, p_-) \preceq_{1-o(1)} G(n, m, p) \quad \text{and} \quad G(n, m, p) \preceq_{1-o(1)} G(n, p_+).$$
Thus we may assume that there exists a probability space on which $G(n, p_{-})$ and $G(n, m, p)$ ($G(n, m, p)$ and $G(n, p_{+})$) are defined and

$$\Pr\{E_{-}\} = 1 - o(1) \quad (\Pr\{E_{+}\} = 1 - o(1)),$$

for

$$E_{-} := \{G(n, p_{-}) \subseteq G(n, m, p)\} \quad (E_{+} := \{G(n, m, p) \subseteq G(n, p_{+})\}).$$

Let $a \in [0, 1]$ and $A$ be an increasing property. If

$$\Pr\{G(n, p_{-}) \in A\} \to a \quad \text{and} \quad \Pr\{G(n, p_{+}) \in A\} \to a,$$

then under the couplings implied by (11)

$$\Pr\{G(n, m, p) \in A\} \leq \Pr\{G(n, m, p) \in A | E_{+}\} \Pr\{E_{+}\} + \Pr\{E_{+}^{c}\} \leq$$

$$\leq \Pr\{G(n, p_{+}) \in A | E_{+}\} \Pr\{E_{+}\} + \Pr\{E_{+}^{c}\} \leq$$

$$\leq \Pr\{\{G(n, p_{+}) \in A\} \cap E_{+}\} + \Pr\{E_{+}^{c}\} =$$

$$= \Pr\{G(n, p_{+}) \in A\} + o(1) = a + o(1)$$

and

$$\Pr\{G(n, m, p) \in A\} \geq \Pr\{G(n, m, p) \in A | E_{-}\} \Pr\{E_{-}\} \geq$$

$$\geq \Pr\{G(n, p_{-}) \in A | E_{-}\} \Pr\{E_{-}\} =$$

$$= \Pr\{\{G(n, p_{-}) \in A\} \cap E_{-}\} =$$

$$\geq \Pr\{G(n, p_{-}) \in A\} + \Pr\{E_{-}\} - \Pr\{\{G(n, p_{-}) \in A\} \cup E_{-}\} \geq$$

$$\geq \Pr\{G(n, p_{-}) \in A\} + \Pr\{E_{-}\} - 1 =$$

$$= \Pr\{G(n, p_{-}) \in A\} + o(1) = a + o(1).$$

Analogous equalities may be formulated for the decreasing properties or if we replace (11) by

$$G(n, m, p_{-}) \preceq_{1-o(1)} G(n, \hat{p}_{+}) \preceq_{1-o(1)} G(n, m, p_{+}).$$

Therefore the main aim of the proof is to show that under the assumptions of the theorems couplings (11) and (12) exist. The existence of the couplings is implied by the following lemma.

**Lemma 5.** Let

$$a_{n}(q) = \begin{cases} 
6 & \text{for } nq^{2} = O(n^{-1/2}); \\
3 \ln n/(- \ln nq^{2} + \ln n) & \text{for } nq^{2} = o(1) \text{ and } o(nq^{2}) = n^{-1/2}; \\
3 \ln n/ \ln \ln n & \text{for } nq^{2} = \Theta(1); \\
3 \ln n/(\ln \ln n - 3 \ln nq^{2}) & \text{for } nq^{2} \to \infty \text{ and } nq^{2} = o(\sqrt[\ln n]{n}); \\
\omega(n)n^{3}q^{6}, \text{ where } \omega(n) \to \infty & \text{for } nq^{2} \to \infty \text{ and } nq^{2} = \Theta(\sqrt[\ln n]{n}) \\
cn^{3}q^{6}, \text{ where } c > 1 & \text{for } nq^{2} \to \infty \text{ and } o(nq^{2}) = \frac{3}{\sqrt[\ln n]{n}};
\end{cases}$$

10
Moreover let \( c_3 > 2 \cdot 3/\sqrt{31} \), \( c_4 > \sqrt{15} \cdot 3 \cdot 4/\sqrt{41} \), \( c_5 > \sqrt{2^2 \cdot 3 \cdot 5^3 \cdot 4 \cdot 5}/ \sqrt{51} \), \( q_k = (1 - \exp(-mp^k(1 - p)^{n-k}))^{1/(2k)} \), for \( k = 2, 3, 4, 5 \) and

\[
p_+ = \begin{cases} 
q_2 + a_n(c_3 q_3) c_3 q_3, & \text{for } p = \Omega(n^{-1} m^{-1/3}) \text{ and } p = o(\min\{n^{-1} m^{-1/4}, n^{-3/7} m^{-1/3}\}); \\
q_2 + \sum_{k=3}^4 a_n(c_k q_k) c_k q_k, & \text{for } p = \Omega(n^{-1} m^{-1/4}) \text{ and } p = o(\min\{n^{-1} m^{-1/5}, n^{-3/7} m^{-1/3}, n^{-9/14} m^{-1/4}\}); \\
q_2 + \sum_{k=3}^5 a_n(c_k q_k) c_k q_k, & \text{for } p = \Omega(n^{-1} m^{-1/6}) \text{ and } p = o(\min\{n^{-1} m^{-1/6}, n^{-3/7} m^{-1/3}, n^{-9/14} m^{-1/4}, n^{-6/7} m^{-1/5}\}). 
\end{cases}
\]

Then

\[
G(n, p_2) \preceq_{1-o(1)} G(n, m, p) \quad \text{and} \quad G(n, p_2) \preceq_{1-o(1)} G(n, p_2).
\]

Now we will prove Theorems 2, 3 and 4 using Lemma 5.

**Proof of Theorems 2, 3 and 4.** It is easy to check that, under the assumptions of Theorem 2(i) about \( p \), we have \( p_+ = (1 + o(1)) q_2 \). Therefore, for some \( \epsilon' = \epsilon'(n) \to 0 \) we have \( p_+ = (1 + \epsilon')(1 - \exp(-mp^2(1 - p)^{m-2})) \). Thus by Lemma 5

\[
G(n, q_2) \preceq_{1-o(1)} G(n, m, p) \preceq_{1-o(1)} G(n, (1 + \epsilon') q_2),
\]

which implies Theorem 2(i).

Analogously in the proof of Theorem 2(ii) we use the fact that, either \( \hat{p} \) is such that we may use Theorem 1 or for some \( \epsilon'(n) \to 0 \) we get

\[
G \left( n, m, \sqrt{-\frac{\ln(1 - \frac{\hat{p}}{1 + \epsilon'})}{m}} \right) \preceq_{1-o(1)} G(n, \hat{p})
\]

and

\[
G(n, \hat{p}) \preceq_{1-o(1)} G \left( n, m, \sqrt{-\frac{\ln(1 - \hat{p})}{(1 - \epsilon') m}} \right).
\]

This implies Theorem 2(ii).

The proofs of Theorems 4 and 5 are basically the same. We only have to calculate the value of \( a(c_k q_k) c_k \) and notice that \( q_k \sim \frac{1}{2} \sqrt{m p^k} \) for \( k = 3, 4 \).

Therefore we need to show Lemma 5 to complete the proof.

