Random Graph Processes

Tamás Makai

Thesis submitted to the University of London
for the degree of Doctor of Philosophy

2012

Declaration of Authorship

I, Tamás Makai, hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly stated.

Signed:

(Tamás Makai)

Date:

Summary

This thesis deals with random graph processes. More precisely it deals with two random graph processes which create H-free graphs. The first of these processes is the random H-elimination process which starts from the complete graph and in every step removes an edge uniformly at random from the set of edges which are found in a copy of H. The second is the H-free random graph process which starts from the empty graph and in every step an edge chosen uniformly at random from the set of edges which when added to the graph would not create a copy of H is inserted.

We consider these graph processes for several classes of graphs H, for example strictly two balanced graphs. The class of strictly two balanced graphs includes among others cycles and complete graphs.

We analysed the H-elimination process, when H is strictly 2-balanced. For this class we show the typical number of edges found at the end of the process. We also consider the subgraphs created by the process and its independence number. We also managed to show the expected number of edges in the H-elimination process when $H = K_4^-$, the graph created from the complete graph on 4 vertices by removing an edge and when $H = K_{3,4}^-$ where $K_{3,4}^-$ is created from the complete bipartite graph with 3 vertices in one partition and 4 vertices in the second partition, by removing an edge.

In case of the H-free process we considered the case when H is the triangle and showed that the triangle-free random graph process only creates sparse subgraphs. Finally we have improved the lower bound on the length of the $K_{3,4}^-$-free random graph process.

Acknowledgements

I would like to thank my supervisor Stefanie Gerke for her constant guidance and encouragement. I would also like to thank my family and friends for their support.

Contents

1 Introduction **7**

2 Preliminaries **10**
2.1 Graphs 10
2.2 Asymptotics 11
2.3 Probability 11
2.4 Graph Densities 11
2.5 Rooted Graphs 13
2.6 Probabilistic Theorems 15

3 Random Graph Processes **18**
3.1 Erdős-Rényi random graph process 18
3.2 The H-free processes 19
3.3 H-elimination process 21
3.4 H-removal process 22

4 Previous Results **23**
4.1 H-free process 23
4.2 H-elimination process 39

5 H-elimination random graph process **41**
5.1 Overlapping rooted graphs 41
5.2 Expected number of edges 43
5.3 Concentration 59
5.4 Independence Number 64
5.5 Subgraphs 69

6 Subgraphs of the triangle-free random graph process **71**

7 The $K_{3,4}^{-}$-free process **77**
7.1 The differential equations 77
7.2 Differential equation proof 86

Bibliography **121**

List of Figures

5.1 The $K_{3,4}^{-}$ . . . 53
5.2 The forbidden rooted graphs . . . 53

6.1 open/partial/complete vertices . . . 71
6.2 Closing multiple open pairs . . . 74

7.1 Overlap at an open pair . . . 90

Chapter 1

Introduction

Random graph processes have received considerable attention lately. These processes provide a natural context for modelling complex networks which evolve over time. They also provide a simple way, and in some cases the only way, to construct a counterexample in extremal graph theory. There are also strong connections to algorithmic graph theory as examining these processes provide good estimates on the typical running time of algorithms.

There are multiple ways to define a random graph process. The first of these is known as the Erdős-Rényi random graph process, which was introduced by Erdős and Rényi [14]. The process starts from the empty graph on n vertices and in every step an edge is selected uniformly at random from the set of non-edges and inserted into the graph. The process stops after m edges have been inserted. Even though we call this a random graph this actually describes a probability space over the graphs on n vertices containing m edges, where each of these graphs is selected with the same probability. The aim is to study the likely structural properties of the graph created by this process when the number of vertices is large. We say that a property holds asymptotically almost surely (a.a.s.) if the probability that the property holds tends to one as the number of vertices tends to infinity. The properties of the Erdős-Rényi random graph process are well known, due to its connection to the Erdős-Rényi random graph, where every edge is inserted independently with probability $p \approx \left(m/\binom{n}{2}\right)$.

One of the significant results of this process was improving the lower bound of the Ramsey number $R(k,t)$. We have that $R(k,t) > n$ if there is a graph on n vertices which neither contains a complete graph on k vertices as a subgraph nor does it contain an independent set of size t i.e. a set where no two vertices in the set are connected. Constructing such graphs is difficult.

The Erdős-Rényi random graph process a.a.s. creates graphs with small independent sets. Therefore if one chooses m as large as possible such that the process a.a.s. has no copies of a complete graph on k vertices one has a lower bound on the Ramsey number. One can improve the bound slightly by allowing a small number of copies of a complete graph on k vertices and removing every edge from every copy as this has almost no affect on the size of the independent sets. In order to increase the number of edges in the final graph the following two modifications were suggested at a conference by Bollobás and Erdős [10]. The processes first appeared in print in Erdős, Suen and Winkler [17].

The first of these processes is called the *H-free random graph process* which starts out from the empty graph on n vertices. In every step of the process we select an edge from the set of non-edges which when added to the graph do not create a copy of H, and insert it into the graph. The process stops once no more edges can be added.

The second process, called the *H-elimination process*, runs in the opposite direction. This process starts from the complete graph on n vertices and in every step of the process an edge selected uniformly at random from the set of edges contained in a copy of H is removed. The process stops after no more copies of H exist.

This thesis deals with these two processes. Similarly to the Erdős-Rényi random graph process we are interested in the likely structural properties of the graph created by these processes. However, in contrast to the Erdős-Rényi random graph process, which always finishes with the complete graph after inserting all of the $\binom{n}{2}$ possible edges, already determining the typical length of the H-free and H-elimination process can prove challenging.

Multiple equivalent definitions exist for both the H-free and H-elimination random graph process. We will discuss these in Chapter 3. The alternative definitions indicate a connection between the Erdős-Rényi, the H-free, and the H-elimination graph processes. This connection was used by Osthus and Taraz [31] and independently by Bollobás and Riordan [9] to establish bounds on the H-free process. These results can also be transferred to the H-elimination process for more details see Chapter 4.

We examine the H-elimination process for a special class of graphs, namely the strictly 2-balanced graphs. Let $v(H)$ and $e(H)$ denote the number of vertices and the number of edges of the graph H respectively. A graph H is strictly 2-balanced if

$$\frac{e(H')-1}{v(H')-2} < \frac{e(H)-1}{v(H)-2}$$

holds for every proper subgraph H' of H which contains more than two vertices. In Chapter 5 we will show the expected number of edges contained in this process and that the number of edges is a.a.s. concentrated around its expectation. Formally we will show that there exists a constant c depending on H such that the final graph created by this process contains a.a.s.

$$(1 + o(1))cn^{2-(v(H)-2)/(e(H)-1)}$$

edges. We also investigate the subgraphs created by this process and have shown that a.a.s. every sparse H-free graph is created by this process, but no dense subgraphs are. We also include bounds on the independence number of the graph created by this process which hold a.a.s.. In case H is a complete graph we show that the independence number of the graph created by the H-elimination process is a.a.s. larger then that created by the H-free process and thus provides a worse bound for the Ramsey number.

The remainder of the thesis deals with the H-free random graph process. When H is a star Ruciński and Wormald [38] have shown that the number of edges in the final graph is a.a.s. the maximal possible number of edges in a H-free graph. The next breakthrough was made by Bohman [4], who showed upper and lower bounds matching up to a constant factor for the number of edges in the triangle-free process which hold a.a.s.. He also determined an upper bound on the independence number of the process and thus was able to recreate Kim's [27] lower bound on the asymptotics of the off-diagonal Ramsey number $R(3, t)$. The results have been extended to several other graphs and classes of graphs [6], [46], [51], [33], [34], [47], [32], for more details see Chapter 4.

The subgraphs of the random H-free graphs have also been considered. When H is a triangle Bohman and Keevash [6] and Wolfovitz [50] established subgraph counts which hold for the early stages of the process. Their results imply that a.a.s. only sparse triangle-free graphs appear during the early stages of the triangle-free process. However denser H-free graphs may appear during the later stages of the process. We will show in Chapter 6 that this is not the case, namely a.a.s. no dense triangle-free graph appears in the triangle-free process. More precisely we show that there exists a constant δ such that if $(e(F)/v(F)) > \delta$ then F is a.a.s. not contained in the random triangle-free graph process.

In the final chapter of the thesis we examine the $K_{3,4}^-$-free random graph process. We apply the differential equation method to this random graph process in order to establish a lower bound on the number of edges contained in the graph. We show that the final graph contains at least $Cn^{3/2}\sqrt{\log\log n}$ edges for some constant C. This improves the previous best known lower bound established by Osthus and Taraz [31].

Chapter 2

Preliminaries

This chapter introduces the notation and definitions used throughout this thesis. It also contains the proof of several simple propositions which will be relevant later.

2.1 Graphs

A graph $G = (V, E)$ is a set of vertices V and a set of edges E where $E \subseteq \{\{u, v\} : u, v \in V, u \neq v\}$. We denote the vertex set of a graph by $V(G)$ and its edge set with $E(G)$ also $v(G) = |V(G)|$ and $e(G) = |E(G)|$. The conjugate of a graph $G = (V, E)$ denoted by $\overline{G}$ is a graph on the same set of vertices, however an edge is present in $\overline{G}$ iff it is not present in G.

There are several important classes of graphs. The complete graph on k vertices is denoted by K_k and the cycle on ℓ vertices is denoted by C_ℓ. Also P_ℓ denotes a path of length ℓ, where the length of a path is determined by the number of edges contained in it. The complete bipartite graph, where one partition has r_1 and the other partition has r_2 vertices is denoted by K_{r_1,r_2}. In many cases we will be interested in graphs which are created from another graph by removing specific edges. Let $\{e_1, e_2, ...e_k\} \subseteq E(G)$ then $G_{\{e_1,e_2,...e_k\}}$ denotes the graph created from G by removing the edges $\{e_1, e_2, ...e_k\}$. Also let $R \subseteq V(G)$ then G_R denotes the graph where every edge between the vertices of R is removed. The notation G^- will be used for G_e when G is symmetric meaning that removing any of its edges gives the same graph, e.g. $P_{\ell-1} = C_\ell^-$.

We are interested in several graph properties. The notation $\delta(G)$ is used for the minimal degree of G and $\Delta(G)$ for the maximal degree. The size of the largest independent set in G, also referred to as the independence number of G, is denoted by $\alpha(G)$.

Let F and G be graphs such that $v(F) \leq v(G)$. For an injective function $f : V(F) \to V(G)$ let $f(E(F))$ denote the set of vertex pairs in G which correspond to the edges of F i.e. $f(E(F)) = \{\{f(u), f(v)\} : \{u, v\} \in E(F)\}$.

We say that there is a copy of F in G if there is an injective function $f : V(F) \to V(G)$ such that $f(E(F)) \subseteq E(G)$. A graph F is a subgraph of G if $V(F) \subseteq V(G)$ and $E(F) \subseteq E(G)$ and is denoted by $F \subseteq G$. We say that F is a proper subgraph of G, denoted by $F \subsetneq G$, if F is a subgraph of G and F is not isomorphic to G. Also for $S \subseteq V(G)$ we denote by $G[S]$ the spanning subgraph of G on S i.e. $V(G[S]) = S$ and $E(G[S]) = \{\{u, v\} \in E(G) : u, v \in S\}$.

The Ramsey-number is denoted by $R(k, l)$. We say that $R(k, l) = n$ if n is the smallest number such that any two colouring of the edges of K_n, with red and blue either contains a blue K_k or a red K_l.

The Turán number denoted by $\mathrm{ex}(n, H)$ is the maximal number of edges that can be found in a graph on n vertices without creating a copy of H.

2.2 Asymptotics

When describing asymptotics we will use the following notation: Let $f, g : \mathbb{N} \to \mathbb{R}^+$ then

$$\begin{aligned}
f(n) = O(g(n)) &\iff \exists C\ s.t.\ f(n)/g(n) \le C \\
f(n) = \Omega(g(n)) &\iff g(n) = O(f(n)) \\
f(n) = \Theta(g(n)) &\iff f(n) = O(g(n))\ and\ g(n) = O(f(n)) \\
f(n) = o(g(n)) &\iff \lim_{n\to\infty} f(n)/g(n) = 0 \\
f(n) = \omega(g(n)) &\iff g(n) = o(f(n))
\end{aligned}$$

2.3 Probability

A probability space is defined by a triple $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is the sample space, $\mathcal{F}$ is a σ-algebra and $\mathbb{P} : \mathcal{F} \to \{0, 1\}$.

In case of a random graph processes described above we consider graphs on n vertices. Let Ω be the set of all maximal sequences of distinct pairs in $E(K_n)$ satisfying the condition of the process. Each element $\omega \in \Omega$ describes a possible evolution of the process. Note that the probability measure on Ω is not necessarily uniform. However in the case of the H-elimination and H-free random graph process it is given by the uniform random choice at each step. Define a family of equivalence classes on the elements of Ω, where two elements belong to the same class in case their first i entries match. Denote the partitions created by the equivalence classes with $\mathcal{P}_i$. The parts of $\mathcal{P}_i$ are called *atoms*. For every i the partition $\mathcal{P}_i$ generates a σ-algebra $\mathcal{F}_i$. The collection $\mathcal{F} = \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq ...$ is called the *natural filtration* of the process. Given a random variable X and a σ-algebra $\mathcal{F}$ we say that X is $\mathcal{F}$ measurable if it is constant on each atom of $\mathcal{F}$. Define the *conditional expectation* of a random variable X as:

$$\mathbb{E}(X|\mathcal{F})(\omega) = \sum_{\omega' \in [\omega]} \mathbb{P}(\omega')X(\omega') \Bigg/ \sum_{\omega' \in [\omega]} \mathbb{P}(\omega'),$$

where $[\omega]$ denotes the atom of $\mathcal{F}$ containing ω. Note that $\mathbb{E}(X|\mathcal{F})$ is a random variable and that it is measurable.

2.4 Graph Densities

Let H be a graph, then we define the density of H as

$$d(H) = \frac{e(H)}{v(H)}$$

and the maximal density of H as

$$m(H) = \max_{H' \subseteq H} d(H').$$

We say that a graph is strictly balanced if for every $H' \subsetneq H$ $d(H') < d(H)$ and it is balanced if $m(H) = d(H)$. The 2-density of a graph with more than 2 vertices is defined similarly

$$d_2(H) = \frac{e(H) - 1}{v(H) - 2}$$

and the maximal 2-density

$$m_2(H) = \max_{\substack{H' \subseteq H \\ v(H') > 2}} d_2(H').$$

A graph is strictly 2-balanced if for every $H' \subsetneq H$ with $v(H') > 2$ we have $d_2(H') < d_2(H)$ and it is 2-balanced if $m_2(H) = d_2(H)$. When determining whether a graph is strictly 2-balanced it is enough to examine the densities of every spanning subgraph i.e. H is strictly 2-balanced if for every $S \subsetneq V(H)$ with $|S| > 2$ we have that $d_2(H[S]) < d_2(H)$. An analogous statement holds for 2-balanced graphs.

Bohman and Keevash [6] have shown that strictly 2-balanced graphs are 2-connected. Osthus and Taraz [31] noted that they are also strictly balanced. In Proposition 2.1 we show that every 2-balanced graph is balanced, however they are not necessarily connected as a complete matching on $2k$ vertices where $k > 1$ is 2-balanced. In Proposition 2.2 we will show that this is the only class of 2-balanced graphs which are not connected. Note that adding an isolated vertex to a graph decreases its 2-density and thus no 2-balanced graph contains an isolated vertex. Let H be a 2-balanced graph, since it does not contain any isolated vertices thus $2e(H) \geq v(H)$ and thus $d_2(H) \geq 1/2$.