**Outline of the proof of Lemma 5.** In the statement of the lemma we have 3 different values of \( p_+ \). They correspond to three cases. In the first case, since \( p = o(n^{-1} m^{-1/4}) \), the hypergraph \( \mathcal{H}_4(n, m, p) \) is empty with high probability. In the second case \( \mathcal{H}_5(n, m, p) \) is empty and in the third case – \( \mathcal{H}_6(n, m, p) \). We will prove Lemma 5 in all three cases at the same time.
The proof will differ only by the value of $K$, which is 3, 4 and 5 in the first, second and third case, respectively.

Let $m = n^a$ and $q_k = (1 - \exp(-mp^k(1-p)^{n-k}))^{1/(n-k)}$. We will prove that under assumptions of Lemma 5 there exists a sequence of couplings

\begin{align}
&\text{(14)} \quad G(n, q_2) \preceq_{1-o(1)} 1 \\
&\text{(15)} \quad G_2(n, m, p) \preceq_{1-o(1)} 1 - o(1) \bigcup_{k=2}^K G_k(n, m, p) \\
&\text{(16)} \quad \preceq_{1-o(1)} 1 - o(1) \bigcup_{k=2}^K H_k\left(n, q_k^{(k)}\right) \\
&\text{(17)} \quad \preceq_{1-o(1)} 1 - o(1) \bigcup_{k=2}^K H_k\left(n, \sum_{k=3}^K a_n(c_k q_k) c_k q_k\right) \\
&\text{(18)} \quad \preceq_{1} 1 - o(1) \bigcup_{k=3}^K a_n(c_k q_k) c_k q_k.
\end{align}

Here

\[G_{H_2}\left(n, q_2^{(2)}\right), \ldots, G_{H_K}\left(n, q_K^{(K)}\right)\]

are independent random hypergraphs, $c_3, c_4, c_5$ are constants given in the statement of the lemma and $a_n(\cdot)$ is a function given in the statement of the lemma.

The left-hand side coupling of (15) is trivial and the coupling (18) is a standard one. The right-hand side coupling in (15) follows by the fact that under the assumptions of Lemma 5

\[\Pr\{\exists w \in W \mid |V(v)| > K\} = O\left(mn^{K+1}p^{K+1}\right) = o(1).\]

Proof of the existence of the remaining couplings will require more effort. The couplings (14) and (16) follow by Lemma 6 and Fact 3.

**Lemma 6.** Let $K \geq 2$ be a constant integer and $p = o(1/n)$, then

\[d_{TV}\left(\bigcup_{k=2}^K H_k\left(n, m, p\right), \bigcup_{k=2}^K H_k\left(n, 1 - \exp(-mp^k(1-p)^{n-k})\right)\right) = o(1),\]

where $H_k\left(n, 1 - \exp(-mp^k(1-p)^{n-k})\right)$ are independent random hypergraphs.

**Lemma 7.** Let $c_3 > 2 \cdot 3/\sqrt{3}$, $q = \Omega(n^{-1})$ and $q = o(n^{-3/7})$.

\[GH_3\left(n, q^3\right) \preceq_{1-o(1)} 1 - o(1) \bigcup a_n(c_3 q) c_3 q,\]

where $a_n(q)$ is defined as in (13).
Lemma 8. Let $c_4 > \sqrt[5]{15} \cdot 3 \cdot 4 / \sqrt[3]{4!}$, $q = \Omega(n^{-2/3})$ and $q = o(n^{-3/7})$.

$$GH_4 (n, q^6) \leq_{1-o(1)} G (n, a_n(c_4q)c_4q) ,$$

where $a_n(q)$ is defined as in (13).

Lemma 9. Let $c_5 > \sqrt[5]{2^2 \cdot 3 \cdot 5^3 \cdot 4 \cdot 5 / \sqrt[5]{5!}}$, $q = \Omega(n^{-1/2})$ and $q = o(n^{-3/7})$.

$$GH_5 (n, q^{10}) \leq_{1-o(1)} G (n, a_n(c_5q)c_5q) ,$$

where $a_n(q)$ is defined as in (13).

The above lemmas will be shown in Section 6. The coupling (17) is a consequence of the above lemmas after substituting $q = q_k$ for $k = 3, \ldots, K$.

Notice that although the expected number of hyperedges in $GH_k (n, q_k(2))$ and cliques in $G (n, q)$ is the same, the function $a_n$ is necessary. We will construct a coupling of two random graph models, the existence of which contradicts the thesis that there exists a constant $C$ such that for all $q$

$$GH_3 (n, q^3) \sim_{1-o(1)} G (n, Cq) .$$

Let $q = o(1)$. For any $e$, a 3-element subset of $V$, define $F_e$ to be the set of bijections assigning to the numbers from the set $\{1, 2, 3\}$ the vertices of $e$ ($|F_e| = 6$). Now, to each $e$, a 3-element subset of $V$, and each function $f \in F_e$ we assign $f$ to $e$ independently of all other functions and sets with probability

$$r = 1 - (1- q^3)^{1/6} \sim \frac{q^3}{6} .$$

Notice that if we add each edge $e$ to the set of edges of the hypergraph with the vertex set $V$ in the case when at least one function from $F_e$ is assigned to $e$, we will get a random variable with the same distribution as $H_3 (n, q^3)$. Moreover we may construct a random subgraph $G_3$ of $GH_3 (n, q^3)$ by adding an edge $(v_1, v_2)$, $v_1, v_2 \in V$, if and only if at least one 3-element subset of $V$ containing $v_1$ and $v_2$ is assigned a function in which $v_1$ and $v_2$ are assigned 1 and 2 or 2 and 1. Notice that, from independent choice of the functions from $F_e$ we get that each edge appears in $G_3$ independently with probability

$$r' = 1 - (1-r)^{2(n-2)} \sim 2nr \sim \frac{1}{3} nq^3 .$$

Therefore

$$G (n, r') \leq_{1} GH_3 (n, q^3)$$

and in the lemmas there should be $a_n = \Omega(nq^2)$. 

13
5 Proof of Lemma 6

Let, for all $2 \le k \le K$,

$$p_k = p^k(1-p)^{n-k}, \quad n_k = \begin{pmatrix} n \end{pmatrix}_k, \quad P_k = \frac{p_k}{\sum_{k=2}^{K} p_k n_k} \quad \text{and} \quad P'_k = 1 - \exp(-mp_k),$$

$M$ have binomial distribution Bin$(m, P)$ ($P = \sum_{k=2}^{K} p_k n_k$), $X(M)$ be defined as in (6) and $Y$ be defined as in (7). Then for all $2 \le k \le K$

$$|E(\mathcal{H}_k(n, m, p))| = X^{(k)}(M) \quad \text{and} \quad |E(H_k(n, 1 - \exp(-mp_k))| = Y^{(k)}.$$

Moreover for any two hypergraphs $H$ and $H'$, such that for all $k \ge 2$ the number of edges of cardinality $k$ in $H$ and $H'$ is the same, we have

$$\Pr \left\{ \bigcup_{k=2}^{K} \mathcal{H}_k(n, m, p) = H \right\} = \Pr \left\{ \bigcup_{k=2}^{K} \mathcal{H}_k(n, m, p) = H' \right\}$$

and

$$\Pr \left\{ \bigcup_{k=2}^{K} H_k(n, 1 - \exp(-mp_k)) = H \right\} = \Pr \left\{ \bigcup_{k=2}^{K} H_k(n, 1 - \exp(-mp_k)) = H' \right\}. $$

Therefore, by Fact 4 the following lemma implies Lemma 6

**Lemma 10.** Let $K \ge 2$ be a constant integer. Let $p = o(1/n)$, $M$ be a random variable with binomial distribution $\text{Bin}(m, \sum_{k=2}^{K} \binom{n}{k} p_k), \ p_k = p^k(1-p)^{n-k}$, $X(M)$ be defined as in (6) for $n_k = \binom{n}{k}$ and $P_k = p_k/(\sum_{k=2}^{K} p_k n_k)$. Let moreover $Y$ be defined as in (7) for $P'_k = 1 - \exp(-mp_k)$. Then

$$d_{TV}(X(M), Y) = o(1).$$

The proof of the above lemma uses the main idea of the proof of Claim 1 in [4]. However, a slight modification of the choice of $M$ and $M'$ enables us to extend the result for $\alpha \le 4$.

**Proof.** We replace the binomial random variable $M$ with the Poisson random variable $M'$ with the same expected value $m \sum_{k=2}^{K} \binom{n}{k} p_k$. By Fact 7 the random variables $X_i^{(j)}(M')$ are independent and

$$\Pr \{X_i^{(k)}(M') = 1\} = 1 - \exp(-mp_k).$$

Therefore in $X(M') = (X^{(2)}(M'), \ldots, X^{(K)}(M'))$, the random variables $X^{(2)}(M'), \ldots, X^{(K)}(M')$ are independent with binomial distribution Bin$(n_2, 1-\exp(-mp_2)), \ldots, \text{Bin}(n_K, 1-\exp(-mp_K))$, respectively. By definition of $Y$, Facts 4 and 6 we have

$$d_{TV}(Y, X(M)) = d_{TV}(X(M'), X(M)) \le 2d_{TV}(M', M) \le \sum_{k=2}^{K} \binom{n}{k} p_k = O \left( \sum_{k=2}^{K} n^k p^k \right) = o(1).$$

\[\square\]
6 Proof of Lemma 7, 8 and 9

In fact, we will show the version of the lemmas in the case of random $k$–partite hypergraphs and graphs. The following fact shows that we may reduce the problem to the $k$–partite case.

First let us introduce additional notation. Let $\mathcal{X}_1, \ldots, \mathcal{X}_k$ be disjoint $n$–element sets and $r \in [0, 1]$. We define $H^{(k)}(n, r)$ to be the hypergraph with a vertex set $\bigcup_{i=1}^{k} \mathcal{X}_i$ and an edge set being the random subset of $\mathcal{E} := \{(x_1, \ldots, x_k) : \forall 1 \leq i \leq k x_i \in \mathcal{X}_i\}$ such that each element from $\mathcal{E}$ is added to $E(H^{(k)}(n, r))$ independently with probability $r$. Let moreover $G^{(k)}(n, r)$ be the random $k$–partite graph with $k$–partition $(\mathcal{X}_1, \ldots, \mathcal{X}_k)$ and each edge appearing with probability $r$. It is simple to prove the following fact.

**Fact 9.** Let $a_n = \Omega(1)$. If

\[
GH^{(k)}(n, r^{(k)}_2) \leq_{1-o(1)} G^{(k)}(n, a_n r),
\]

then,

\[
GH_k\left(n, 1 - (1 - r^{(k)}_2)k\right) \leq_{1-o(1)} G(n, 1 - (1 - a_n r)^k(k-1))
\]

and if $a_n r = o(1)$, then for any constant $c > k - 1$

\[
GH_k\left(n, k!r^{(k)}_2\right) \leq_{1-o(1)} G(n, c a_n r).
\]

**Proof.** Let $\mathcal{X}_i = \{x_{i}^{(1)}, \ldots, x_{i}^{(n)}\}$, for $1 \leq i \leq k$, and $\mathcal{V} = \{v_1, \ldots, v_n\}$. For a given instance of $H^{(k)}(n, r^{(k)}_2)$ (or $G^{(k)}(n, a_n r)$) we may get an instance of a hypergraph $H_k\left(n, 1 - (1 - r^{(k)}_2)\right)$ (or a graph $G\left(n, 1 - (1 - a_n r)^k(k-1)\right)$) with the vertex set $\mathcal{V}$ by merging for all $1 \leq j \leq n$ all the vertices $x_{j}^{(i)}$, $1 \leq i \leq k$, into $v_j$ and deleting the edges with less then $k$ (or 2) vertices. \(\Box\)

In the remaining part of the section we will prove the following lemmas.

**Lemma 11.** Let $q = \Omega(n^{-1})$ and $q = o(n^{-1/3})$.

\[
GH^{(3)}(n, q^3) \leq_{1-o(1)} G^{(3)}(n, a_n(q)q),
\]

where $a_n(q)$ is defined as in (13).

The above lemma is a generalization of Theorem 1.7 from [8], where it was shown in the case $(\ln n/n^2)^{1/3} \ll q = o(n^{-3/5})$ and $a_n = 17$. Moreover we will prove the analogue of the lemma for the case of 4-partite and 5-partite hypergraphs.

**Lemma 12.** Let $q = \Omega(n^{-2/3})$, $c'_4 > \sqrt[3]{15}$ and $q = o(n^{-2/5})$.

\[
GH^{(4)}(n, q^6) \leq_{1-o(1)} G^{(4)}(n, a_n(c'_4 q)),
\]

where $a_n(q)$ is defined as in (13).

**Lemma 13.** Let $q = \Omega(n^{-1/2})$, $c'_5 > \sqrt[3]{2^3 \cdot 5^3}$ and $q = o(n^{-2/5})$.

\[
GH^{(5)}(n, q^{10}) \leq_{1-o(1)} G^{(5)}(n, a_n(c'_5 q)),
\]

where $a_n(q)$ is defined as in (13).
We will prove in detail Lemma 11. The proof of the remaining lemmas is analogous, therefore we will only sketch them.

Proof of Lemma 11 First we will divide the hypergraph $H^{(3)}(n, q^3)$ into $n$, independent, edge disjoint hypergraphs $H(x)$, $x \in X_k$, such that

$$H^{(3)}(n, q^3) = \bigcup_{x \in X_3} H(x).$$

For any $x \in X_3$, $H(x)$ will be a hypergraph with the vertex set $\{x\} \cup X_1 \cup X_2$ and an edge set consisting of those edges from $E(H^{(3)}(n, q^3))$, which contain $x$.

Next, for each $x$ separately, we will compare $GH(x)$ with an auxiliary graph $T(x)$. By definition, $T(x)$ will be a graph with the same vertex set as $H(x)$ and with an edge set constructed by the following procedure. First we add each edge $(x, y)$, $y \in X_1 \cup X_2$ independently with probability $Cq$, to the edge set, where

$$C = C(q) = \begin{cases} c, & \text{where } c > 5, \\ \omega(n), & \text{where } \omega(n) \to \infty, \\ cnq^2, & \text{where } c > 1, \\ q, & \text{where } nq^2 = o(1); \\ \omega(n), & \text{where } \omega(n) \to \infty, \\ cnq^2, & \text{where } c > 1, \\ q, & \text{where } nq^2 = \Theta(n); \end{cases}$$

(We assume, that $\omega(n)$ tends slowly to infinity and $c$ is close to 5 and 1, respectively)

Then independently with probability $q$ we add to the edge set each edge $(x_1, x_2) \in X_1^* \times X_2^*$, where, for each $1 \leq i \leq 2$, $X_i^*$ is the set of vertices form $X_i$ connected by an edge with $x$.