Proposition 2.1

Every 2-balanced graph is balanced.

PROOF Let H be a 2-balanced graph and $H' \subsetneq H$. Since H has no isolated vertices, removing k vertices from H removes at least $\lceil k/2 \rceil$ edges. Therefore $e(H) - e(H') \geq (v(H) - v(H'))/2$ or equivalently $2e(H') + v(H) \leq 2e(H) + v(H')$.

Since H is 2-balanced we have that:

$$\frac{e(H') - 1}{v(H') - 2} \leq \frac{e(H) - 1}{v(H) - 2}$$

$$e(H')v(H) - 2e(H') - v(H) \leq e(H)v(H') - v(H') - 2e(H)$$

$$e(H')v(H) \leq e(H)v(H')$$

$$\frac{e(H')}{v(H')} \leq \frac{e(H)}{v(H)}$$

completing the proof. ∎

Excluding the case when H is a maximal matching on $2k$ vertices every 2-balanced graph has maximal 2-density larger than $1/2$ as it must contain a path with 2 edges and $d_2(P_2) = 1$. Using this fact we can show that every 2-balanced graph which is not a matching is connected. Define the family of functions

$$f_d(H) = v(H) - 2 - (e(H) - 1)/d.$$

This family is very useful when analysing densities as $f_d(H) > 0$ iff $d_2(H) < d$, $f_d(H) < 0$ iff $d_2(H) > d$ and $f_d(H) = 0$ iff $d_2(H) = d$.

Proposition 2.2

Every 2-balanced graph H with $d_2(H) > 1/2$ is connected.

PROOF Assume for contradiction that H is 2-balanced with $d_2(H) > 1/2$ and not connected. Therefore one can partition $V(H)$ into V_1, V_2 such that there is no edge joining a vertex in V_1 to a vertex in V_2. Then we have that

$$\begin{aligned}
0 > -2 + \frac{1}{d_2(H)} &= f_{d_2(H)}(H) - 2 + \frac{1}{d_2(H)} = v(H) - 2 - \frac{e(H) - 1}{d_2(H)} - 2 + \frac{1}{d_2(H)} \\
&= |V_1| - 2 - \frac{e(H[V_1]) - 1}{d_2(H)} + |V_2| - 2 - \frac{e(H[V_2]) - 1}{d_2(H)} \\
&= f_{d_2(H)}(H[V_1]) + f_{d_2(H)}(H[V_2]).
\end{aligned}$$

Therefore at least one of the terms is less then 0 and without loss of generality we may assume that it is the first. Thus $0 > f_{d_2(H)}(H[V_1])$ which implies that $d_2(H[V_1]) > d_2(H)$ contradicting the fact that H is 2-balanced. ∎

2.5 Rooted Graphs

Let F be a graph and let $R \subseteq V(F)$, then we call (R, F) a *rooted graph*. We say that (R, F) is *nontrivial* if every vertex in the root has a non-root neighbour. In most cases we will be considering rooted graphs where R forms an independent set.

Definition 2.1

Let (R, F) be a rooted graph and G be a graph such that $v(F) \leq v(G)$ and let $S \subseteq V(G)$ with $|S| = |R|$. For an injective function $\phi : R \to V(G)$ we say that there is a copy of (R, F) with respect to ϕ in G if there exists an injective function $f : V(F) \to V(G)$ such that $f(E(F)) \subseteq E(G)$ and $f|_R = \phi$.

In many cases it is enough that there is a copy of (R, F) with respect to ϕ in G for some $\phi : R \to S$ where $S \subseteq V(G)$. In this case we say that there is a copy of (R, F) rooted at S in G.

Definition 2.2

(R', F') is a rooted subgraph of (R, F) denoted by $(R', F') \subseteq (R, F)$ if (R', F') is a rooted graph, F' is a subgraph of F and $R' = R \cap V(F')$.

We say that (R', F') is a proper rooted subgraph of (R, F) denoted by $(R', F') \subsetneq (R, F)$ in case (R', F') is a rooted subgraph of (R, F) and F' is a proper subgraph of F.

The following density function can be defined for rooted graphs:

$$d(R, F) = \frac{e(F)}{v(F) - |R|}$$

and the maximal density

$$m(R, F) = \max_{R \subsetneq S \subseteq V(F)} d(R, F[S]).$$

We say that a rooted graph is strictly balanced if for every $R \subsetneq S \subsetneq V(F)$ we have $d(R, F[S]) < d(R, F)$ and it is balanced if $d(R, F[S]) \leq d(R, F)$. In the following we show that if (R, F) is a non-trivial strictly balanced rooted graph then $d(R', F') < d(R, F)$ for every $(R', F') \subsetneq (R, F)$.

Proposition 2.3

Let (R, F) be a nontrivial strictly balanced rooted graph and $(R', F') \subsetneq (R, F)$ then $d(R', F') < d(R, F)$.

PROOF In case $R' = R$ then the statement holds as either $V(F') = V(F)$ in which case $e(F') < e(F)$ or $d(R, F') \leq d(R, F[V(F')]) < d(R, F)$.

Now if $R' \subsetneq R$ then create F'' by adding the vertices in $R \backslash R'$ to F', formally $F'' = (V(F') \cup R, E(F'))$. Obviously $d(R', F') = d(R, F'')$. Due to the fact that (R, F) is nontrivial we have that $(R, F'') \subsetneq (R, F)$ and we have already shown that $d(R, F'') < d(R, F)$. ∎

Note that this is not true for trivial rooted graphs, as removing an isolated vertex from the roots does not necessarily have an affect on the density.

Proposition 2.4

Let (R, F) be a balanced rooted graph and $(R', F') \subsetneq (R, F)$ then $d(R', F') \leq d(R, F)$.

PROOF Similarly as before let F'' be the graph created from F' by adding all the vertices in $R \backslash R'$. Obviously $d(R', F') = d(R, F'')$. The result follows as $(R, F'') \subseteq (R, F)$ and since (R, F) is balanced $d(R, F'') \leq d(R, F[V(F'')]) \leq d(R, F)$. ∎

When determining whether a rooted graph is strictly balanced it is easier to work with the following family of functions

$$f_d(R, F) = v(F) - |R| - \frac{e(F)}{d}$$

as $f_d(R, F) = v(F) - |R| - \frac{e(F)}{d} > 0$ iff $d(R, F) < d$ and similarly $f_d(R, F) = v(F) - |R| - \frac{e(F)}{d} < 0$ iff $d(R, F) > d$. Note that if R is an independent set then

$f_d(R,F) = 0$ iff $d(R,F) = d$ or $F = (R,\emptyset)$. Therefore if (R,F) is a nontrivial strictly balanced rooted graph such that R forms an independent set then by Lemma 2.3 we have that $f_{d(R,F)}(R',F') > 0$ for every $(R',F') \subsetneq (R,F)$.

There is a connection between strictly balanced rooted graphs and strictly 2-balanced graphs, namely H is strictly 2-balanced iff for every $r \in E(H)$ the rooted graph (r, H_r) is strictly balanced. Similarly H is 2-balanced iff for every $r \in E(H)$ the rooted graph (r, H_r) is balanced.

Note that rooted graphs generalise the notion of a graph as every graph can be considered as a rooted graph where the roots form an empty set. In this case the notion of density of the rooted graph and the notion of density of the graph coincide.

2.6 Probabilistic Theorems

In this section we introduce some probabilistic Lemma's and Theorem's which will be used during this thesis. The following two Lemmas give lower bounds on the probability that no event from a set of multiple dependent events happen.

Lemma 2.5 (Harris's Lemma[23])

Let Ω be a finite universal set and let E be a random subset of Ω given by

$$\mathbb{P}(e \in E) = p_e$$

where these events are mutually independent over $e \in \Omega$. Let $\{A_i\}_{i\in I}$ be subsets of Ω, where I is a finite index set. Let X_i be the indicator random variable for the event that $A_i \subseteq E$. Then for $X = \sum_{i\in I} X_i$ we have that:

$$\mathbb{P}(X = 0) \geq \prod_{i\in I} \mathbb{P}(X_i = 0).$$

Lemma 2.6

With the setup of Lemma 2.5, if in addition $\mathbb{E}(X_i) < 1$ for every $i \in I$ then

$$\mathbb{P}(X = 0) \geq \exp\left(-\frac{\mathbb{E}(X)}{\max_{i\in I}(1 - \mathbb{E}(X_i))}\right).$$

PROOF Lemma 2.5 gives us that $\mathbb{P}(X = 0) \geq \prod_{i\in I} \mathbb{P}(X_i = 0)$. We also have that

$$\prod_{i\in I} \mathbb{P}(X_i = 0) = \prod_{i\in I}(1 - \mathbb{E}(X_i)) \geq \prod_{i\in I} \exp\left(-\frac{\mathbb{E}(X_i)}{1 - \mathbb{E}(X_i)}\right) \geq \exp\left(-\frac{\mathbb{E}(X)}{\max_{i\in I}(1 - \mathbb{E}(X_i))}\right)$$

completing the proof. ∎

The following two Theorems give upper bounds on the probability that no event from a set of multiple dependent events happens. We need the following definition for the first of these. We say that L is a *dependency graph* for a family of random variables $\{X_i\}_{i\in I}$ if $V(L) = I$ such that if A and B are two disjoint subsets of I and L contains no edge between A and B, then the families $\{X_i\}_{i\in A}$ and $\{X_i\}_{i\in B}$ are mutually independent.

Theorem 2.7 (Suen's inequality [24])
Let X_i, $i \in I$ be a finite family of Bernoulli random variables having dependency graph L. For $i, j \in I$ we write $i \sim j$ if $\{i, j\} \in E(L)$. Let

$$\Delta = \sum_{i\sim j} \mathbb{E}(X_i X_j)$$

and

$$\delta = \max_{i\in I} \sum_{i\sim k} \mathbb{E}(X_k).$$

Then for $X = \sum_{i\in I} X_i$ we have that:

$$\mathbb{P}(X = 0) \leq \exp\left(-\mathbb{E}(X) + \Delta e^{2\delta}\right).$$

Theorem 2.8 (Janson's inequality[25])
Let Ω be a finite universal set and let E be a random subset of Ω given by

$$\mathbb{P}(e \in E) = p_e$$

where these events are mutually independent over $e \in \Omega$. Let $\{A_i\}_{i\in I}$ be subsets of Ω, where I is a finite index set. Let $\mathcal{B}_i$ be the event that $A_i \subseteq E$. For $i, j \in I$ we write $i \sim j$ if $i \neq j$ and $A_i \cap A_j \neq \emptyset$. Let

$$\mu = \sum_{i\in I} \mathbb{P}(\mathcal{B}_i)$$

and

$$\Delta = \sum_{i\sim j} \mathbb{P}(\mathcal{B}_i \cap \mathcal{B}_j).$$

Then:

$$\mathbb{P}\left(\bigcap_{i\in I} \overline{\mathcal{B}_i}\right) \leq \exp\left(-\mu + \Delta\right)$$

and

$$\mathbb{P}\left(\bigcap_{i\in I} \overline{\mathcal{B}_i}\right) \leq \exp\left(-\frac{\mu^2}{2\Delta}\right).$$

Note that we can assign indicator variables X_i to the events $\mathcal{B}_i$ and in such a case the event that none of the $\mathcal{B}_i$ happen is equivalent to the event that $X = \sum X_i = 0$.

Note that if one considers a set of random variables which fulfil the conditions of Janson's inequality, then Suen's inequality can also be applied to them, with the same value of Δ. Thus Janson's inequality gives a stronger bound, but Suen's inequality can be applied more generally. The final Theorem in this section gives an upper bound on the probability that out of a set of events many mutually independent events happen.

Theorem 2.9 ([18])
Let $\{\mathcal{A}_i\}_{i\in I}$ be a set of events in an arbitrary probability space with I finite. Denote $\mu = \sum_{i\in I} \mathbb{P}(\mathcal{A}_i)$. Let Y denote the size of the largest set $J \subseteq I$ such that the set of events $\{\mathcal{A}_j\}_{j\in J}$ are mutually independent and the event $\bigcap_{j\in J} \mathcal{A}_j$ holds. Then

$$\mathbb{P}(Y \geq s) \leq \frac{\mu^s}{s!} \leq \left(\frac{\mathrm{e}\mu}{s}\right)^s .$$

Chapter 3

Random Graph Processes

We start this chapter by showing equivalent definitions for the Erdős-Rényi random graph process. Later we show how these definitions can be modified to create equivalent definitions for the H-free and H-elimination random graph processes.

3.1 Erdős-Rényi random graph process

Definition 3.1

The Erdős-Rényi random graph process *starts out from the empty graph on n vertices and in every step it selects an edge uniformly at random from the set of non-edges and inserts it into the graph.*

Denote the graph created after m vertices were inserted by $G_{n,m}$. There are several alternative definitions for this random graph process. The first of these orders the edges before starting the process, instead of selecting the next edge after every step, formally:

Definition 3.2

Consider the empty graph on n vertices and select a random permutation of the $\binom{n}{2}$ possible edges. Insert the first m edges of this permutation to create the random graph $G_{n,m}$.

Originally Erdős and Rényi [13] considered the following equivalent random graph model.

Definition 3.3

Let $G_{n,m}$ be the graph selected uniformly at random from all labelled graphs on n vertices containing m edges.

The properties of $G_{n,m}$ are well understood mostly due to its connection to $G_{n,p}$ when $m \approx \binom{n}{2}p$.

Definition 3.4

Let $G_{n,p}$ be the random graph on n vertices where each of the $\binom{n}{2}$ edges is present with probability p independently of the presence or absence of the other edges.

A well known equivalent definition for this random graph model $G_{n,p}$ is the following:

Definition 3.5
Consider the complete graph on n vertices and assign to each of its $\binom{n}{2}$ edges a birthtime uniformly and independently distributed on $[0,1]$. An edge is present in $G_{n,p}$ if its birthtime is at most p.

This definition suggests the previously described connection between $G_{n,p}$ and $G_{n,m}$ as ordering the edges according to their birthtimes and inserting the first m edges into the graph creates $G_{n,m}$. Also if $m = \binom{n}{2}p$ such that $m = \omega(1)$ then there are a.a.s. m edges with birthtimes less then p.

Łuczak [29] showed that for a graph property $\mathcal{Q}$ if one knows the asymptotic probability for $G_{n,m}$ and this is the same for every $m \approx p\binom{n}{2}$ than one also has it for $G_{n,p}$.

Theorem 3.1 ([29])
Let $\mathcal{Q}$ be an arbitrary graph property and $p = p(n) \in [0,1]$. If for every graph $G_{n,m}$ such that $m = \binom{n}{2}p + O\left(\sqrt{\binom{n}{2}p(1-p)}\right)$ it holds that $\mathbb{P}(G_{n,m}$ has $\mathcal{Q}) \to a$ as $n \to \infty$, then also $\mathbb{P}(G_{n,p}$ has $\mathcal{Q}) \to a$ as $n \to \infty$.

In the same paper Łuczak [29] also showed the other direction, but only if the graph property is monotone. A graph property is *increasing* if it is preserved under the addition of edges and it is *decreasing* if it is preserved under the removal of edges. A graph property is monotone if it is either increasing or decreasing.