We will prove that

$$(21) \quad GH(x) \approx_{1-o(1/n)} T(x).$$

By definition of $T(x)$, (21) is equivalent to

$$(22) \quad H^*(x) \approx_{1-o(1/n)} T^*(x),$$

where $H^*(x)$ is a graph with the vertex set $X_1 \cup X_2$ and an edge set $\{(x_1, x_2) : (x_1, x_2, x) \in H(x)\}$ and $T^*(x)$ is a subgraph of $T(x)$ induced on $X_1 \cup X_2$.

Then we will show, that

$$(23) \quad \bigcup_{x \in X_3} T(x) \approx_{1-o(1)} G^{(3)}(n, a_n(q)q).$$

To get (23) we will show that

$$(24) \quad \bigcup_{x \in X_3} T^*(x) \approx_{1-o(1)} H^{(2)}(n, a_n(q)q)$$

and $H^{(2)}(n, a_n(q)q)$ is independent of the choice of $X_i^*$. This by definition of $a_n(q)$, $T(x)$ and $T^*(x)$ is equivalent to (23).

Therefore the main parts of the proof will be to show (22) and (24).

Proof of (22)

We will divide the proof into four cases
1. \( q = O(\ln n/n) \),
2. \( \ln n/n \ll q \) and \( q = O((n^2 \ln n)^{1/3}) \),
3. \( (n^2 \ln n)^{1/3} \ll q \) and \( q = o(n^{-1/2}) \),
4. \( q = \Omega(n^{-1/2}) \) and \( q = o(n^{-1/3}) \),

**CASE 1**

In this case we will use the fact that with probability \( 1 + o(1/n) \) the graph \( H^*(x) \) consists of at most one edge. Namely the probability that the graph \( H^*(x) \) has more than one edge is at most

\[
\binom{n^2}{2} q^6 = O\left(n^4 q^6\right) = o\left(\frac{1}{n}\right)
\]

Moreover, for large \( n \),

\[
\Pr \{ \exists x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2 (x_1, x_2) \in E(T^*(x)) \} \geq \sum_{x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2} \Pr \{ (x_1, x_2) \in E(T^*(x)) \} - \sum_{x_1, x'_1 \in \mathcal{X}_1, x_2, x'_2 \in \mathcal{X}_2} \Pr \{ (x_1, x_2), (x'_1, x'_2) \in E(T^*(x)) \}
\]

\[
= n^2 (Cq)^2 q - \binom{n}{2} (Cq)^3 q^2 - 2n \binom{n}{2} (Cq)^3 q^2 = Cn^2 q^3 (1 - O(n^2 q^3 + nq^2)) \geq n^2 q^3 \geq \Pr \{ \exists x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2 (x_1, x_2) \in E(H^*(x)) \}.
\]

This gives an obvious coupling

\( H^*(x) \preceq_{1-o(1/n)} T^*(x) \).

**CASES 2, 3 and 4**

In the three latter cases we will use the fact that the number of the vertices in \( \mathcal{X}_i^* \) is sharply concentrated around its expected value. Let \( r \in [0, 1] \), we denote by \( H^{(2)}_r (C' nq, r) \) a 2–partite graph with 2–partition \((\mathcal{X}_1, \mathcal{X}_2)\) and with an edge set constructed by the following procedure. Let

\[
C' = \begin{cases} 
5, & \text{for } nq^2 = o(1); \\
\omega'(n), & \text{for } nq^2 = \Theta(1); \\
c'nq^2, & \text{where } c' > 1 \text{ for } nq^2 \to \infty,
\end{cases}
\]

be such that \( C/C' > 1 \) is a constant. First, for each \( 1 \leq i \leq 2 \) independently, we choose \( \mathcal{X}_i' \) uniformly at random from all \( C' nq \) element subsets of \( \mathcal{X}_i \) and then we add each edge \( (x_1, x_2) \in \mathcal{X}_1' \times \mathcal{X}_2' \) to the edge set independently with probability \( r \). By Chernoff bound \( H^{(2)}_r (C' nq, q) \preceq_{1-o(1/n)} T^*(x) \).
Therefore, since in $H^*(x)$ each edge appears independently, in order to show \((22)\) we will show that in all three cases

$$H^{(2)}(n, q^3) \preceq_{1-o(1/n)} H^{(2)}_s(C'nq, q^3).$$

**CASE 2**

In this case we will use the fact that $H^{(2)}(n, q^3)$ (i.e. also $H^*(x)$) with high probability does not contain many edges except the maximum matching. First we will show auxiliary lemmas.

**Lemma 14.** Let $r = o(1/(C'nq))$, $ln n = o(nq)$ and $N_2(r)$ be the random variable denoting the size of the maximum matching in $H^{(2)}_s(C'nq, r)$, then

\[(25)\]

$$N'_2(r) \preceq_{1-o(1/n)} N_2(r),$$

where $N'_2(r)$ has binomial distribution $Bin(C'nq, s_2(r))$ and $s_2(r) \sim C'nqr$.

**Proof of Lemma 14.** Let $H$ be a hypergraph chosen according to the probability distribution of $H^{(2)}_s(C'nq, r)$. Define $H'$ to be the hypergraph with a vertex set $X_1$ and an edge set $\{(x_1): x_1 \in X_1 \text{ and } \exists_{x_2 \in X_2}(x_1, x_2) \in E(H)\}$. Notice that $H'$ is chosen according to the probability distribution of $H^{(1)}_s(C'nq, 1 - (1 - r)^{C'nq})$ ($H^{(1)}_s(C'nq, 1 - (1 - r)^{C'nq})$ is a hypergraph with a vertex set $X_1'$ and the edge set constructed by first choosing $X_1'$ uniformly at random from all $C'nq$-element subsets of $X_1$ and then adding to the edge set each $x_1 \in X_1'$ independently with probability $1 - (1 - r)^{C'nq}$. Let $H''$ be the subhypergraph of $H$ such that for each edge $(x_1) \in E(H')$ we pick uniformly at random an edge from $E(H)$ containing $x_1$ and add it to the edge set of $H''$. Notice that the maximum matching in $H$ is at least of the size of the set of non isolated vertices in $X_2$ in $H''$. Moreover the edge set of $H''$ may be alternatively constructed in the following way (i.e. this construction will give us the same probability distribution). First we pick an integer according to the binomial distribution $Bin(C'nq, 1 - (1 - r)^{C'nq})$, then, given the value of the picked integer, we pick a subset $X_1''$ uniformly at random from all subsets of $X_1$ of this cardinality. Independently we choose $X_1''$, uniformly at random from all $C'nq$-element subsets of $X_2$. Then to each vertex $x_1 \in X_1''$, to create an edge, we add one vertex, chosen uniformly at random from the set $X_1''$. For all $x_1 \in X_1''$ the choices of the second vertex are independent with repetition. Therefore by the above construction, \((10)\) and Fact 8

$$X(M) \preceq_{1-o(1/n)} X(C'nq) \preceq_{1} N_2,$$

where $X(M)$ and $X(C'nq)$ are defined as in \((6)\) for $K = 2$, $n_2 = C'nq$, $P_2 = (1 - (1 - r)^{C'nq})/(C'nq)$ and $M$ with Poisson distribution $Po(C'nq - \sqrt{3C'nq \ln n})$. Thus by Fact 7 $X(M)$ has binomial distribution $Bin(C'nq, s_2(r))$, where

$$s_2(r) = 1 - \exp(-(C'nq - \sqrt{3C'nq \ln n})(1 - (1 - r)^{C'nq})/C'nq) = 1 - \exp((1 + o(1))C'nqr) \sim C'nqr.$$
Using the above lemma we will show the existence of the coupling between the random variable $M_2$ denoting the size of an edge set in $H^{(2)}(n, q^3)$ and $N_2$.