Theorem 3.2 ([29])
Let $\mathcal{Q}$ be a monotone graph property and $0 \leq m \leq \binom{n}{2}$. If for every graph $G_{n,p}$ such that $p = m/\binom{n}{2} + O\left(\sqrt{m(m-\binom{n}{2})/\binom{n}{2}^3}\right)$ it holds that $\mathbb{P}(G_{n,p}$ has $\mathcal{Q}) \to a$ as $n \to \infty$, then also $\mathbb{P}(G_{n,m}$ has $\mathcal{Q}) \to a$ as $n \to \infty$.

Similar results were known beforehand for the cases when the probability that the graph property appears tended to 0 or 1 see for example [8].

3.2 The H-free processes

Definition 3.6 (H-free random graph process)
Fix a finite graph H. Start out from the empty graph on n vertices denoted by $G_f(H)_{n,0}$. In step i consider $O(i)$ the set of non-edges which when inserted into $G_f(H)_{n,i}$ would not create a copy of H. $G_f(H)_{n,i+1}$ is constructed by adding an edge chosen uniformly at random from $O(i)$. The process stops when no more edges can be inserted.

The set $O(i)$ is called the set of *open pairs* at step i and let $Q(i) = |O(i)|$. The set of non-edges which when inserted would create a copy H in $G_f(H)_{n,i}$ is called the set of *closed pairs* at step i and is denoted by $C(i)$. The graph created by this process is denoted by $G_f(H)_n$.

Similarly to the Erdős-Rényi random graph process there are several alternative definitions. Erdős, Suen and Winkler [17] gave the following equivalent definitions. The first of these is based on ordering the edges randomly at the start of the process.

Definition 3.7 (H-free random graph process)
Fix a finite graph H. Consider a random permutation on the edges of the complete graph on n vertices. Start from the empty graph on n vertices and consider the edges in the permutation one at a time. An edge is inserted into the graph if it does not form a copy of H with the previously inserted edges. The process stops after all edges have been considered.

Proposition 3.3
The above definitions of the H-free random graph process are equivalent.

PROOF Note that it is enough to show that at any step of Definition 3.7 we have that the next edge which will actually be inserted is chosen uniformly at random from the set of edges which when inserted into the graph would not create a copy of H. Denote the edges of the complete graph on n vertices with $e_1, ..., e_{\binom{n}{2}}$ and let σ be a permutation on $\left[\binom{n}{2}\right]$. Fix $i < \binom{n}{2}$ and let j_1, j_2 be such that $j_1, j_2 \neq \sigma(j)$ for any $j \leq i$ and inserting either the edge e_{j_1} or the edge e_{j_2} would not create a copy of H. Now we are interested in the number of permutations S_1 such that for all $\sigma_1 \in S_1$ we have that $\sigma_1(j) = \sigma(j)$ for $j \leq i$ and the next edge to be inserted is j_1. Define S_2 analogously. Our aim is to show that $|S_1| = |S_2|$. Let $\sigma_1 \in S_1$ and let σ_2 be the permutation created from σ_1 by switching j_1 and j_2 i.e. $\sigma_2(\sigma_1^{-1}(j_1)) = j_2$ and $\sigma_2(\sigma_1^{-1}(j_2)) = j_1$. Note that $\sigma_2 \in S_2$ and that due to symmetry this operation is a bijection between the two sets. ∎

These first two models resemble the random graph process $G_{n,m}$. Introducing birthtimes gives us a model resembling $G_{n,p}$. In this case birthtimes are used to determine the order the edges are chosen.

Definition 3.8 (H-free random graph process)
Fix a finite graph H. Consider the complete graph on n vertices and assign to each of the possible $\binom{n}{2}$ edges a birthtime uniformly distributed in $[0, 1]$. Now start out from an empty graph on n vertices and increase p gradually. Every time a new edge is born consider adding this edge to the graph. In case adding this edge would create a copy of H in the graph then discard the edge, otherwise insert it into the graph. Edges with equal birthtime (which occur

with probability zero) are considered in arbitrary order. Denote with $G_f(H)_{n,p}$ the graph where all the edges with birthtime at most p have been considered.

Note that the term H-free can be replaced with any graph property preserved by the deletion of edges, i.e. the next edge is chosen uniformly at random from the set of non-edges which when added to the graph the chosen property is still satisfied. In addition to the H-free random graph process, planar [21], bipartite [17] and cycle-free [1] random graph processes have been studied.

3.3 H-elimination process

Definition 3.9 (H-elimination random graph process)

Fix a graph H. Start out from the complete graph on n vertices and in every step of the process select an edge uniformly at random from the set of edges contained in a copy of H and remove it from the graph. The process stops once no more copies of H exist.

Bollobás and Erdős [10] noted that the definition above actually describes a family of processes as the edge which will be removed does not have to be selected according to a uniform distribution, e.g. the edge could be selected in proportion to the number of copies of H it is contained in. However the following equivalent definitions only apply when the following edge is selected uniformly. Similarly to the case of the H-free process this is equivalent to the following definition.

Definition 3.10 (H-elimination random graph process)

Fix a finite graph H. Start from the complete graph on n vertices. Take a random permutation of the edges of K_n and traverse the edges one at a time. An edge is removed from the graph if it is contained in a copy of H otherwise it remains in the graph.

Erdős, Suen and Winkler [17] observed that in this second definition if one has a copy of H in the graph then none of its edges have been traversed, as every traversed edge has either been removed or it is not in a copy of H. Therefore when deciding whether an edge is in a copy of H one only has to consider the edges which have not yet been traversed. Thus we have that an edge remains in the graph iff it does not form a copy of H with the edges which have not yet been traversed. Note that this is equivalent to the process which starts out from the empty graph and in every step of the process and edge is inserted iff it does not form a copy of H with the edges which have not yet been traversed. Reversing the direction in which the edges are traversed leads to the following definition.

Definition 3.11 (H-elimination random graph process)

Fix a finite graph H. Consider a random permutation on the edges of the complete graph on n vertices. Start from the empty graph on n vertices and consider the edges in the permutation one at a time. An edge is added if it does not form a copy of H with the previously traversed edges, otherwise it is discarded.

The difference between this definition and Definition 3.7 for the random H-free graph process is that in this case an edge is discarded if it forms a copy of H with the previously traversed edges, instead of the previously inserted edges. This implies that for any permutation the H-elimination process gives a subgraph of the H-free process. Similarly as in the previous section we can construct our process by assigning birth times to the edges.

Definition 3.12 (H-elimination random graph process)
Fix a finite graph H. Consider the complete graph on n vertices and assign to each of the possible $\binom{n}{2}$ edges a birthtime uniformly distributed in $[0, 1)$. Now start out from an empty graph on n vertices and increase p gradually. Every time a new edge is born consider adding this edge to the graph. In case adding this edge forms a copy of H with the previously born edges then discard the edge, otherwise insert it into the graph. Edges with equal birthtime (which occur with probability zero) are considered in arbitrary order.

Note that choosing the edge uniformly at random from $[0, 1]$ or from $[0, 1)$ is equivalent as with probability 1 none of the edges was assigned birthtime 1. The graph created after every edge with assigned value less then p has been considered is denoted by $G_e(H)_{n,p}$.

3.4 H-removal process

A third option to create an H-free graph is the H-removal process, which similarly to the previous process starts out from the complete graph on n vertices and in every step a copy of H is selected uniformly at random and its edges are removed from the graph. The process stops once no more copies of H exist.

The number of edges present in this process has been investigated when H is a triangle. Recently Bohman, Frieze and Lubetzky [5] have shown that the random triangle-removal process terminates a.a.s. with $n^{3/2+o(1)}$ edges, improving previous results by Spencer [43], Rödl and Thoma [36] and Grable [22].

Chapter 4

Previous Results

4.1 H-free process

The first question concerning the H-free random graph process is the number of edges contained in the process. The first results relied on comparing the H-free random graph process to $G_{n,p}$. These results already managed to approximate the number of edges in the process up to a logarithmic factor. In many cases these are still the best known bounds. The first half of this section discusses the ideas behind these approximations. In many cases applying the differential equation method improves these bounds. In some cases it gives lower and upper bounds which differ only by a constant factor and for a special class of graphs it gives even tighter concentration.

4.1.1 Bounds using approximations

Erdős, Suen and Winkler [17] observed that Definition 3.8 implies that the graph created from $G_{n,p}$ by removing every edge contained in a copy of H is a subgraph of $G_f(H)_{n,p}$ and thus also of $G_f(H)_n$. We will denote the graph created from $G_{n,p}$ by removing every copy of H with $G(H)_{n,p}$. This observation was later used by Osthus and Taraz [31] and Bollobás and Riordan [9] to give lower and upper bounds for special cases of the H-free random graph process. We will give the results in Osthus and Taraz [31] as these are more general and their upper bounds are more accurate.

Theorem 4.1 ([31])

Let H be a balanced graph then a.a.s.

$$e(G_f(H)_n) = \Omega\left(n^{2-1/m(H)}\right).$$

Note that this implies a lower bound for every 2-balanced graph H.

Theorem 4.2 ([31])

Let H be a strictly 2-balanced graph then a.a.s.

$$\Delta(G_f(H)_n) = O\left(n^{1-1/d_2(H)} \log^{1/(\Delta(H)-1)}\right).$$

The proofs analyse $G(H)_{n,p}$ where p is chosen in such a way that the expected number of copies of H in $G_{n,p}$ is a small fraction of the expected number of edges found in $G_{n,p}$ i.e. $p = cn^{-1/m(H)}$ where c is a suitably chosen constant. Based on concentration results on the number of edges and the number of copies of a balanced graph in $G_{n,p}$ one can deduce $e(G(H)_{n,p}) = \Omega(n^{2-1/d_2(H)})$.

For the upper bound consider x a vertex with maximal degree in H and denote the neighbourhood of x with $N(x)$. Now consider the graph H' which is created from H by removing x. Note that for any vertex $v \in V(G_f(H)_n)$ and any $S \subseteq N(v)$ with $|S| = \Delta(H)$ there can be no copy of $(N(x), H')$ rooted at S in $G_f(H)_n$, except when this copy contains v, as any such copy would imply that a copy of H is also present in $G_f(H)_n$.

Osthus and Taraz [31] show that any set of $Cn^{1-1/d_2(H)} \log^{1/(\Delta(H)-1)} n$ vertices in $V(G(H)_{n,p})$ a.a.s. contains a subset S such that $|S| = \Delta(H)$ and a copy of $(N(x), H')$ is rooted at S in $G(H)_{n,p}$. This clearly implies the upper bound on the degree of every vertex and thus also on $e(G_f(H)_n)$.

In many cases these are still the best known results on the number of edges in the H-free process.

Instead of removing every edge found in a copy of H Wolfovitz [49] calculated, for strictly 2-balanced graphs, the probability that an edge is actually removed. More precisely he gave an upper bound on the probability that assuming $f \in E(G_{n,p})$ we have that $f \notin E(G_f(H)_{n,p})$ when $p = c(\log n)^{1/8e(H)} n^{-1/d_2(H)}$ for a suitably chosen constant c. This improved the lower bounds on $e(G_f(H)_n)$, although only in expectation.

Theorem 4.3

Let H be a strictly 2-balanced graph then

$$\mathbb{E}(e(G_f(H)_n)) = \Omega(n^{2-1/d_2(H)}(\log \log n)^{1/(e(H)-1)}).$$

Warnke [47] showed the following connection between $G_f(H)_{n,m}$ and $G_{n,m}$ for decreasing graph properties:

Lemma 4.4 ([47])

Suppose $\mathcal{Q}$ is a decreasing graph property and that $\lambda = \lambda(n) \geq 2$ is a parameter. Then for every $1 \leq m \leq \binom{n}{2}/\lambda$, setting $M = m\lambda$, we have

$$\mathbb{P}[G_f(H)_{n,m} \notin \mathcal{Q} \text{ and } |O(m)| \geq n^2/\lambda] \leq \mathbb{P}[G_{n,M} \notin \mathcal{Q}] + e^{-m/4}.$$

Further using the connection between $G_{n,m}$ and $G_{n,p}$ Warnke [47] showed that

Theorem 4.5 ([47])

Let $m = m(n) = \omega(1)$ and $p = m/\binom{n}{2}$. Suppose that a.a.s. $|O(i)| > \binom{n}{2}/\lambda$ for some $\lambda = \lambda(m)$ and that $\mathcal{Q}$ is a decreasing graph property. Then for the H-free process we have

$$\mathbb{P}[G_f(H)_{n,m} \notin \mathcal{Q}] \leq \mathbb{P}[G_{n,p\lambda} \notin \mathcal{Q}] + o(1).$$

Later we will use a slightly modified version of this Theorem namely

Theorem 4.6

Let $m = m(n) = \omega(\log n)$ and $p = m/\binom{n}{2}$. Then for any decreasing graph property $\mathcal{Q}$ we have that:

$$\mathbb{P}[G_f(H)_{n,m} \notin \mathcal{Q} \text{ and } |O(m)| \geq n^2/\lambda] \leq \mathbb{P}[G_{n,p\lambda} \notin \mathcal{Q}] + o(n^{-2}).$$

4.1.1.1 Hamiltonicity

One can also answer the question whether there is a Hamiltonian cycle in $G_f(H)_n$ using similar methods. For this we need the following statements on random graphs:

Theorem 4.7 ([7])

The threshold function for the appearance of a subgraph S in $G_{n,p}$ is $p = n^{-1/m(S)}$.

Theorem 4.8 ([35])

The threshold function for the appearance of a Hamiltonian cycle in $G_{n,p}$ is $p = n^{-1}\log n$.

When $d(H) > 1$ the Hamiltonicity of the H-free process follows from the fact that $G_{n,p} = G(H)_{n,p} = G_f(H)_{n,p}$ as long as no copy of H is present in $G_{n,p}$. Since this holds a.a.s. when $p = o(n^{-1/d(H)})$ and because $d(H) > 1$ this holds at the time when the Hamiltonian cycle is formed in $G_{n,p}$. Therefore there is also a Hamiltonian cycle in $G_f(H)_n$.

The case when H is a cycle was solved by Allen [2] and Osthus [30] using the following result of Sudakov and Vu [44]:

Theorem 4.9 ([44])

For every fixed $\varepsilon > 0$ and $p \geq \log^4 n/n$ the random graph $G_{n,p}$ a.a.s. has the following property. If G' is a subgraph of $G_{n,p}$ with maximum degree $\Delta(G') \leq (1-\varepsilon)np$ then $G - G'$ contains a Hamiltonian cycle.

Theorem 4.10

Let H be a cycle then a.a.s. $G_f(H)_n$ is Hamiltonian.

PROOF The difference compared to the previous case is that a small number of copies of H are already present in $G_{n,p}$ when the Hamiltonian cycle is formed at $p = \log n/n$. Select $\hat{p} = \log^4 n/n$. Next we will show that the graph $G' = G_{n,\hat{p}} - G(H)_{n,\hat{p}}$ a.a.s. has maximum degree 2. Note that once we have shown this the result follows from Theorem 4.9.

In order to show that the maximum degree of G' is 2 we will show that $G_{n,\hat{p}}$ does not contain two vertex overlapping copies of H. Let F be a graph created from two vertex overlapping copies of H. Note that the density of F is larger than 1 as $\delta(F) \geq 2$ and $\Delta(F) > 2$ thus $d(F) \geq ((v_F - 1)\delta(F) + \Delta(F))/(2v_F) > 1$. Therefore according to Theorem 4.7 a.a.s. no copy of such a graph is present in

$G_{n,\hat{p}}$. The result follows as there are only finite number of ways two copies of H can overlap. ∎

4.1.2 Bounds using the Differential Equation Method

Applying the differential equation method to the H-free random graph process brought the next breakthrough. The differential equation method was developed by Wormald [52]. The main idea is to identify a collection of random variables whose one step expected changes can be approximated in terms of the random variables in the collection. These expressions yield an autonomous system of ordinary differential equations, and in case some additional conditions are met then the random variables in our collection are tightly concentrated around the trajectory given by the solution of the o.d.e.. There are several different sets of conditions which when fulfilled result in tight concentration of the random variables for example Theorem 5.1 in Wormald [52] or Lemma A.1 in Warnke [46] which is an improved version of Lemma 7.3 in Bohman and Keevash [6].

Next we present the version by Warnke [46]. The notation $\pm$ is used in two ways. First for $b \geq 0$ we denote by $a \pm b$ the closed interval $[a-b, a+b]$. For brevity we use the notation $x = a \pm b$ for $x \in [a-b, a+b]$ and $x \pm y = a \pm b$ for $[x-y, x+y] \subseteq [a-b, a+b]$. The second is the notation $\pm_1$. Every expression containing this symbol is an abbreviation for two different statements the first is where every $\pm_1$ is replaced by $+$ and the second is when it is replaced by $-$. For example $x^{\pm_1} = a^{\pm_1} \pm b$ holds if both $x^+ = a^+ \pm b$ and $x^- = a^- \pm b$ hold. For the sake of brevity we will omit the labels of the $\pm$ sign whenever there is no danger of confusion.

In case all of the conditions of the following Lemma are satisfied for every value of $i \leq m$ then we have bounds on the random variables $X_\sigma(i)$ which hold a.a.s. for every $i \leq m$.

Lemma 4.11 ('Differential Equation Method'[46])

Suppose that $m = m(n)$ and $s = s(n)$ are positive parameters. Let $\mathcal{V} = \mathcal{V}(n)$ be a set, and $\{\mathcal{I}_j\}_{j \in \mathcal{V}}$ be a family of sets, where $\mathcal{I}_j = \mathcal{I}_j(n)$. For every $0 \leq i \leq m$ set $t = t(i) := i/s$. Suppose we have a filtration $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$ and random variables $X_\sigma(i)$ and $Y_\sigma^{\pm}(i)$ which satisfy the following conditions. Assume that for all $j \in \mathcal{V}$ and $\sigma \in \mathcal{I}_j$ the random variables $X_\sigma(i)$ are non-negative and $\mathcal{F}_i$-measurable for all $0 \leq i \leq m$, and that for all $0 \leq i < m$ the random variables $Y_\sigma^{\pm}(i)$ are non-negative, $\mathcal{F}_{i+1}$-measurable and satisfy

$$X_\sigma(i+1) - X_\sigma(i) = Y_\sigma^+(i) - Y_\sigma^-(i).$$

In addition, suppose that for each $j \in \mathcal{V}$ and $\sigma \in \mathcal{I}_j$ we have positive parameters $u_\sigma = u_\sigma(n)$, $\lambda_\sigma = \lambda_\sigma(n)$, $\beta_\sigma = \beta_\sigma(n)$, $\tau_\sigma = \tau_\sigma(n)$, $s_\sigma = s_\sigma(n)$ and $S_\sigma = S_\sigma(n)$, as well as functions $x_\sigma(t)$ and $f_\sigma(t)$ that are smooth and non-negative for $t \geq 0$. For all $0 \leq i^* \leq m$, let $\mathcal{G}_{i^*}$ denote the event that for all $0 \leq i \leq i^*$, $j \in \mathcal{V}$ and $\sigma \in \mathcal{I}_j$, we have

$$X_\sigma(i) = \left(x_\sigma(t) \pm \frac{f_\sigma(t)}{s_\sigma}\right) S_\sigma.$$

Moreover, assume that we have an event $\mathcal{H}_i \in \mathcal{F}_i$ for all $0 \le i \le m$ with $\mathcal{H}_{i+1} \subseteq \mathcal{H}_i$ for all $0 \le i < m$. Finally, suppose that for n large enough the following conditions hold:

1. *(Trend hypothesis) For all $0 \le i < m$, $j \in \mathcal{V}$ and $\sigma \in \mathcal{I}_j$, whenever $\mathcal{G}_i \cap \mathcal{H}_i$ holds we have*
$$\mathbb{E}\big[Y_\sigma^{\pm_1}(i) \mid \mathcal{F}_i\big] = \left(y_\sigma^{\pm_1}(t) \pm \frac{h_\sigma(t)}{s_\sigma}\right)\frac{S_\sigma}{s},$$
where $y_\sigma^{\pm}(t)$ and $h_\sigma(t)$ are smooth non-negative functions such that
$$x'_\sigma(t) = y_\sigma^{+}(t) - y_\sigma^{-}(t) \qquad \text{and} \qquad f_\sigma(t) \ge 2\int_0^t h_\sigma(\tau)\, d\tau + \beta_\sigma.$$

2. *(Boundedness hypothesis) For all $0 \le i < m$, $j \in \mathcal{V}$ and $\sigma \in \mathcal{I}_j$, whenever $\mathcal{G}_i \cap \mathcal{H}_i$ holds we have*
$$Y_\sigma^{\pm}(i) \le \frac{\beta_\sigma^2}{s_\sigma^2 \lambda_\sigma \tau_\sigma} \cdot \frac{S_\sigma}{u_\sigma}.$$

3. *(Initial conditions) For all $j \in \mathcal{V}$ and $\sigma \in \mathcal{I}_j$ we have*
$$X_\sigma(0) = \left(x_\sigma(0) \pm \frac{\beta_\sigma}{3 s_\sigma}\right) S_\sigma.$$

4. *(Bounded number of variables) For all $j \in \mathcal{V}$ and $\sigma \in \mathcal{I}_j$ we have*
$$\max\{|\mathcal{V}|, |\mathcal{I}_j|\} \le e^{u_\sigma}.$$

5. *(High probability event) The event $\mathcal{H}_i$ satisfies*
$$\mathbb{P}[\exists i \le m :\ \mathcal{G}_i \cap \neg\mathcal{H}_i] = o(1).$$

6. *(Additional technical assumptions) For all $j \in \mathcal{V}$ and $\sigma \in \mathcal{I}_j$ we have $u_\sigma = \omega(1)$ as well as*
$$s \ge \max\left\{15 u_\sigma \tau_\sigma (s_\sigma \lambda_\sigma/\beta_\sigma)^2, 9 s_\sigma \lambda_\sigma/\beta_\sigma\right\}, \quad s/(18 s_\sigma \lambda_\sigma/\beta_\sigma) < m \le s \cdot \tau_\sigma/1944,$$
$$\sup_{0 \le t \le m/s} y_\sigma^{\pm}(t) \le \lambda_\sigma, \qquad \int_0^{m/s} |x''_\sigma(t)|\, dt \le \lambda_\sigma,$$
$$h_\sigma(0) \le s_\sigma \lambda_\sigma \qquad \text{and} \qquad \int_0^{m/s} |h'_\sigma(t)|\, dt \le s_\sigma \lambda_\sigma.$$

Then $\mathcal{G}_m \cap \mathcal{H}_m$ holds with high probability.

The fact that the $x_\sigma(t)$ are the solutions of a system of differential equations is implicitly contained in the trend hypothesis. In order to verify the trend hypothesis we need that $\mathbb{E}(Y_\sigma^{\pm_1}(i)|\mathcal{F}_i)$ can be expressed using the random variables $X_\sigma(i)$. We also have that $X_\sigma(i)$ is concentrated around $x_\sigma(t)S_\sigma$. Removing all of the error terms leaves us with a system of differential equations. In general this system of differential equations is determined, usually with the help of heuristics, before attempting to verify the conditions of Lemma 4.11. In this section we will show the heuristics used in determining these equations for several H-free processes.

When analysing the H-free process using the differential equation method Definition 3.6 will be used. One of the key observations behind the analysis of the H-free process is that until one has roughly $n^{2-\varepsilon}$ open edges for some small $\varepsilon > 0$, depending only on H, the H-free random graph process resembles the Erdős-Rényi random graph process.

4.1.2.1 Star-free processes

In the special case when H is a star the H-free process is equivalent to the degree bounded random graph process, or d-process where $d = \Delta(H) - 1$. In the d-process, one starts out from the empty graph on n vertices and in every step selects uniformly at random a pair of non-adjacent vertices which have degree less then d and adds this edge to the graph. Obviously the process contains at most $\lfloor dn/2 \rfloor$ edges. Wormald [52], extending an argument of Ruciński and Wormald [38], has shown that one can track the random variables $D_0(i), \ldots, D_d(i)$ accurately, where $D_j(i)$ is the number of vertices of degree j in $G_f(H)_{n,i}$ when $i \leq d(n/2 - n^{1-\varepsilon})$ for some $\varepsilon > 0$. Note that at this stage $|O(d(n - n^{1-\varepsilon}))| = \Omega(n^{2-2\varepsilon})$ as we have $\Omega(n^{1-\varepsilon})$ vertices with degree less then d.

Note that

$$Q(i) = \frac{1}{2}\left(\sum_{k=0}^{d-1} D_k(i)\right)^2 + O(n) = \frac{1}{2}(n - D_d(i))^2 + O(n)$$

where the $O(n)$ term follows from the edges already present between the vertices with degree less then d. Let $D_j^+(i)$ be the number of vertices v which had degree $j-1$ in $G_f(H)_{n,i}$, but have degree j in $G_f(H)_{n,i+1}$. Similarly let $D_j^-(i)$ be the number of vertices v which had degree j in $G_f(H)_{n,i}$, but have degree $j+1$ in $G_f(H)_{n,i+1}$. Then we have that:

$$\mathbb{E}(D_j^+(i)) = \frac{D_{j-1}(i) \sum_{k=0}^{d-1} D_k(i) + O(n)}{Q(i)} = \frac{2D_{j-1}(i)}{n - D_d(i)} + O(n^{-1})$$

for $j > 0$ as $Q(i) = \Omega(n^{2-2\varepsilon})$. Since inserting an edge can increase the degree of any vertex by at most one we have that $D_j^+(i) = D_{j-1}^-(i)$ and thus:

$$\mathbb{E}(D_j^-(i)) = \frac{2D_j(i)}{n - D_d(i)} + O(n^{-1})$$

for $j < d$. Clearly $D_0(0) = n$ and $D_j(0) = 0$ for $j > 0$.

In the following we will show the heuristics involved in determining the differential equations needed for the proof. Let $t = i/n$ and $D_j(i) \approx z_j(t(i))n$ where $\approx$ means that we approximate the value of $D_j(i)$ with $z_j(t(i))n$. Then we can estimate the derivative of z_0 as follows:

$$\begin{aligned}
z_0'(t) &= z_0'(t(i))(t(i+1)-t(i))n \\
&\approx (z_0(t(i+1)) - z_0(t(i)))n \\
&\approx (D_0(i+1)) - (D_0(i)) \\
&\approx D_0^+(i) - D_0^-(i) \\
&\approx \mathbb{E}(D_0^+(i)) - \mathbb{E}(D_0^-(i)) \\
&\approx -\frac{2D_0(i)}{n - D_d(i)} \\
&\approx -\frac{2z_0(t(i))n}{n - z_d(t(i))n} \\
&= -\frac{2z_0(t(i))}{1 - z_d(t(i))}.
\end{aligned}$$

A similar argument for the remaining variables gives us the following differential equations and initial conditions:

$$\begin{aligned}
z_0'(t) &= -\frac{2z_0(t)}{1 - z_d(t)} & z_0(0) &= 1 \\
z_j'(t) &= \frac{2z_{j-1}(t) - 2z_j(t)}{1 - z_d(t)} & z_j(0) &= 0 \; for \; 0 < j < d \\
z_d'(t) &= \frac{2z_{d-1}(t)}{1 - z_d(t)} & z_d(0) &= 0.
\end{aligned}$$

The simple way to find the solution for this system of differential equations is by exploiting the connection between $G_f(H)_{n,i}$ and $G_{n,i}$, and thus the connection between $G_f(H)_{n,i}$ and $G_{n,p}$. Note that $i = nt \approx \binom{n}{2}p$ thus $p \approx 2t/n$. When $p = O(n^{-1})$ then the degree of every vertex can be approximated with independent Poisson random variables with parameter $\lambda = np$ thus the expected number of vertices with degree j is $n(np)^j e^{-np}/j!$. This suggests the following solutions for the differential equations:

$$\begin{aligned}
z_j(t) &= \frac{(2t)^j e^{-2t}}{j!} \text{ for } 0 \le j < d \\
z_d(t) &= 1 - \sum_{k=0}^{d-1} z_k(t)
\end{aligned}$$

and one can quickly check that these solutions satisfy the differential equations. Wormald [52] has shown that a.a.s. the random variables remain close to the solutions given by these differential equations.

Definition 4.1

Let $\mathcal{G}_m$ be the event that

$$D_k(i) = nz(i/n) + o(n)$$

holds for all $0 \le k \le d$ and $0 \le i < m$.

Theorem 4.12 ([52])

Let H be a star. Then there exists an $\varepsilon > 0$ depending only on H such that $\mathcal{G}_m$ holds a.a.s. for $m = n(d/2 - \varepsilon)$.

Wormald [52] actually states that this can be extended until $d(n/2 - n^{1-\varepsilon})$ edges have been inserted into the process for some small ε. Ruciński and Wormald [38] showed that after this point the minimum degree in the process a.a.s. increases and soon after the minimum degree reaches $d-1$ a.a.s. all the vertices with degree $d-1$ form an independent set. Thus the process can continue until every vertex has degree d and the graph contains $\lfloor dn/2 \rfloor$ edges.

Ruciński and Wormald [39] also show that when the $d \geq 3$ the degree bounded random graph process is a.a.s. connected. However Telcs, Wormald and Zhou [45] showed that this does not hold in the case when $d = 2$, as in this case the degree bounded process is a.a.s. not connected.

4.1.2.2 Strictly 2-balanced graphs

However when H is not a star this simplification can not be applied anymore. Bohman and Keevash [6] in order to track the number of open pairs, also needed to track the formation of rooted graphs. Let (R, Γ) be a rooted graph where R forms an independent set and J is a spanning subgraph of Γ. Denote by $\Xi_{\phi,J,\Gamma}(i)$ the set of injective functions $f : V(\Gamma) \to V(G_f(H)_{n,i})$ such that $f|_R = \phi$, $f(E(J)) \subseteq E(G_f(H)_{n,i})$ and $f(E(\Gamma)\backslash E(J)) \subseteq O(i)$. Note that $\Xi_{\phi,J,\Gamma}(i)$ describes the copies of the rooted graph (R, J) rooted at $\phi(R)$ such that the edges needed to extend this copy into a copy of (R, Γ) can still be inserted. Let $X_{\phi,J,\Gamma}(i) = |\Xi_{\phi,J,\Gamma}(i)|$.

In addition to the number of open pairs Bohman and Keevash [6] track the random variables $X_{\phi,J,\Gamma}(i)$ if one of the following two conditions holds:

- a) $m(R, \Gamma) < d_2(H)$ and Γ does not contain H as a subgraph
- b) $m(R, \Gamma) = d_2(H)$, $E(J) \subsetneq E(\Gamma)$ and the graph Γ' created from Γ by adding all the edges $\{a, b\} \in R \times R$ such that $\phi(a)\phi(b) \in E(G_f(H)_{n,i})$ does not contain a copy of H.