**Lemma 15.** Let $C' = 5$, $M_2$ have binomial distribution $\text{Bin}(n^2, q^3)$ and let $N_2$ be the size of the maximum matching in $H^{(2)}_*(C' nq, q)$. Then

$$M_2 \preceq_{1-o(1/n)} N_2.$$ 

**Proof.** By previous lemma and Fact 1 it is sufficient to show

$$M_2 \preceq_{1-o(1/n)} N'_2,$$

where $N'_2$ has binomial distribution $\text{Bin}(C' nq, s_2(q))$ for some $s_2(q) \sim C'nq^2$.

Notice that

$$M_2 = \sum_{i=1}^{nq} \xi_i,$$

where $\xi_i$ are independent with distribution $\text{Bin}\left(\frac{n}{q}, q^3\right)$;

$$N'_2 = \sum_{i=1}^{nq} \zeta_i,$$

where $\zeta_i$ are independent with distribution $\text{Bin}(C', s_2(q))$.

Since $s_2(q) \sim C'nq^2$, for large $n$ we have

$$\forall 1 \leq l \leq 4 \Pr\{\xi_i = l\} \leq \frac{1}{l!}(nq^2)^l \leq \frac{(C')^l}{l!} s_2^l (1 - s_2)^{C' - l} = \Pr\{\zeta_i = l\}$$

and

$$\Pr\{\xi_i > 4\} \leq \left(\frac{3}{5}\right)^{q^5} \leq (nq^2)^5 = \frac{1}{n^2q} \left(nq^2\right)^7 q^2 = o\left(\frac{1}{n^2q}\right).$$

Therefore, for all $1 \leq i \leq nq$ it is simple to construct the probability measure on $\mathbb{N} \times \mathbb{N}$, which existence implies

$$\xi_i \preceq_{1-o(1/n^2p)} \zeta_i.$$ 

This by Fact 2 implies (26).

Since we will compare the sizes of the maximum matchings in $H^{(2)}(n, q^3)$ and $H^{(2)}_*(C'nq, q)$, we will introduce the additional notation. Let $G$ be a set of 2–partite graphs with 2–partition $(X_1, X_2)$. We define

- $M(l)$ - the subset of $G$ containing all graphs with a maximum matching of cardinality $l$;
- $M_1(l)$ - the subset of $M(l)$ containing all graphs with the maximum degree 1;
- $M_2(l)$ - the subset of $M(l)$ containing all graphs with the maximum degree 2 and exactly one vertex of degree 2

and

$$M_1 = \bigcup_{l=0}^{n} M_1(l), \quad M_2 = \bigcup_{l=0}^{n} M_2(l).$$
For \( q = o(n^{-2/3}) \)

\[
\Pr \{ H^{(2)}(n, q^3) \notin \mathcal{M}_1 \} \leq 2n \left( \frac{n}{2} \right) q^6 = O(n^4q^6n^{-1}) = o\left( \frac{1}{n} \right)
\]

and for \( q = O\left( (n^{-2} \ln n)^{1/3} \right) \)

\[
\Pr\{ H^{(2)}(n, q^3) \notin \mathcal{M}_1 \cup \mathcal{M}_2 \} \leq 2\left( \frac{n}{2} \right) \left( \left( \left( \frac{n}{2} \right) q^6 \right)^2 \right) + n^2 \left( \left( \frac{n-1}{2} \right) q^6 \right)^2 + n^2 q^3 \left( (n-1) q^3 \right)^2 + 2n \left( \frac{n}{3} \right) q^9 =
\]

\[
= O(n^6q^{12} + n^4q^9) = O\left( (n^2q^3)^4 n^{-2} + (n^2q^3)^3 n^{-2} \right) = o\left( \frac{1}{n} \right)
\]

Now let

\[
\mu : \mathbb{N} \times \mathbb{N} \to [0, 1]
\]

be the probability measure associated with the coupling of \( M_2 \) and \( N_2 \) from Lemma \[15\]. Let \( \mathcal{G} \) be the set of all 2–partite graphs with 2-partition \((X_1, X_2)\). Starting with the probability measure \( \mu \) we will construct the coupling, which implies for large \( n \)

\[
H^{(2)}(n, q^3) \asymp_{1-o(1/n)} H^{(2)}_*(C'q^2, q).
\]

We introduce additional notation. Let \( H^{(2)}(\mathcal{M}_1) \) be the random graph constructed in the following way. First we sample \( H \) according to the probability distribution of \( H^{(2)}(n, q^3) \) and if \( H \) is not in \( \mathcal{M}_1 \), then we replace it by a graph chosen uniformly at random from \( \mathcal{M}_1(|E(H)|) \). Moreover let \( H^{(2)}(\mathcal{M}_1 \cup \mathcal{M}_2) \) be the random graph constructed by sampling \( H \) according to the probability distribution of \( H^{(2)}(n, q^3) \) and replacing it by a graph chosen uniformly at random from \( \mathcal{M}_1(|E(H)|) \) in the case if it is not contained in \( \mathcal{M}_1 \cup \mathcal{M}_2 \). The sizes of the edge sets of \( H^{(2)}(\mathcal{M}_1) \) and \( H^{(2)}(\mathcal{M}_1 \cup \mathcal{M}_2) \) are random variables \( M_2(\mathcal{M}_1) \) and \( M_2(\mathcal{M}_1 \cup \mathcal{M}_2) \), respectively. Obviously \( M_2(\mathcal{M}_1) \) and \( M_2(\mathcal{M}_1 \cup \mathcal{M}_2) \) have the same distribution as \( M_2 \). For any event \( \mathcal{A} \), denote by \( H^{(2)}(\mathcal{M}_1)_{\mathcal{A}} \), \( H^{(2)}(\mathcal{M}_1 \cup \mathcal{M}_2)_{\mathcal{A}} \) and \( H^{(2)}_*(C'q^2, q)_{\mathcal{A}} \) the graphs \( H^{(2)}(\mathcal{M}_1) \), \( H^{(2)}(\mathcal{M}_1 \cup \mathcal{M}_2) \) and \( H^{(2)}_*(C'q^2, q) \) under condition \( \mathcal{A} \).