We will refer to these variables as the tracked variables. Let $t(i) = i/n^{2-1/d_2(H)}$ and similarly as before we would like to show that these random variables are a.a.s. concentrated around some functions for the beginning of the process namely for the first $\mu n^{2-1/d_2(H)} \log^{1/(e(H)-1)} n$ steps, where μ is a constant depending on H. Our aim is to approximate the random variables as follows:

$$X_{\phi,J,\Gamma}(i) \approx x_{R,J,\Gamma}(t(i)) n^{f_{d_2(H)}(R,J)}$$
$$Q(i) \approx q(t(i)) n^2.$$

Recall that $f_{d_2(H)}(R, J) = v(J) - |R| - e(J)/2$. Note that the expected number of copies of (R, J) rooted at $S \subseteq V(G_{n,p})$, with $|R| = |S|$, when $p = n^{-1/d_2(H)+o(1)}$ is $n^{f_{d_2(H)}(R,J)+o(1)}$. Since we expect the H-free random graph process to create a graph which resembles $G_{n,p}$ when $p = n^{-1/d_2(H)+o(1)}$ we choose this as our scaling factor.

Note that $Q(i) = X_{\phi_0,\overline{e},e}(i)/2$ where $\phi_0 : \emptyset \to V(G_f(H)_n)$ and e and $\overline{e}$ are graphs on 2 vertices with the edge present and absent, respectively.

Similarly as before we will use differential equations to determine the values of these functions based on the expected change in the variables.

Denote by $C_{u,v}(i)$ the number of pairs in $O(i)$ which would become closed if the edge $\{u, v\}$ were inserted in step i. We will use the following heuristic:

$$\begin{aligned} C_{u,v}(i) &\approx \sum_{\substack{r,e\in E(H)\\ r\neq e}} \sum_{\phi: r\to\{u,v\}} X_{\phi,H_{\{r,e\}},H_r}(i)/\mathrm{aut}(H) \\ &\approx \frac{2}{\mathrm{aut}(H)} \sum_{\substack{r,e\in E(H)\\ r\neq e}} x_{r,H_{\{r,e\}},H_r}(t) n^{1/d_2(H)}. \end{aligned}$$

Note that this approximation assumes that $C_{u,v}(i)$ is concentrated around the same value for every open pair $\{u, v\}$. Define

$$c(t) = \frac{2}{\mathrm{aut}(H)} \sum_{\substack{r,e\in E(H)\\ r\neq e}} x_{r,H_{\{r,e\}},H_r}(t)$$

and let $Q(i)^+$ be the size of $O(i+1)\backslash O(i)$ and $Q(i)^-$ be the size of $O(i)\backslash O(i+1)$. Define $X^+_{\phi,J,\Gamma}(i)$ and $X^-_{\phi,J,\Gamma}(i)$ similarly. Then we have that:

$$\begin{aligned} \mathbb{E}(Q(i)^+) &= 0 \\ \mathbb{E}(Q(i)^-) &\approx \frac{1}{Q(i)} \sum_{\{u,v\}\in O(i)} C_{u,v}(i) \approx c(t) n^{1/d_2(H)} \\ \mathbb{E}(X^+_{\phi,J,\Gamma}(i)) &\approx \frac{1}{Q(i)} \sum_{e\in E(J)} X_{\phi,J_e,\Gamma}(i) \approx \sum_{e\in E(J)} \frac{1}{q(t)n^2} x_{R,J_e,\Gamma}(t) n^{f_{d_2(H)}(R,J_e)} \\ &\approx \sum_{e\in E(J)} \frac{1}{q(t)n^{2-1/d_2(H)}} x_{R,J_e,\Gamma}(t) n^{f_{d_2(H)}(R,J)} \\ \mathbb{E}(X^-_{\phi,J,\Gamma}(i)) &\approx \frac{1}{Q(i)} \sum_{f\in\Xi_{\phi,J,\Gamma}(i)} \sum_{e\in E(\Gamma)\backslash E(J)} C_{f(e)}(i) \\ &\approx \frac{1}{q(t)n^{2-1/d_2(H)}} (e(\Gamma) - e(J)) c(t) x_{R,J,\Gamma}(t) n^{f_{d_2(H)}(R,J)}. \end{aligned}$$

Recall that $t(i+1) - t(i) = n^{-2+1/d_2(H)}$ and thus we have the following differential equations:

$$\begin{aligned} q'(t) &= -c(t) \\ q(t)x'_{R,J,\Gamma}(t) &= \sum_{e\in E(J)} x_{R,J_e,\Gamma}(t) - (e(\Gamma) - e(J)) c(t) x_{R,J,\Gamma}(t) \\ c(t) &= \frac{2}{\mathrm{aut}(H)} \sum_{\substack{r,e\in E(H)\\ r\neq e}} x_{r,H_{\{r,e\}},H_r}(t). \end{aligned}$$

Clearly $x_{R,J,\Gamma}(0) = 1$ if J is the empty graph and $x_{R,J,\Gamma}(0) = 0$ otherwise. Also $q(0) = 1/2$.

Similarly to the star-free random graph process the similarities to $G_{n,p}$ can be exploited, this time with $p \approx 2tn^{-1/d_2(H)}$. Obviously the edges are distributed independently and we will assume that the open pairs are also distributed independently. This suggests the following solution:

$$x_{r,J,\Gamma}(t) = (2t)^{e(J)}(2q(t))^{e(\Gamma)-e(J)}$$

it follows that

$$c(t) = \frac{2e(H)(e(H)-1)}{\mathrm{aut}(H)} 2^{e(H)-1} t^{e(H)-2} q(t)$$

and thus

$$2q(t) = \exp\left(\int_0^t -c(\tau)d\tau\right) = \exp\left(-2^{e(H)}\frac{e(H)}{\mathrm{aut}(H)} t^{e(H)-1}\right).$$

Fix constants μ, ε, V and W satisfying $0 < \mu \ll \varepsilon \ll 1/W \ll 1/V \ll 1/e_H$. The notation $0 \leq \alpha \ll \beta$ means that there is an increasing function $f(x)$ such that the following hold for every $0 < \alpha < f(\beta)$:

- $e(\mu \log^{1/(e(H)-1)} n) \leq n^{\varepsilon}$
- $q(\mu \log^{1/(e(H)-1)} n)^{-V} \leq n^{\varepsilon}$

where $e(t) = e^{P(t)} - 1$ with $P(t) = W(t^{e(H)-1} + t)$. Also define $s_e = n^{1/(2e(H))-\varepsilon}$ and $m = \mu n^{2-1/d_2(H)} \log^{1/(e(H)-1)} n$.

Definition 4.2

Let $\mathcal{G}_m$ be the event that for every tracked variable and every $i \leq m$ the following holds:

$$X_{\phi,J,\Gamma}(i) = (1 \pm e(t)/s_e)(x_{R,J,\Gamma}(t(i)) \pm 1/s_e)n^{f_{d_2(H)}(R,J)}.$$

Theorem 4.13 ([6])

For strictly 2-balanced graphs $\mathcal{G}_m$ holds a.a.s..

Theorem 4.14 ([6])

For strictly two balanced H we have that $G_f(H)_n$ contains a.a.s.

$$\Omega(n^{2-1/d_2(H)} \log^{1/(e(H)-1)} n)$$

edges.

PROOF Since $\mathcal{G}_m$ holds a.a.s. we have that

$$Q(i) = \left(1 \pm \frac{e(t)}{s_e}\right)\left(q(t(i)) \pm \frac{1}{s_e}\right)n^2.$$

Note that $e(t(m))/s_e = o(1)$ and that $s_e = o(q(t(m)))$. We also have that $q(t(m)) = \Omega(n^{-\varepsilon})$ thus $Q(i) = \Omega(q(t(m)n^2)) = \Omega(1)$. The statement follows from the fact that the H-free process finishes only when there are no more open pairs. ∎

Note that in order to set up the differential equations it would be enough to consider the random variables $X_{\phi,J,H_r}(i)$ for every $J \subsetneq H_r$ and $r \in E(H)$ as these variables determine $C_{u,v}(i)$.

Before showing how these results can be extended to determine upper bounds on the size of the maximal independent set and on the maximum degree we first show how this result was modified by Picollelli [32] to show a lower bound on the diamond-free process.

4.1.2.3 The Diamond

The diamond or K_4^- is a 2-balanced graph however it is not strictly 2-balanced as $K_3 \subsetneq K_4^-$ and $d_2(K_3) = d_2(K_4^-) = 2$. Similarly to the previous cases the aim is to track $Q(i)$ for which we need to determine $\mathbb{E}(Q(i)^-)$. However unlike the strictly 2-balanced case $C_{u,v}(i)$ is not concentrated around one value for every $\{u,v\} \in O(i)$. This is due to the fact that there is a large difference in the value of $C_{u,v}(i)$ based on whether the codegree of $\{u,v\} \in O(i)$ is 0 or 1. Note that when the codegree of $\{u,v\}$ is larger than one then it is already closed.

Partition $O(i)$ into two parts $O_0(i)$ and $O_1(i)$ where $O_0(i)$ is the set of open pairs which have codegree zero and $O_1(i)$ is the set of open pairs which have codegree 1. Also let $Q_0(i) = |O_0(i)|$ and $Q_1(i) = |O_1(i)|$. Note that this is similar to what was done in the star free process where the vertices with degree less than d were partitioned into classes with degree exactly k for every $k < d$. The edges are also partitioned into two groups namely the edges not contained in a triangle denoted by $E_0(i)$ and the edges contained in a triangle denoted by $E_1(i)$. We are interested in the sets $\Xi_{\phi,J,\Gamma}(i)$ and the random variables $X_{\phi,J,\Gamma}(i)$ only in the case when $\Gamma = (K_3)_r$ for $r \in E(K_3)$ and $\phi(r) \in O(i)$. Therefore we simplify the notation. For the rest of this section fix $r \in K_3$ and let $\{e_1, e_2\} = E(K_3)\backslash r$. Let $\Xi_{\phi,E(J)}(i) = \Xi_{\phi,J,(K_3)_r}(i)$ and $X_{\phi,E(J)}(i) = X_{\phi,J,(K_3)_r}(i)$. In order to track $Q_0(i)$ and $Q_1(i)$ we need to partition the set $\Xi_{\phi,E(J)}(i)$ according to whether the codegree of $f(e_1)$ and $f(e_2)$ is zero or one. Thus for fixed ϕ and J we are looking at 4 sets $\Xi_{\phi,E(J),D}(i)$ and thus 4 random variables $X_{\phi,E(J),D}(i)$, where $D \subseteq \{e_1, e_2\}$. Formally $\Xi_{\phi,E(J),D}(i)$ is the set of maps $f \in \Xi_{\phi,E(J)}(i)$ which satisfy that $f(e) \in O_1(i) \cup E_1(i)$ iff $e \in D$. Similarly to the strictly 2-balanced case let $t = i/n^{2-1/m_2(H)} = i/n^{3/2}$ and

$$Q_0(i) \approx q_0(t)n^2$$
$$Q_1(i) \approx q_1(t)n^2$$
$$X_{\phi,E(J),D}(i) \approx x_{r,E(J),D}(t)n^{f_2(r,J)}.$$

Similarly to the strictly 2-balanced case the $n^{f_2(r,J)}$ term follows from the fact that we expect our graph to look similar to $G_{n,p}$ when $p = n^{-1/2+o(1)}$. The expected number of copies in (R,J) rooted at S in $G_{n,p}$, when $p = n^{-1/2+o(1)}$ is $n^{f_2(R,J)+o(1)}$. We also assume that each partition of $\Xi_{\phi,E(J)}(i)$ contains a large fraction of the elements.

We modify the definition of $C_{u,v}(i)$ slightly. For $u,v \in O_1(i)$ it still denotes the number of edges which when inserted into the graph would close $\{u,v\}$. However in the case when $\{u,v\} \in O_0(i)$ this is the number of open pairs which when inserted into the graph would remove $\{u,v\}$ from $O_0(i)$, therefore in addition to the edges which when inserted would close $\{u,v\}$ the edges which would increase the codegree of $\{u,v\}$ are also counted. Thus for $\{u,v\} \in O_0(i)$ we have that:

$$\begin{aligned} C_{u,v}(i) &= \sum_{\phi:r\to\{u,v\}} \sum_{D\subseteq\{e_1,e_2\}} X_{\phi,e_1,D}(i) \\ &\approx 2 \sum_{D\subseteq\{e_1,e_2\}} x_{r,e_1,D}(t)\sqrt{n} \\ &= c_0(t)\sqrt{n}. \end{aligned}$$

For $\{u,v\} \in O_1(i)$ let $g(\{u,v\})$ denote the vertex connected to both u and v. Assume that inserting the edge $\{w_1,w_2\}$ closes $\{u,v\}$ and in particular increases the codegree of $\{u,g(\{u,v\})\}$. Then we have that $\{w_1,w_2\} \in O_1(i)$ as inserting this edge created a triangle. Therefore when $\{u,v\} \in O_1(i)$ we have that:

$$\begin{aligned} C_{u,v}(i) = &\sum_{\phi:r\to\{u,v\}} \sum_{D\in\{e_1,e_2\}} X_{\phi,e_1,D}(i) \\ &+ \left(\sum_{\phi:r\to\{u,g(\{u,v\})\}} X_{\phi,e_1,e_2}(i) + \sum_{\phi:r\to\{v,g(\{u,v\})\}} X_{\phi,e_1,e_2}(i) \right) \\ \approx &\left(2 \sum_{D\subseteq\{e_1,e_2\}} x_{r,e_1,D}(t) + 4x_{r,e_1,e_2}(t) \right) \sqrt{n} \\ = &\, c_1(t)\sqrt{n}. \end{aligned}$$

Let $D_{u,v}(i)$ denote the number of edges which when inserted into the graph would increase the codegree of $\{u,v\}$ by one and does not close $\{u,v\}$ in the process.

$$D_{u,v}(i) = \sum_{\phi:r\to\{u,v\}} X_{\phi,e_1,\emptyset}(i) \approx 2x_{r,e_1,\emptyset}(t)\sqrt{n} = d(t)\sqrt{n}.$$

Thus we can express the expected change as:

$$\begin{aligned} \mathbb{E}(Q_0^+(i)) &= 0 \\ \mathbb{E}(Q_0^-(i)) &= \frac{1}{Q(i)} \sum_{\{u,v\}\in O_0(i)} C_{u,v}(i) \approx \frac{q_0(t)c_0(t)}{q(t)}\sqrt{n} \\ \mathbb{E}(Q_1^+(i)) &= \frac{1}{Q(i)} \sum_{\{u,v\}\in Q_0(i)} \sum_{\phi:r\to\{u,v\}} X_{\phi,e_1,\emptyset}(i) \approx 2\frac{q_0(t)}{q(t)} x_{r,e_1,\emptyset}(t)\sqrt{n} \\ \mathbb{E}(Q_1^-(i)) &= \frac{1}{Q(i)} \sum_{\{u,v\}\in O_1(i)} C_{u,v}(i) \approx \frac{q_1(t)c_1(t)}{q(t)}\sqrt{n}. \end{aligned}$$

Also

$$\begin{aligned}
\mathbb{E}(X^+_{\phi,E(J),D}(i)) &= \frac{1}{Q(i)}\left(\sum_{e\in E(J)} X_{\phi,E(J)\backslash e,D}(i) + \sum_{e\in D}\sum_{f\in\Xi_{\phi,E(J),D\backslash e}(i)} D_{f(e)}(i)\right)\\
&\approx \frac{1}{q(t)}\left(\sum_{e\in E(J)} x_{r,E(J)\backslash e,D}(t) + \sum_{e\in D} x_{r,E(J),D\backslash e}(t)d(t)\right)\frac{n^{f_2(R,J)}\sqrt{n}}{n^2}\\
\mathbb{E}(X^-_{\phi,E(J),D}(i)) &= \frac{1}{Q(i)}\sum_{e\in E(\Gamma)\backslash E(J)}\sum_{f\in\Xi_{\phi,E(J),D}(i)} C_{f(e)}(i)\\
&\approx \frac{x_{r,E(J),D}(t)}{q(t)}\left((2-|E(J)\cup D|)c_0(t) + |D\backslash E(J)|c_1(t)\right)\frac{n^{f_2(R,J)}\sqrt{n}}{n^2}.
\end{aligned}$$