In the case \( q = o(n^{-2/3}) \) the construction of the coupling is simple. By the above calculation

\[
H^{(2)}(n, q^3) \asymp_{1-o(1/n)} H^{(2)}(\mathcal{M}_1),
\]

Therefore we will only show

\[
H^{(2)}(\mathcal{M}_1) \asymp_{1-o(1/n)} H^{(2)}_*(C'q^2, q).
\]

Let \((l_1, l_2) \in \mathbb{N} \times \mathbb{N}\) be chosen according to the probability measure \( \mu \). If \( l_1 > l_2 \), then we sample \( H^{(2)}(\mathcal{M}_1)_{|M_2=l_1|} \) and \( H^{(2)}_*(C'q^2, q)_{|N_2=l_2|} \) independently. And if \( l_1 \leq l_2 \), then first we sample an instance of \( H^{(2)}_*(C'q^2, q)_{|N_2=l_2|} \) and then choose its subgraph uniformly at random from all its subgraphs contained in \( \mathcal{M}_1(l_2) \). Then, from the chosen subgraph, we delete \( l_2 - l_1 \) edges chosen uniformly at random. In this way we get an edge set of \( H^{(2)}(\mathcal{M}_1)_{|M_2=l_1|} \).
Now we will construct a coupling in the case \( q = \Omega(n^{-2/3}) \) and \( q = O((n^{-2} \ln n)^{1/3}) \). Let

\[
P_1(l) = \Pr\{H^{(2)}(\mathcal{M}_1 \cup \mathcal{M}_2)^{[M_2=\ell]} \in \mathcal{M}_1\}; \quad P_2(l) = \Pr\{H^{(2)}_*(C'nq, q)^{[N_2=\ell]} \in \mathcal{M}_1\};
\]

\[
Q_1(l) = \Pr\{H^{(2)}(\mathcal{M}_1 \cup \mathcal{M}_2)^{[M_2=\ell]} \notin \mathcal{M}_1\}; \quad Q_2(l) = \Pr\{H^{(2)}_*(C'nq, q)^{[N_2=\ell]} \notin \mathcal{M}_1\}.
\]

We will show that

\[
H^{(2)}(\mathcal{M}_1 \cup \mathcal{M}_2) \preceq_{1-o(1/n)} H^{(2)}_*(C'nq, q).
\]

Let \((l_1, l_2) \in \mathbb{N} \times \mathbb{N}\) be chosen according to the probability measure \(\mu\). If \(l_1 > l_2\) or \(l_2 \geq \omega(n) \ln n\) (where \(\omega(n)\) is a sequence tending slowly to infinity), then we construct a pair of graphs from \(\mathcal{G}\) by sampling independently \(H^{(2)}(\mathcal{M}_1 \cup \mathcal{M}_2)^{[M_2=l_1]}\) and \(H^{(2)}_*(C'nq, q)^{[N_2=l_2]}\). If \(1 \leq l_1 \leq l_2 < \omega(n) \ln n\), then we sample \(H\), a second graph in a pair, according to the probability distribution of \(H^{(2)}_*(C'nq, q)^{[N_2=l_2]}\). If \(H \in \mathcal{M}_1\), then we choose a first graph uniformly at random from all subgraphs of \(H\) contained in \(\mathcal{M}_1(l_1)\). If \(H \notin \mathcal{M}_1\), then with probability \((P_1(l_1) - P_2(l_2))/Q_2(l_2)\) we choose a first graph uniformly at random from all subgraphs of \(H\) contained in \(\mathcal{M}_1(l_1)\) and with probability \(Q_1(l_1)/Q_2(l_2)\) we choose a first graph uniformly at random from all subgraphs of \(H\) contained in \(\mathcal{M}_2(l_1)\).

According to this construction the first graph is chosen according to the probability distribution of \(H^{(2)}(\mathcal{M}_1 \cup \mathcal{M}_2)\) and the second according to the probability distribution of \(H^{(2)}_*(C'nq, q)\). Moreover

\[
\mu(\{(l_1, l_2) : l_1 > l_2\}) = o\left(\frac{1}{n}\right).
\]

In addition, the size of the maximum matching (i.e. \(N_2\)) is at most the number of edges of \(H^{(2)}_*(C'nq, q)\), which has binomial distribution with expected value \((C'n)^2q^2 = O(\ln n)\). Thus by Chernoff bound \((\ref{eq:chernoff})\)

\[
\mu(\{(l_1, l_2) : l_2 \geq \omega(n) \ln n\}) = o\left(\frac{1}{n}\right)
\]

Therefore this is a desired coupling and it is well defined for large \(n\) if \(P_1(l_1) \geq P_2(l_2)\) for large \(n\) and \(l_1 \leq l_2\). Calculations show that for a given \(l < \omega(n) \ln n\) and \(\omega(n)\) tending slowly to infinity

\[
Q_1(l) \leq 1 - \frac{\binom{n}{l}^2(l!)}{(n^2)^l} = 1 - \prod_{i=0}^{l-1} \left(\frac{(n-i)^2}{n^2-i}\right) = 1 - \prod_{i=0}^{l-1} \left(1 - \frac{2ni-i^2}{n^2-i}\right) \leq 1 - \prod_{i=0}^{l-1} \left(1 - \frac{2l}{n}\right) \leq 2l^2/n.
\]
and

\[ Q_2(l) = \Pr\{H_s^{(2)}(C'q, q) \in M_1 \} \geq \frac{\Pr\{H_s^{(2)}(C'q, q) \notin M_1 \} - \Pr\{H_s^{(2)}(C'q, q) \notin M_1 \cup M_2 \}}{1 - \Pr\{H_s^{(2)}(C'q, q) \in M_1 \cup M_2 \}} = \]

\[ = \frac{\Pr\{H_s^{(2)}(C'q, q) \notin M_1 \} + \Pr\{H_s^{(2)}(C'q, q) \in M_2 \}}{\Pr\{H_s^{(2)}(C'q, q) \in M_1 \}} = \]

\[ = \Pr\{H_s^{(2)}(C'q, q) \in M_2(l) \} = \Omega(n_2^{\ast}) = \Omega(n^{1/3}), \]

since

\[ \Pr\{H_s^{(2)}(C'q, q) \in M_2(l) \} = \]

\[ = \binom{C'nq}{l} \binom{C'nq}{l} \left( \frac{l+1}{l} \right) \left( \frac{q}{1-q} \right)^{l+1} (1-q)^{C'nq} = \]

\[ = \binom{C'nq}{l}^2 \binom{C'nq-l}{l+1} \cdot \left( \frac{l+1}{l} \right) \left( \frac{q}{1-q} \right)^{l} \frac{q}{1-q} (1-q)^{C'nq} = \]

\[ = \Pr\{H_s^{(2)}(C'q, q) \in M_1(l) \}(1+o(1)) \frac{C'nq^2l}{2}. \]

Hence \( Q_1(l_1) = o(Q_2(l_2)) \) uniformly over all \( 1 \leq l_1 \leq l_2 \leq \omega(n) \ln n \) and \( \omega(n) \) such that \( \omega(n) \ln n = o(nq) \).

**CASE 3 and 4**

In this case we will use the fact that the number of edges in \( H^{(2)}(n, q^3) \) and \( H_s^{(2)}(C'q, q) \) is sharply concentrated around its expected value. Therefore in order to compare \( H^{(2)}(n, q^3) \) and \( H_s^{(2)}(C'q, q) \) we will compare the auxiliary random graphs. In fact we will construct a coupling of the degree sequences of some multigraphs of which \( H^{(2)}(n, q^3) \) and \( H_s^{(2)}(C'q, q) \) are underlying graphs.