This gives us the following system of differential equations:

$$\begin{aligned}
q_0'(t) &= -\frac{q_0(t)c_0(t)}{q(t)}\\
q_1'(t) &= \frac{q_0(t)}{q(t)}d(t) - \frac{q_1(t)c_1(t)}{q(t)}\\
x'_{r,E(J),D}(t) &= \frac{1}{q(t)}\left(\sum_{e\in E(J)} x_{r,E(J)\backslash e,D}(t) + \sum_{e\in D} x_{r,E(J),D\backslash e}(t)d(t)\right)\\
&\quad - \frac{x_{r,E(J),D}(t)}{q(t)}\left((2-|E(J)\cup D|)c_0(t) + |D\backslash E(J)|c_1(t)\right)\\
c_0(t) &= 2\sum_{D\subseteq\{e_1,e_2\}} x_{r,e_1,D}(t)\\
c_1(t) &= \left(2\sum_{D\subseteq\{e_1,e_2\}} x_{r,e_1,D}(t) + 4x_{r,e_1,e_2}(t)\right)\\
d(t) &= 2x_{r,e_1,\emptyset}(t)
\end{aligned}$$

and initial conditions $q_0(0) = 1/2$, $q_1(t) = 0$, $x_{r,E(J),D} = 1$ if $E(J) = D = \emptyset$ and 0 otherwise. Similarly to earlier occasions we solve these differential equations by comparing the diamond free random graph process to $G_{n,p}$ when $p = 2t/\sqrt{n}$. We have that the codegree for any given pair of vertices in $G_{n,p}$ is Poisson distributed with parameter np^2 which would indicate that $q_0(t) = e^{-4t^2}$ and $q_1(t) = 4t^2e^{-4t^2}$. However the approximation for $q_1(t)$ is incorrect as this considers all edges, however for a pair of vertices which have exactly one mutual neighbour to be open, it is required that neither edge involved in the creation of the mutual neighbour is contained in a triangle. One could try expanding this argument by multiplying with an e^{-4t^2} term i.e. the probability that a given pair of vertices have codegree zero, for both edges. However this is unlikely to hold, as we have strongly limited the number of triangles which can appear, due to the fact that every edge can be found in at most one triangle. Recall that by definition $|E(i)| = tn^{3/2}$. Let $E_0(i) \approx r(t)n^{3/2}$ and thus $E_1(i) \approx (t - r(t))n^{3/2}$. Then replacing the $4t^2$ term with $4(r(t))^2$ and thus

$q_1(t) = 4(r(t))^2 e^{-4t^2}$ should be the correct term. Note that $|E_0(i)|$ increases by 1 every time an element in $O_0(i)$ is selected and it decreases by 2 every time an element in $O_1(i)$ is selected. This implies the following differential equation for $r(t)$

$$r'(t) = \frac{q_0(t) - 2q_1(t)}{q(t)} = \frac{q_0(t) - 2q_1(t)}{q_0(t) + q_1(t)} = \frac{1 - 8r(t)^2}{1 + 4r(t)^2}$$

and this differential equation with the initial condition $r(t) = 0$ holds if $r(t)$ satisfies:

$$8t + 4r(t) - 3\sqrt{2} \cdot \operatorname{arctanh}(2\sqrt{2} \cdot r(t)) = 0.$$

If one examines the solutions of the differential equations when H is a strictly 2-balanced graph one multiplies a $2t$ term for every edge present in the graph and a $2q(t)$ term for every edge still needed. Similarly here one multiplies a $2r(t)$ term for every edge not in a triangle, a $2(t - r(t))$ term for every edge in a triangle, a $2q_0(t)$ term for every needed edge with codegree zero and a $2q_1(t)$ term for every needed edge with codegree one. Thus the solutions of the above variables can be expressed as follows:

$$\begin{aligned} q_0(t) &= e^{-4t^2}/2 \\ q_1(t) &= 4(r(t))^2 e^{-4t^2}/2 = 2(r(t)^2)e^{-4t^2} \\ x_{r,E(J),D}(t) &= (2q_0)^{2-|E(J))\cup D|}(2q_1)^{|D\setminus E(J)|}(2r(t))^{|E(J)\setminus D|}(2(t - r(t)))^{|E(J)\cap D|}. \end{aligned}$$

Let K be a sufficiently large constant and define

$$e(t) = e^{K(t^2+t)}, \; e_q(t) = 1 + \int_0^t Ke(\tau)d\tau, \; s_e = n^{1/6}.$$

Let $0 \le \varepsilon \le 1/40$ and let $\mu = \mu(\varepsilon, K) > 0$ be sufficiently small so that $e(\mu\sqrt{\log n}) < n^\varepsilon$ for all sufficiently large n. Let $m = \mu\sqrt{\log n} n^{3/2}$.

Definition 4.3

Let $\mathcal{G}_m$ be the event that the following bounds hold for every $i \le m$, $E(J) \subseteq \{e_1, e_2\}$, $D \subseteq \{e_1, e_2\}$ and $f \in V(G_f(K_4^-)_{n,i}) \times V(G_f(K_4^-)_{n,i})$

$$Q_j(i) = (q_j(t) \pm \frac{e_q(t)}{s_e})n^2$$

$$\sum_{\phi:r\to f} X_{\phi,E(J),D} = \left(2x_{r,E(J),D}(t) \pm \frac{e(t)e^{4t^2(e(J)-2)}}{s_e}\right) n^{f_2(r,J)}$$

Theorem 4.15 ([32])

For the diamond-free random graph process $\mathcal{G}_m$ holds a.a.s..

4.1.2.4 Independent sets

Before discussing the upper bounds on the number of edges we will first show an upper bound on the size of the largest independent set. Similar ideas will be used when establishing an upper bound on the number of edges. Actually an upper bound is shown for $\alpha(G_f(H)_{n,m})$ and since adding edges can only decrease the independence number this is an upper bound on the size of any independent set at the end of the process. Note that in case H is a triangle this actually gives an upper bound on the length of the triangle-free random graph process as in a triangle-free graph the neighbourhood of any vertex forms an independent set.

The main idea behind the proof is to give a lower bound on the number of open pairs in any set of size $3\mu^{-1}(\log n)^{1-1/(e(H)-1)}n^{1/d_2(H)}$. Consider a set I of this size, and let $O_I(i)$ denote the set of open pairs in I at step i. Again the differential equation method is used to track the random variable $|O_I(i)|$. However using the differential equation method for this random variable is not straightforward as inserting certain edges into the graph can cause a large change in this random variable, e.g. when H is a triangle then adding an edge connecting a vertex in I to a vertex with many neighbours in I could close a large number of open pairs in $O_I(i)$. Therefore instead of tracking $|O_I(i)|$ one tracks a random variable $|O'_I(i)|$, which do not take the affect of these large changes into account. Note that $|O'_I(i)| > |O_I(i)|$, however in special cases, e.g. when H is a cycle or when H is a complete graph, one can give an upper bound on $|O'_I(i)| - |O_I(i)|$.

Theorem 4.16 ([6])

Let H be a cycle or a complete graph. Let $Q_I(i)$ denote the number of open pairs in the H-free process at step i in $I \subset V(G_f(H)_{n,i})$, where $|I| = 3\mu^{-1}(\log n)^{1-1/(e(H)-1)}n^{1/d_2(H)}$ then:

$$Q_I(i) = (1 \pm e(t)n^{-2\varepsilon})(2q(t) \pm 2n^{-2\varepsilon})(3\mu^{-1}(\log n)^{1-1/(e(H)-1)}n^{1/d_2(H)})^2/2.$$

Now one can calculate the probability that none of these edges are chosen in the first m steps and applying the union bound gives:

Theorem 4.17 ([6])

Let H be a cycle or a complete graph then every set of size $3\mu^{-1}(\log n)^{1-1/(e(H)-1)}n^{1/d_2(H)}$ contains an edge by time m.

4.1.3 Upper bounds

The idea for finding an upper bound on the number of edges is similar to finding the size of the maximal independent set. However due to technical difficulties, so far the order of magnitude has only been determined for cycles, the K_4 and the diamond. In general one tries to show that after m steps a.a.s. there is a copy of $(N(v), H[V(H)\backslash v])$ for some $v \in V(h)$ rooted in every vertex set of size $Cn^{1-1/d_2(H)}\log^{1/e(H)-1} n$. When H is the K_4 we want to show that there is a triangle in every set of size $Cn^{3/5}\log^{1/5} n$ and in case H is a cycle of length ℓ then we want to show that in any set of size $Cn^{1/(\ell-1)}\log^{1/(\ell-1)} n$ there is a pair of vertices connected by a path of length $\ell - 2$. In case H is the diamond we want to show that there is an edge with codegree one in every set of size $Cn^{1/2}\sqrt{\log n}$.

The main idea is to track the number of open pairs which would complete a copy of such a graph. Similarly to the independent set one ignores the edges which would cause large changes in this random variable and tries to estimate the maximal affect of the ignored edges.

Theorem 4.18 ([46],[51])

We have a.a.s. that

$$\Delta(G_f(K_4)_n) = O(n^{3/5}\sqrt[5]{\log n}).$$

Theorem 4.19 ([47],[34])

We have a.a.s. that

$$\Delta(G_f(C_\ell)_n) = O((n\log n)^{1/(\ell-1)}).$$

Theorem 4.20 ([32])

We have a.a.s. that

$$\Delta(G_f(K_4^-)_n) = O(\sqrt{n\log n}).$$

4.1.4 Ramsey and Turán numbers

As we mentioned in the introduction this graph process is strongly connected to the Ramsey number. More precisely it is connected to the asymptotics of the off-diagonal Ramsey number $R(k,t)$ when k is fixed and t tends to infinity. Note that showing that there exists a K_k-free graph on n vertices which has an independent set of size t implies that $R(k,t) > n$. Therefore it follows from Theorem 4.17 that

Theorem 4.21 ([6])

For fixed $k \geq 3$ we have that

$$R_{k,t} = \Omega\left(t^{\frac{k+1}{2}}(\log(t))^{\frac{1}{k-2}-\frac{k+1}{2}}\right).$$

Currently these are the best known lower bounds.

Recall that the Turán number $\text{ex}(n,H)$ is the maximal number of edges contained in an H-free graph on n vertices. Note that the number of edges in the H-free random graph process give a lower bound for the asymptotics of the Turán number.

Theorem 4.22

For every strictly 2-balanced H we have that

$$\text{ex}(n,H) = \Omega(n^{2-1/d_2(H)}\log^{1/(e(H)-1)} n).$$

In most cases deterministic methods give better lower bounds. For general H the Erdős-Stone theorem [16] gives the exact asymptotics for the Turán number, except when H is bipartite. Deterministic constructions also exist for complete bipartite graphs. It is known that $\text{ex}(n,K_{s,t}) = O(n^{2-1/s})$ [28], [19]. Lower bounds matching this upper bound up to a constant factor are known for the $K_{2,2}$ [15], the $K_{3,3}$ [11] and the $K_{s,t}$ when $t > (s-1)!+1$ [3].

Note that $\text{ex}(n,H') \leq \text{ex}(n,H)$ when $H' \subseteq H$. Thus we have that

$$\Omega(n^{5/3}) = \text{ex}(n,K_{3,3}) \leq \text{ex}(n,K_{3,4}^-) \leq \text{ex}(n,K_{3,4})) = O(n^{5/3}).$$

Still in many cases Theorem 4.22 is the best known lower bound. For example when $H = K_{r,r}$ for $r \geq 5$.

4.2 H-elimination process

Although the H-elimination process has not been studied in detail, many results originally shown for the H-free process still apply. Note that removing every edge in every copy of H from $G_{n,p}$, in addition to creating a subgraph of the H-free process, one also creates a subgraph of the H-elimination process. Therefore the previous results imply a lower bound when H is balanced, namely $e(G_e(H)_{n,1}) = \Omega(n^{2-1/d_2(H)})$. Recall that the H-elimination process creates a subgraph of the H-free process, this implies that every upper bound for the H-free process is also an upper bound for the H-elimination process. Tighter upper bounds can be shown using the following Theorem of Spencer [42].

Fix a rooted graph (R, F) such that R forms an independent set. Let $\mathrm{Ext}(R, F)$ be the event that for every injective function $\phi : R \to V(G_{n,p})$ there exists a copy of (R, F) with respect to ϕ in $G_{n,p}$. Spencer [42] proved thresholds for the event $\mathrm{Ext}(R, F)$.

Theorem 4.23 ([42])

Let (R, F) be a non-trivial strictly balanced rooted graph. Let c_1 be the number of graph automorphisms of F such that the roots are fixed points. Let c_2 be the number of bijections on R which can be extended to an automorphism over F. Let $\lambda > 0$ be arbitrary and fixed. Let $p = p(n)$ satisfy

$$n^{v(F)-|R|}p^{e(F)}/c_1 = \ln(n^{|R|}/(c_2\lambda)).$$

Then

$$\mathbb{P}(\mathrm{Ext}(R, H)) \to e^{-\lambda}.$$

Select $r \in E(H)$ for H strictly 2-balanced. Then inserting any pair of vertices into $G_{n,p}$ when $p = cn^{-1/d_2(H)} \log^{1/(e(H)-1)} n$ a.a.s. creates a copy of H. This implies that there are a.a.s. $O(n^{2-1/d_2(H)}) \log^{1/(e(H)-1)} n$ edges in the H-elimination process, when H is strictly 2-balanced.

A slightly weaker version of this theorem is valid for any graph.

Theorem 4.24 ([42])

Let (R, F) be a rooted graph such that R forms an independent set.

If for every $(R', F') \subseteq (R, F)$ such that $m(R, F) = d(R', F')$ we have that R' is a set of isolated vertices then $p = n^{-1/m(R,F)}$ is a threshold function for $\mathrm{Ext}(R, F)$ in the sense that

$$\text{If } p = o(n^{-1/m(R,F)}) \text{ then } \mathbb{P}(\mathrm{Ext}(R, F)) \to 0$$
$$\text{If } p = \omega(n^{-1/m(R,F)}) \text{ then } \mathbb{P}(\mathrm{Ext}(R, F)) \to 1.$$

Otherwise let $(R', F') \subseteq (R, F)$ be the rooted subgraph with the least number of edges such that $m(R, F) = d(R', F')$ and R' is not a set of isolated vertices. Then $p = n^{-1/m(R,F)} \log^{1/e(F')} n$ is a threshold function for $\mathrm{Ext}(R, F)$ in the sense that there *exists constants* $0 < c < C$ *such that*

$$\text{If } p \leq cn^{-1/m(R,F)} \log^{1/e(F')} n \text{ then } \mathbb{P}(\mathrm{Ext}(R,F)) \to 0$$
$$\text{If } p \geq Cn^{-1/m(R,F)} \log^{1/e(F')} n \text{ then } \mathbb{P}(\mathrm{Ext}(R,F)) \to 1.$$

This clearly gives an upper bound for the H-elimination process in general, after examining (r, H_r) for every $r \in E(H)$. Let $m = \min_{r \in E(H)} m(r, H_r)$. Now let $r_1, r_2, ..r_k$ be the set of edges where this minimum is achieved. If there exists $i \leq k$ such that for every $(r_i, F) \subseteq (r_i, H_{r_i})$ with $d(r_i, F) = m$ we have that r is a set of isolated vertices, then for any $g(n) = \omega(1)$ function the H-elimination process terminates with $O(n^{-1/m(H)})g(n)$ edges. On the other hand if this does not hold, then let s_i be the minimal number of edges found in $(r_i, F) \subseteq (r_i, H_i)$ such that $d(r_i, F) = m$ and r_i is not an isolated set. Let $s = \max s_i$. Then the process terminates a.a.s. with $O(n^{-1/m(H)} \log^{1/s} n)$ edges.

Chapter 5

H-elimination random graph process

In contrast to the H-free process one does not need to use the differential equation method to analyse the H-elimination process. First we will show bounds on the expected number of edges created by the H-elimination process. Afterwards we will show that the number of edges is concentrated around its expectation in case H is strictly 2-balanced. Finally we will examine the independence number and the subgraphs present in the H-elimination process when H is strictly 2-balanced. In this chapter we will be using Definition 3.12, where we assign a birthtime to every edge independently and uniformly from $[0, 1)$ and consider the edges in order of the birthtimes. Let p_e denote the value assigned to edge e and note that edge e is inserted into the graph if no copy of (r, H_r) is rooted at e in G_{n,p_e} for any $r \in E(H)$. Since in most cases the appearance of different copies of (r, H_r) rooted at e are not independent it is vital to study how these copies can overlap.

5.1 Overlapping rooted graphs

In this section we are considering how two rooted graphs can overlap.

Definition 5.1

Let (R_1, F_1) and (R_2, F_2) be two (not necessarily different) nontrivial rooted graphs. Define $(R_1, F_1) \oplus (R_2, F_2)$ as the set of triples (g_1, g_2, F) where F is a graph and $g_i : V(F_i) \to V(F)$ are injective functions such that the following hold:

- $F = (g_1(V(F_1)) \cup g_2(V(F_2)), g_1(E(F_1)) \cup g_2(E(F_2)))$
- $g_1(E(F_1)) \cap g_2(E(F_2)) \neq \emptyset$
- $g_1(R_1, F_1) \neq g_2(R_2, F_2)$

In addition denote the vertex sets $S_{g_1,g_2,F} = g_1(R_1) \cup g_2(R_2)$, $R_{g_1,g_2,F} = g_1(R_1) \cap g_2(R_2)$ and the graph $T_{g_1,g_2,F} = (g_1(V(F_1)) \cap g_2(V(F_2)), g_1(E(F_1)) \cap g_2(E(F_2)))$.

$S_{g_1,g_2,F}$ is the "union" of the roots, $R_{g_1,g_2,F}$ is the "intersection" of the roots, $T_{g_1,g_2,F}$ is the "intersection" of the graphs and F is the "union" of the graphs. When the roots of a rooted graph form an empty set then we have a graph and in this case the rooted graph being strictly balanced is equivalent to the fact that the graph is strictly balanced. It is known that the union of two balanced graphs with the same density is at least as dense as the original graphs, however when both are strictly balanced the density of the union is denser then the original graphs. The following Lemmas generalise these statements for balanced and strictly balanced rooted graphs.

Lemma 5.1

For any two rooted graphs (R_1, F_1), (R_2, F_2) and $(g_1, g_2, F) \in (R_1, F_1) \oplus (R_2, F_2)$ we have for any value of d that

$$f_d(S_{g_1,g_2,F}, F) = f_d(R_1, F_1) + f_d(R_2, F_2) - f_d(R_{g_1,g_2,F}, T_{g_1,g_2,F}).$$

PROOF Applying the inclusion exclusion principle gives that:

$$\begin{aligned} f_d(S_{g_1,g_2,F}, F) &= v(F) - |S_{g_1,g_2,F}| - \frac{e(F)}{d} \\ &= v(F_1) + v(F_2) - v(T_{g_1,g_2,F}) - |R_1| - |R_2| + |R_{g_1,g_2,F}| \\ &- \frac{e(F_1) + e(F_2) - e(T_{g_1,g_2,F})}{d} \\ &= v(F_1) - |R_1| - \frac{e(F_1)}{d} + v(F_2) - |R_2| - \frac{e(F_2)}{d} \\ &- \left(v(T_{g_1,g_2,F}) - |R_{g_1,g_2,F}| - \frac{e(T_{g_1,g_2,F})}{d} \right) \\ &= f_d(R_1, F_1) + f_d(R_2, F_2) - f_d(R_{g_1,g_2,F}, T_{g_1,g_2,F}) \end{aligned}$$

completing the proof. ■

Lemma 5.2

Let (R_1, F_1) be a strictly balanced nontrivial rooted graph and (R_2, F_2) be a rooted graph such that $d = d(R_1, F_1) \leq d(R_2, F_2)$. Then we have for every $(g_1, g_2, F) \in (R_1, F_1) \oplus (R_2, F_2)$ that:

$$f_d(R_{g_1,g_2,F}, T_{g_1,g_2,F}) \geq 0$$

where equality holds only if $(R_{g_1,g_2,F}, T_{g_1,g_2,F}) = g_1(R_1, F_1)$.

PROOF Due to the fact that $R_{g_1,g_2,F} \subseteq g_1(R_1)$ we have that $f_d(R_{g_1,g_2,F}, T_{g_1,g_2,F}) \geq f_d(g_1(R_1) \cap V(T_{g_1,g_2,F}), T_{g_1,g_2,F})$ and equality holds only if $R_{g_1,g_2,F} = g_1(R_1) \cap V(T_{g_1,g_2,F})$. Note that $(g_1(R_1) \cap V(T_{g_1,g_2,F}), T_{g_1,g_2,F}) \subseteq g_1(R_1, F_1)$. Since (R_1, F_1) is a strictly balanced rooted graph we have that $f_d(g_1(R_1) \cap V(T_{g_1,g_2,F}), T_{g_1,g_2,F}) \geq 0$. Proposition 2.3 gives us that equality holds only if $(g_1(R_1) \cap V(T_{g_1,g_2,F}), T_{g_1,g_2,F}) = g_1(R_1, F_1)$. The statement follows. ■

Lemma 5.3

Let (R_1, F_1) be a balanced rooted graph and (R_2, F_2) be a rooted graph such that $d = d(R_1, F_1) \leq d(R_2, F_2)$. Then we have for every $(g_1, g_2, F) \in (R_1, F_1) \oplus (R_2, F_2)$ that:

$$f_d(R_{g_1,g_2,F}, T_{g_1,g_2,F}) \geq 0.$$

PROOF The proof is similar to the strictly balanced case. Since $R_{g_1,g_2,F} \subseteq g_1(R_1)$ thus $f_d(R_{g_1,g_2,F}, T_{g_1,g_2,F}) \geq f_d(g_1(R_1) \cap V(T_{g_1,g_2,F}), T_{g_1,g_2,F})$. Note that $(g_1(R_1) \cap V(T_{g_1,g_2,F}), T_{g_1,g_2,F}) \subseteq g_1(R_1, F_1)$. Since (R_1, F_1) is a balanced rooted graph we have by Proposition 2.4 that $f_d(g_1(R_1) \cap V(T_{g_1,g_2,F}), T_{g_1,g_2,F}) \geq 0$. ∎

Lemma 5.4

Let $(R_1, F_1), (R_2, F_2)$ be nontrivial strictly balanced rooted graphs such that $d(R_1, F_1) = d(R_2, F_2)$. Then we have for any $d \leq d(R_1, F_1)$ and for every $(g_1, g_2, F) \in (R_1, F_1) \oplus (R_2, F_2)$ that:

$$f_d(S_{g_1,g_2,F}, F) < 0.$$

PROOF Obviously $f_d(S_{g_1,g_2,F}, F) \leq f_{d'}(S_{g_1,g_2,F}, F)$ if $d \leq d'$ so we may assume that $d = d(R_1, F_1)$. According to Lemma 5.1 for any $(g_1, g_2, F) \in (R_1, F_1) \oplus (R_2, F_2)$ we have that $f_d(S_{g_1,g_2,F}, F) = f_d(R_1, F_1) + f_d(R_2, F_2) - f_d(R_{g_1,g_2,F}, T_{g_1,g_2,F})$. We have that $f_d(R_1, F_1) = 0$ and $f_d(R_2, F_2) = 0$. Thus it follows that $f_d(S_{g_1,g_2,F}, F) = -f_d(R_{g_1,g_2,F}, T_{g_1,g_2,F})$. Lemma 5.2 gives us that $f_d(R_{g_1,g_2,F}, T_{g_1,g_2,F}) \geq 0$ where equality holds only if $(R_{g_1,g_2,F}, T_{g_1,g_2,F}) = g_1(R_1, F_1)$.

Assume for contradiction that $(R_{g_1,g_2,F}, T_{g_1,g_2,F}) = g_1(R_1, F_1)$. This implies that $g_1(R_1, F_1) \subseteq g_2(R_2, F_2)$. However the case $g_1(R_1, F_1) = g_2(R_2, F_2)$ is ruled out by the definition, therefore $g_1(R_1, F_1) \subsetneq g_2(R_2, F_2)$. However this implies that (R_2, F_2) a nontrivial strictly balanced graph has a rooted subgraph with the same density which is a contradiction with Proposition 2.3. ∎

Lemma 5.5

Let $(R_1, F_1), (R_2, F_2)$ be balanced rooted graphs such that $d(R_1, F_1) = d(R_2, F_2)$. Then we have for any $d \leq d(R_1, F_1)$ and for every $(g_1, g_2, F) \in (R_1, F_1) \oplus (R_2, F_2)$ that:

$$f_d(S_{g_1,g_2,F}, F) \leq 0.$$

PROOF Similarly to the strictly balanced case we may assume that $d = d(R_1, F_1)$. Lemma 5.1 implies $f_d(S_{g_1,g_2,F}, F) = f_d(R_1, F_1) + f_d(R_2, F_2) - f_d(R_{g_1,g_2,F}, T_{g_1,g_2,F})$. Again $f_d(R_1, F_1) = 0$ and $f_d(R_2, F_2) = 0$. Thus we have that $f_d(S_{g_1,g_2,F}, F) = -f_d(R_{g_1,g_2,F}, T_{g_1,g_2,F})$ and Lemma 5.3 gives us that $f_d(R_{g_1,g_2,F}, T_{g_1,g_2,F}) \geq 0$. ∎

5.2 Expected number of edges

In this section we consider the expected number of edges contained in the final graph of the H-elimination process. We give lower and upper bounds which match up to a constant factor when H is 2-balanced, and we give the exact asymptotics when H is strictly 2-balanced and when $K = K_4^-$ and $H = K_{3,4}^-$.

Let $\mathcal{H}_e$ denote the event that an edge e is inserted into the H-elimination random graph process. Recall that $\mathcal{H}_e$ holds if no copy of (r, H_r) is rooted at e in G_{n,p_e} for any $r \in E(H)$.

Erdős, Suen and Winkler [17] have considered the case when H is a triangle. Note that in this case for $\mathcal{H}_e$ to hold e must have no mutual neighbours in G_{n,p_e}. Since the appearance of the individual mutual neighbours of e are independent thus

$$\mathbb{P}(e \in E(G_e(K_3)_{n,1})|p_e) = (1 - p_e^2)^{n-2}$$

and thus

$$\mathbb{P}(e \in E(G_e(K_3)_{n,1})) = \int_0^1 (1 - p_e^2)^{n-2} dp_e = 4^{n-2}\frac{((n-2)!)^2}{(2n-3)!} = (1 + o(1))\frac{\sqrt{\pi}}{2}n^{-1/2}.$$

Therefore the expected number of edges in the process is $(1 + o(1))\sqrt{\pi}n^{3/2}/4$. We will generalise this result.

Let (R, F) be a rooted graph and $S \subseteq V(G_{n,p})$ with $|S| = |R|$. Let $I_{S,(R,F)}$ be the set of injective functions $f : V(F) \to V(G_{n,p})$ such that $f(R) = S$. Note that if for some function $f \in I_{S,(R,F)}$ we have that $f(E(F)) \subseteq E(G_{n,p})$ then there is a copy of (R, F) rooted at S in $G_{n,p}$. Now consider a set of rooted graphs $M = \{(R_i, F_i)\}_{i=1}^k$ where $|R_i| = |R_j|$ for every $i, j = 1..k$. We define

$$I_{S,M} = \bigcup_{(R,F)\in M} \bigcup_{f\in I_{S,(R,F)}} ((R, F), f).$$

However it can happen for some $((R, F), f), ((R', F'), f') \in I_{S,M}$ that $f(E(F)) = f'(E(F'))$. Based on this we can split the elements of $I_{S,M}$ into equivalence classes and let $I'_{S,M}$ denote the set where one representative of each of these classes is taken. We will also need the notation $[a]_b = a(a-1)...(a-b+1)$.

5.2.1 2-balanced graphs

We start by showing bounds on the expected number of edges in the random H-elimination graph process when H is 2-balanced.

Lemma 5.6

Let H be a 2-balanced random graph and let p_e be the probability assigned to edge e. Assume $n^{-1/d_2(H)} \le p_e \le n^{-1/d_2(H)} \log^2 n$ then there exist constants $c, C > 0$ depending only on H and not on p_e such that

$$\exp(-C(n^{v(H)-2}p_e^{e(H)-1})) \le \mathbb{P}(\mathcal{H}_e|p_e) \le \exp\left(-cn^{1/d_2(H)}p_e\right).$$

PROOF Let $M = \{(r, H_r)\}_{r\in E(H)}$. Since $|I_{e,(r,H_r)}| = 2[n]_{v(H)-2}$ for $r \in E(H)$ thus we have that $|I_{e,M}| = 2e(H)[n]_{v(H)-2} = (1 + o(1))2e(H)n^{v(H)-2}$. Note that every f, such that $((r, H_r), f) \in I_{e,M}$ for some (r, H_r), is counted $\mathrm{aut}(H)$ times. Therefore $|I'_{e,M}| = (1+o(1))2e(H)n^{v(H)-2}/\mathrm{aut}(H)$. For $((R, F), f) \in I'_{S,M}$ let $X_{(R,F),f,p_e}$ be the indicator random variable for the event that $f(E(F)) \subseteq E(G_{n,p_e})$. Let

$$X_{p_e} = \sum_{((R,F),f)\in I'_{S,M}} X_{(R,F),f,p_e}$$

then we have that

$$\mathbb{E}(X_{p_e}) = (1+o(1))\frac{2e(H)}{\mathrm{aut}(H)} n^{v(H)-2} p_e^{e(H)-1}.$$

Note that

$$\mathbb{P}(\mathcal{H}_e|p_e) = \mathbb{P}\left(\{X_{p_e} = 0\}\right).$$

Since $\mathbb{E}(X_{(R,F),f,p_e}) = p_e^{e(H)-1} = o(1)$ applying Lemma 2.6 gives

$$\mathbb{P}(X_{p_e} = 0) \geq \exp\left(-\frac{\mathbb{E}(X_{p_e})}{1-p_e^{e(H)-1}}\right) \geq \exp\left(-C(n^{v(H)-2}p_e^{e(H)-1})\right)$$

for some constant C depending only on H (e.g. $C = 2e(H)+1$ works for every H). For the upper bound Theorem 2.8 will be used. First we need to calculate the value of Δ. Let $\tau = n^{1/d_2(H)}p_e$ and note that $1 \leq \tau \leq \log^2 n$ then

$$\begin{aligned}
\Delta &\leq \sum_{r_1,r_2\in E(H)} \sum_{\substack{(g_1,g_2,F)\in((r_1,H_{r_1})\oplus(r_2,H_{r_2}))\\ g_1(r_1)=g_2(r_2)}} n^{v(F)-2}p_e^{e(F)} \\
&\leq \sum_{r_1,r_2\in E(H)} \sum_{(g_1,g_2,F)\in((r_1,H_{r_1})\oplus(r_2,H_{r_2}))} n^{f_{d_2(H)}(S_{g_1,g_2,F},F)}\tau^{e(F)} \\
&\leq c_1\tau^{2e(H)-3}
\end{aligned}$$

since $f_{d_2(H)}(S_{g_1,g_2,F},F) \leq 0$ according to Lemma 5.5 and $e(F) \leq 2(e(H)) - 3$. Note that $c_1 \leq \sum_{r_1,r_2\in E(H)} |(r_1,H_{r_1}) \oplus (r_2,H_{r_2})|$ and thus depends only on H. We already established that $\mathbb{E}(X_{p_e}) \geq c_2 n^{v(H)-2}p_e^{e(H)-1} = c_2\tau^{e(H)-1}$ for any $c_2 < 2(e(H))/\mathrm{aut}(H)$. Applying Theorem 2.8 gives us that

$$\begin{aligned}
\mathbb{P}(X_{p_e} = 0) &\leq \exp\left(-\frac{(c_2\tau^{e(H)-1})^2}{2c_1\tau^{2e(H)-3}}\right) \\
&= \exp\left(-\frac{c_2^2}{2c_1}\tau\right) \\
&= \exp\left(-cn^{1/d_2(H)}p_e\right)
\end{aligned}$$

where $c = c_2^2/(2c_1)$ depends only on H as c_1 and c_2 depend only on H. ∎

Note that the proof actually works for a much wider range of p_e. The lower bound actually holds for $p_e = o(1)$ for any fixed $\varepsilon > 0$ and the upper bound holds for $n^{-1/d_2(H)} \leq p_e \leq 1$. However the Lemma above is enough to show the asymptotics of the probability that the edge e will be inserted.