By Fact 4 graphs \( H^{(2)}(n, q^3) \) and \( H_s^{(2)}(C'q, q) \) are underlying graphs of multigraphs \( H^{**}(x) \) and \( T^{**}(x) \) with the Poisson distribution of the number of edges with expected value \(-n^2 \ln(1 - q^3)\) and \(-(C'q)^3 \ln (1 - q)\), respectively and an edge set constructed by independently choosing one by one with repetition edges from the edge set of the complete 2–partite graph with 2–partition \((X_1, X_2)\) and \((X'_1, X'_2)\), respectively. By Chernoff bound \( H^{**}(x) \preceq 1-o(1/n) H^{***}(x) \) and \( T^{**}(x) \preceq 1-o(1/n) T^{***}(x) \),

where \( H^{***}(x) \) is the multigraph with \( C_1 n^2 q^3 \) (\( C_1 > 1 \) is a constant) edges and \( T^{***}(x) \) is the multigraph with \((C')^2 n^2 q^3\) edges (where \( C'/C' > 1 \) is a constant). Notice that choosing one edge is equivalent to choosing its 2 vertices independently from each set of 2–partition.
Therefore we may independently choose the degree sequence in each set $X_i$ ($X'_i$) and then create the multigraph with a given degree sequence. Thus, by Fact 2 to prove that

$$H^{***}(x) \preceq_{1-o(1/n)} T^{***}(x),$$

we have to prove for each $X_i$, $1 \leq i \leq 2$, that

$$(D^{(1)}_1, \ldots, D^{(1)}_n) \preceq_{1-o(1/n)} (D^{(5)}_1, \ldots, D^{(5)}_n),$$

where $D^{(1)}_j$ is the random variable denoting the degree of the $j$-th vertex in $X_i$ in $H^{***}(x)$ and $D^{(5)}_j$ is the random variable denoting the degree of the $j$-th vertex in $X_i$ in $T^{***}(x)$.

We introduce the auxiliary urn models. Assume that we have $n$ urns. Let $D^{(i)} = (D^{(i)}_1, \ldots, D^{(i)}_n)$ be a random vector in which $D^{(i)}_j$ stays for a number of balls in the $j$-th urn in the $i$-th model. Let $1 < C_1 < C_2$, $C''/C''' > 1$, $C'/C'' > 1$ be constants such that $C_2 < C'''$.

- In the 1–st model we throw $C_1 n^2 q^3$ balls one by one independently, with repetition, to the urn chosen uniformly at random from $n$ urns.

- In the 2–nd model the number of thrown balls has Poisson distribution $P(C_2 n^2 q^3)$, i.e. by Fact 7 $D^{(2)}_j$ has Poisson distribution $P(C_2 n q^3)$ (for $K = 2$, $P_2 = \frac{1}{n}$, $n_2 = n$).

- In the 3–rd model $D^{(3)}_j = D_j \cdot D'_j$, where $D_j$ is a Bernoulli random variable with probability of success $C'''q$ and $D'_j$ has Poisson distribution $P(C''' n q^2)$.

- In the 4–th model first we select $C'' n q$ urns from the set of all urns and the number of balls thrown to the selected urns has Poisson distribution $P((C'''^2 n^2 q^3)$, i.e. for the urns not selected $D^{(4)}_j = 0$ and for the selected urns $D^{(4)}_j$ has Poisson distribution $P(C''' n q^2)$ (by Fact 7 for $K = 2$, $P_2 = \frac{1}{C''' n q^2}$, $n_2 = C''' n q$).

- In the 5–th model first we select $C'' n q$ urns and we throw $(C''^2 n^2 q^3$ balls one by one independently to the urn chosen uniformly at random from the set of selected urns.

By Chernoff bound

$$D^{(1)} \preceq_{1-o(1/n)} D^{(2)} \text{ and } D^{(3)} \preceq_{1-o(1/n)} D^{(4)} \preceq_{1-o(1/n)} D^{(5)}.$$  

Moreover, by Fact 2 if for large $n$

$$(27) \quad D^{(2)}_j \preceq_{1-o(1/n^2)} D_j \cdot D'_j,$$

then for large $n$

$$(28) \quad D^{(2)} \preceq_{1-o(1/n)} D^{(3)}$$

The constants may be chosen such that

$$C''' = \begin{cases} 
4, & \text{for } nq^2 = o(1); \\
\omega'''(n), & \text{for } nq^2 = \Theta(1); \\
c''' nq^2, & \text{for } nq^2 \rightarrow \infty,
\end{cases}$$
where \(c'''' > 1\) and \(\omega'''(n)\) is a function tending slowly to infinity.

For large \(n\)

\[
\Pr \left\{ \mathbb{D}_j^{(2)} \geq 1 \right\} = 1 - \exp(-C_2nq^3) \leq C''''q \left( 1 - \exp\left( -\left( C''''nq^2 \right) \right) \right) = \Pr \left\{ \mathbb{D}_j^{(3)} \geq 1 \right\}.
\]

Moreover, for \(t \geq 2\), \(nq^2 = o(1)\) and large \(n\)

\[
\Pr \left\{ \mathbb{D}_j^{(2)} \geq t \right\} \sim \frac{(C_2nq^3)^t}{t!} = o \left( C''''q \left( C''''nq^2 \right)^t \right) = o \left( \Pr \left\{ \mathbb{D}_j^{(3)} \geq t \right\} \right)
\]

This implies (27) for \(nq^2 = o(1)\).

Let now \(nq^2 = \Omega(1)\). By Chernoff bound, if we estimate the number of urns with at least one ball and compare it to the number of balls we will get, that with probability \(1 - o(1/n)\) the number of urns with at least 2 balls in the 3-rd model is \(o(n^2q^3)\) and \(\Omega(n^2q^3)\) in the 2-nd model. Therefore, since urns with at least 2 balls are uniformly distributed, the coupling is easy to construct.

**Proof of (24)**

Let \(C = C(q)\) be defined as in (20). Define \(X_n(x_1, x_2) = |\{x \in \mathcal{X}_3 : x_1 \in \mathcal{X}_1^*(x), \ldots, x_2 \in \mathcal{X}_2^*(x)\}|. \ X_n(x_1, x_2)\) has binomial distribution \(\text{Bin}(n, (Cq)^2)\). Then for

\[
\mathbb{E}X_n = C^2nq^2 = \begin{cases} cnq^2, & \text{for } nq^2 = o(1); \\ \omega^2(n)nq^2, & \text{for } nq^2 = \Theta(n); \\ cn^3q^5, & \text{for } nq^2 \rightarrow \infty. \end{cases}
\]

Therefore by Lemma 3

\[
\Pr \{ \exists (x_1, x_2) X_n(x_1, x_2) \geq a'_n(q) \} \leq n^2 \Pr \{ X_n(x_1, x_2) \geq a'_n(q) \} = o(1),
\]

where

\[
a'_n(q) = \begin{cases} 3 \ln n/(\ln \ln n - \ln(nq^2)), & \text{for } nq^2 = o(1); \\ 3 \ln n/\ln n, & \text{for } nq^2 = \Theta(n); \\ 3 \ln n/(\ln \ln n - \ln(n^3q^6)), & \text{for } nq^2 \rightarrow \infty \text{ and } nq^2 = o(\sqrt[3]{\ln n}); \\ \omega_1(n)n^3q^6, & \text{for } nq^2 \rightarrow \infty \text{ and } nq^2 = \Theta(\sqrt[3]{\ln n}); \\ cn^3q^5, & \text{for } nq^2 \rightarrow \infty \text{ and } o(nq^2) = \sqrt[3]{\ln n}, \end{cases}
\]

(since \(\omega(n)\) may tend to infinity arbitrarily slowly).

Thus

\[
\bigcup_{x \in \mathcal{X}_{k+1}} T^*(x) \preceq_{1-o(1)} H^{(2)}(n, a_n(q)q).
\]

\[\square\]
The proof of the remaining two lemmas is analogous. However we only need to consider the analogues of CASE 2 and 3, which shortens the proof.

Proof of Lemma 12 and 13. Let \( k = 4 \) or \( k = 5 \). Similarly to the previous proof we will divide the hypergraph \( H^{(k)} \left( n, q \binom{k}{2} \right) \) into \( n \) edge disjoint hypergraphs \( H(x), x \in \mathcal{X}_k \), such that

\[
H^{(k)} \left( n, q \binom{k}{2} \right) = \bigcup_{x \in \mathcal{X}_k} H(x).
\]

For any \( x \in \mathcal{X}_k \), \( H(x) \) will be a hypergraph with the vertex set \( \{x\} \cup \bigcup_{i=1}^{k-1} \mathcal{X}_i \) and an edge set consisting of those edges from \( E \left( H^{(k)} \left( n, q \binom{k}{2} \right) \right) \), which contain \( x \).

\( T(x) \) will be an auxiliary hypergraph, which has the same vertex set as \( H(x) \) and an edge set constructed by the following procedure. First we add each edge \((x, y), y \in \bigcup_{i=1}^{k-1} \mathcal{X}_i \) independently with probability \( Cq \) \((C > 5)\) to the edge set and then independently with probability \( q \binom{k-1}{2} \) we add to the edge set each edge \((x_1, \ldots, x_{k-1}) \in \mathcal{X}_1^* \times \cdots \times \mathcal{X}_{k-1}^*\), where, for each \( 1 \leq i \leq k - 1 \), \( \mathcal{X}_i^* \) is the set of vertices connected by an edge with \( x \). Therefore we have to prove that

\[
(29) \quad GH(x) \preceq_{1-o(1/n)} GT(x)
\]

and that

\[
(30) \quad \bigcup_{x \in \mathcal{X}_k} GT(x) \preceq_{1-o(1)} G^{(k)} (n, a(c_k q) c_k q).
\]

To get (30) we will show that if \( T^*(x) \) is a subhypergraph of \( T(x) \) induced on \( \bigcup_{i=1}^{k-1} \mathcal{X}_i \), then for \( c > 5(k - 1) \)

\[
(31) \quad \bigcup_{x \in \mathcal{X}_k} T^*(x) \preceq_{1-o(1)} H^{(k-1)} \left( n, cq \binom{k-1}{2} \right),
\]

where \( H^{(k-1)} \left( n, cq \binom{k-1}{2} \right) \) is independent of choice of \( \mathcal{X}_i^* \), and then we will use the lemma for \((k-1)\) to arrive at

\[
\bigcup_{x \in \mathcal{X}_k} GT^*(x) \preceq_{1-o(1)} G^{(k-1)} (n, a_n(c'_k q) c'_k q),
\]

which by definition of \( T(x) \) and \( T^*(x) \) implies (30).

The proof of (29) in CASE 2 and 3 is similar to the proof of (22). We will only sketch the proof of (31).

Define \( X_n = X_n(x_1, \ldots, x_{k-1}) = |\{x \in \mathcal{X}_k : x_1 \in \mathcal{X}_1^*(x) \cdots x_{k-1} \in \mathcal{X}_{k-1}^*(x)\}| \). It has binomial distribution \( \text{Bin}(n, (Cq)^{k-1}) \) and for large \( n \)

\[
\mathbb{E}X_n = C^3 n q^3 \leq n^{-1/5} \quad \text{and} \quad \mathbb{E}X_n = C^4 n q^4 \leq n^{-3/5}.
\]

Therefore, since

\[
\frac{\ln n}{\ln \ln n - \ln n^{-1/5}} \sim 5 \quad \text{and} \quad \frac{\ln n}{\ln \ln n - \ln n^{-3/5}} \sim \frac{5}{3},
\]

25
by Lemma 3 for any constant $c_4'' > 15$ and $c_5'' > 20/3$

$$\Pr\{\exists (x_1, \ldots, x_{k-1}) X_n(x_1, \ldots, x_{k-1}) \geq c_k''\} \leq n^{k-1} \Pr\{X_n(x_1, \ldots, x_{k-1}) \geq c_k''\} = o(1).$$

Thus in the case $k = 4$, for any constant $c_4' > \sqrt[3]{15}$, we have

$$\bigcup_{x \in X_4} T^*(x) \preceq_{1-o(1)} H^{(3)}(n, (c_4'q)^3).$$

Thus, by Lemma 11

$$\bigcup_{x \in X_5} T^*(x) \preceq_{1-o(1)} G^{(3)}(n, a_n(c_4'q)c_4'q).$$

Finally, since $a_n(c_4'q) \geq 5$,

$$GH^{(4)}(n, q^6) \preceq_{1-o(1)} G^{(4)}(n, a_n(c_4'q)c_4'q).$$

Analogously, for $k = 5$ and $c > \sqrt[3]{20/3}$

$$\bigcup_{x \in X_5} T^*(x) \preceq_{1-o(1)} H^{(4)}(n, (cq)^6).$$

Therefore by Lemma 12 for $c_5' > \sqrt[3]{20/3\sqrt{15}} = \sqrt[3]{2^2 \cdot 3 \cdot 5^3}$

$$\bigcup_{x \in X_5} T^*(x) \preceq_{1-o(1)} G^{(4)}(n, a_n(c_5'q)c_5'q),$$

which implies the thesis. \qed

**Acknowledgments**

I would like to thank Andrzei Ruciński for the suggestion to read the article [8].

**References**

[1] A. D. Barbour, L. Holst, and S. Janson. *Poisson Approximation*. Oxford University Press, 1992.

[2] M. Bloznelis, J. Jaworski, and K. Rybarczyk. Component evolution in a secure wireless sensor network. *Networks*, 53(1):19–26, 2009.

[3] B. Bollobás. *Random Graphs*. Academic Press, 1985.

[4] J. A. Fill, E. R. Scheinerman, and K. B. Singer-Cohen. Random intersection graphs when $m = \omega(n)$: An equivalence theorem relating the evolution of the $G(n, m, p)$ and $G(n, p)$ models. *Random structures and Algorithms*, 16:156–176, 2000.
[5] E. Godehardt and J. Jaworski. Two models of random intersection graphs for classification. In Studies in Classification, Data Analysis and Knowledge Organization, pages 67–81. Springer, Berlin–Heidelberg–New York, 2003.

[6] S. Janson, T. Łuczak, and A. Ruciński. Random Graphs. Wiley, 2001.

[7] M. Karoński, E. R. Scheinerman, and K.B. Singer-Cohen. On random intersection graphs: The subgraph problem. Combinatorics, Probability and Computing, 8:131–159, 1999.

[8] J. H. Kim. Perfect matchings in random uniform hypergraphs. Random Structures and Algorithms, 23(2):111 – 223, 2003.

[9] T. Łuczak. On the equivalence of two basic models of random graphs. In Karoński M., Jaworski J., and Ruciński A., editors, Random Graphs 87’, pages 151–158. John Wiley & Sons, 1990.

[10] K. B. Singer-Cohen. Random intersection graphs. PhD thesis, Department of Mathematical Sciences, The Johns Hopkins University, 1995.

[11] D. Stark and K. Rybarczyk. Poisson approximation of the number of cliques in random intersection graphs. submitted.