Lemma 5.7

Let H be a 2-balanced graph then:

$$\mathbb{P}(\mathcal{H}_e) = \Theta(n^{-1/d_2(H)}).$$

PROOF Let $p_0 = n^{-1/d_2(H)}$ and $p_1 = n^{-1/d_2(H)} \log^2 n$. Clearly

$$\begin{aligned}\mathbb{P}(\mathcal{H}_e) &= \int_0^1 \mathbb{P}(\mathcal{H}_e|p_e)dp_e \\ &= \int_0^{p_0} \mathbb{P}(\mathcal{H}_e|p_e)dp_e + \int_{p_0}^{p_1} \mathbb{P}(\mathcal{H}_e|p_e)dp_e + \int_{p_1}^{1} \mathbb{P}(\mathcal{H}_e|p_e)dp_e.\end{aligned}$$

Since $0 \leq \mathbb{P}(\mathcal{H}_e|p_e) \leq 1$ thus

$$0 \leq \int_0^{p_0} \mathbb{P}(\mathcal{H}_e|p_e)dp_e \leq p_0 = n^{-1/d_2(H)}.$$

Note that

$$\begin{aligned}\int_{p_0}^{p_1} \exp(-C(n^{v(H)-2}p_e^{e(H)-1}))dp_e &= n^{-1/d_2(H)} \int_1^{\log^2 n} \exp(-C\tau^{e(H)-1})d\tau \\ &= (1+o(1))n^{-1/d_2(H)} \int_1^{\infty} \exp(-C\tau^{e(H)-1})d\tau \\ &= \Theta\left(n^{-1/d_2(H)}\right).\end{aligned}$$

Also

$$\begin{aligned}\int_{p_0}^{p_1} \exp\left(-cn^{1/d_2(H)}p_e\right) dp_e &= n^{-1/d_2(H)} \int_1^{\log^2 n} \exp(-c\tau)d\tau \\ &= (1+o(1))n^{-1/d_2(H)} \int_1^{\infty} \exp(-c\tau)d\tau \\ &= \Theta\left(n^{-1/d_2(H)}\right).\end{aligned}$$

Applying Lemma 5.6 gives us that

$$\int_{p_0}^{p_1} \exp(-C(n^{v(H)-2}p_e^{e(H)-1}))dp_e \leq \int_{p_0}^{p_1} \mathbb{P}(\mathcal{H}_e|p_e)dp_e$$

and that

$$\int_{p_0}^{p_1} \mathbb{P}(\mathcal{H}_e|p_e)dp_e \leq \int_{p_0}^{p_1} \exp(-c(n^{1/d_2(H)}p_e))dp_e.$$

Thus

$$\int_{p_0}^{p_1} \mathbb{P}(\mathcal{H}_e|p_e)dp_e = \Theta\left(n^{-1/d_2(H)}\right).$$

We also have that

$$\begin{aligned}\int_{p_1}^{1} \mathbb{P}(\mathcal{H}_e|p_e)dp_e \leq \int_{p_1}^{1} \mathbb{P}(\mathcal{H}_e|p_1)dp_e &\leq \mathbb{P}(\mathcal{H}_e|p_1) \\ &\leq \exp(-c\log^2 n) = o(n^{-4})\end{aligned}$$

and the result follows. ∎

Corollary 5.8

When H is a 2-balanced graph then

$$\mathbb{E}(e(G_e(H)_{n,1})) = \Theta(n^{2-1/d_2(H)}).$$

5.2.2 Strictly 2-balanced graphs

In the case when H is strictly 2-balanced these bounds can be improved. However instead of just showing bounds on $\mathbb{P}(\mathcal{H}_e|p_e)$ we will show a slightly more general bound, which we will need later.

Lemma 5.9

Let H be a strictly 2-balanced random graph. Let $M = \{r, H_r\}_{r\in E(H)}$ and $Y \subseteq I'_{e,M}$. Define the event $\mathcal{Y}$ as the event that for every $((r, H_r), f) \in Y$ we have that $f(E(H)) \not\subseteq E(G_{n,p})$.

If in addition we have that $|Y| = (1 + O(n^{-\varepsilon}))|I'_{e.M}|$ for some fixed $\varepsilon \geq 0$ and that $p \leq p_0 = n^{-1/d_2(H)} \log n$ then

$$\mathbb{P}(\mathcal{Y}) = (1 + o(1)) \exp\left(-\frac{2e(H)}{\text{aut}(H)}(n^{v(H)-2}p^{e(H)-1})\right).$$

Note that when $p_e \leq p_0$ then by setting $Y = I'_{e,M}$ and $p = p_e$ this implies that

$$\mathbb{P}(\mathcal{H}_e|p_e) = (1 + o(1)) \exp\left(-\frac{2e(H)}{\text{aut}(H)}(n^{v(H)-2}p_e^{e(H)-1})\right).$$

The proof is similar to the 2-balanced case with the exception that in this case $\Delta = o(1)$ and thus sharpening the remaining estimates gives us a tighter bound.

PROOF Let $M = \{r, H_r\}_{r\in E(H)}$. Then we have that $|I_{e.M}| = 2e(H)[n]_{v(H)-2} = (1 + O(n^{-1}))2e(H)n^{v(H)-2}$. As in the 2-balanced case every f, such that $((r, H_r), f) \in I_{e,M}$ for some (r, H_r), is counted $\text{aut}(H)$ times and thus $|I'_{e.M}| = |I_{e,M}|/\text{aut}(H)$. Set $\delta = \min\{\varepsilon, 1/(2d_2(H))\}$ and note that this implies that $0 < \delta < 1$. Therefore

$$|Y| = (1 + O(n^{-\delta}))\frac{2e(H)n^{v(H)-2}}{\text{aut}(H)}.$$

For $((R, F), f) \in Y$ let $X_{(R,F),f,p}$ be the indicator random variable for the event that $f(E(F)) \subseteq E(G_{n,p})$. Let

$$X_p = \sum_{((R,F),f)\in Y} X_{(R,F),f,p}$$

then we have that

$$\mathbb{E}(X_p) = (1 + O(n^{-\delta}))\frac{2e(H)}{\text{aut}(H)}n^{v(H)-2}p^{e(H)-1}.$$

Note that $\mathbb{E}(X_{(R,F),f,p}) = p^{e(H)-1}$ and that $1/(1-p^{e(H)-1}) = 1+O(p_0) = 1+O(n^{-\delta})$. Applying Lemma 2.6 gives

$$\begin{aligned}\mathbb{P}(X_p = 0) &\geq \exp\left(-\frac{\mathbb{E}(X_p)}{1-p^{e(H)-1}}\right) \\ &\geq \exp\left(-(1+O(n^{-\delta}))\frac{2e(H)}{\mathrm{aut}(H)}\left(n^{v(H)-2}p^{e(H)-1}\right)\right).\end{aligned}$$

For the upper bound Theorem 2.8 will be used. First we need to calculate the value of Δ. We have that

$$\begin{aligned}\Delta &\leq \sum_{r_1,r_2\in E(H)} \sum_{\substack{(g_1,g_2,F)\in((r_1,H_{r_1})\oplus(r_2,H_{r_2}))\\ g_1(r_1)=g_2(r_2)}} n^{v(F)-2}p^{e(F)} \\ &\leq \sum_{r_1,r_2\in E(H)} \sum_{\substack{(g_1,g_2,F)\in((r_1,H_{r_1})\oplus(r_2,H_{r_2}))\\ g_1(r_1)=g_2(r_2)}} n^{v(F)-2}p_0^{e(F)} \\ &\leq \sum_{r_1,r_2\in E(H)} \sum_{(g_1,g_2,F)\in((r_1,H_{r_1})\oplus(r_2,H_{r_2}))} n^{f_{d_2(H)}(S_{g_1,g_2,F},F)}\log^{e(F)} n \\ &= \sum_{r_1,r_2\in E(H)} \sum_{(g_1,g_2,F)\in((r_1,H_{r_1})\oplus(r_2,H_{r_2}))} n^{f_{d_2(H)}(S_{g_1,g_2,F},F)+o(1)} \\ &= o(1).\end{aligned}$$

The last step follows from Lemma 5.4 since $f_{d_2(H)}(S_{g_1,g_2,F}, F) < 0$ as when H is strictly 2-balanced graph then (r, H_r) is a nontrivial strictly balanced rooted graph such that $d(r, H_r) = d_2(H)$.

We already established that $\mathbb{E}(X_p) = (1+O(n^{-\delta}))2e(H)n^{v(H)-2}p^{e(H)-1}/\mathrm{aut}(H)$. Applying Theorem 2.8 gives us that

$$\begin{aligned}\mathbb{P}(X_p = 0) &\leq \exp\left(-(1+O(n^{-\delta}))\frac{2e(H)n^{v(H)-2}p^{e(H)-1}}{\mathrm{aut}(H)} + o(1)\right) \\ &= (1+o(1))\exp\left(-(1+O(n^{-\delta}))\frac{2e(H)n^{v(H)-2}p^{e(H)-1}}{\mathrm{aut}(H)}\right).\end{aligned}$$

Therefore we have shown that

$$\mathbb{P}(X_p = 0) = (1+o(1))\exp\left(-(1+O(n^{-\delta}))\frac{2e(H)n^{v(H)-2}p^{e(H)-1}}{\mathrm{aut}(H)}\right).$$

Note that $O(n^{-\delta}n^{v(H)-2}p^{e(H)-1}) = O(n^{-\delta}\log^{e(H)-1} n) = o(1)$ and thus

$$\mathbb{P}(X_p = 0) = (1+o(1))\exp\left(-\frac{2e(H)n^{v(H)-2}p^{e(H)-1}}{\mathrm{aut}(H)}\right).$$

■

Lemma 5.10
Let H be a strictly 2-balanced graph then we have that:

$$\mathbb{P}(\mathcal{H}_e) = (1+o(1))n^{-1/d_2(H)}\,\Gamma\left(\frac{e(H)}{e(H)-1}\right)\Big/\left(\frac{2e(H)}{\mathrm{aut}(H)}\right)^{1/(e(H)-1)}$$

where $\Gamma(x)$ is the gamma function.

PROOF Let $p_0 = n^{-1/d_2(H)}\log n$ then:

$$\mathbb{P}(\mathcal{H}_e) = \mathbb{P}(\mathcal{H}_e \cap \{p_e \le p_0\}) + \mathbb{P}(\mathcal{H}_e \cap \{p_e > p_0\}).$$

When $p_e \le p_0$ Lemma 5.9 can be applied, thus:

$$\begin{aligned}
\mathbb{P}(\mathcal{H}_e \cap \{p_e \le p_0\}) &= \int_0^{p_0} \mathbb{P}(\mathcal{H}_e | p_e) dp_e \\
&= (1+o(1)) \int_0^{p_0} \exp\left(-\frac{2e(H)p_e^{e(H)-1}n^{v(H)-2}}{\mathrm{aut}(H)}\right) dp_e \\
&= (1+o(1))n^{-1/d_2(H)} \int_0^{\log n} \exp\left(-\frac{2e(H)t^{e(H)-1}}{\mathrm{aut(H)}}\right) dt \\
&= (1+o(1))n^{-1/d_2(H)} \int_0^{\infty} \exp\left(-\frac{2e(H)t^{e(H)-1}}{\mathrm{aut}(H)}\right) dt
\end{aligned}$$

and

$$\int_0^{\infty} \exp\left(-\frac{2e(H)t^{e(H)-1}}{\mathrm{aut}(H)}\right) dt = \Gamma\left(\frac{e(H)}{e(H)-1}\right)\Big/\left(\frac{2e(H)}{\mathrm{aut}(H)}\right)^{1/(e(H)-1)}.$$

On the other hand $\mathbb{P}(\mathcal{H}_e \cap \{p_e > p_0\}) = \int_{p_0}^1 \mathbb{P}(\mathcal{H}_e|p_e)dp_e \le \int_{p_0}^1 \mathbb{P}(\mathcal{H}_e|\{p_e = p_0\})dp_e \le \mathbb{P}(\mathcal{H}_e|\{p_e = p_0\})$ and

$$\mathbb{P}(\mathcal{H}_e|\{p_e = p_0\}) \le (1+o(1))\exp\left(-\log^{e(H)-1} n\right) \le \exp(-\log^2 n + o(1)) = o(n^{-4}),$$

as every strictly 2-balanced graph contains at least 3 edges. ■

Corollary 5.11
When H is a strictly 2-balanced graph then

$$\mathbb{E}(e(G_e(H)_{n,1})) = (1+o(1))\frac{1}{2}n^{2-1/d_2(H)}\,\Gamma\left(\frac{e(H)}{e(H)-1}\right)\Big/\left(\frac{2e(H)}{\mathrm{aut}(H)}\right)^{1/(e(H)-1)}.$$

5.2.3 The K_4^-

We have shown the exact asymptotics of the expected number of edges found in the H-elimination process when H is strictly 2-balanced. Next we show similarly tight bounds on the expected number of edges in the H-elimination process when $H = K_4^-$ and when $H = K_{3,4}^-$. Recall that the K_4^- is a 2-balanced graph, however it is not strictly 2-balanced as $d_2(K_4^-) = d_2(K_3) = 2$ and $K_3 \subsetneq K_4^-$.

Theorem 5.12
Let $H = K_4^-$. *Assume that* $p_e \le p_0 = n^{-1/2}\log n$ *then*

$$\mathbb{P}(\mathcal{H}_e|p_e) = (1+o(1))(\exp(-np_e^2) + np_e^2\exp(-3np_e^2))$$

PROOF Let $e = \{u, v\}$. Note that for $\mathcal{H}_e$ to hold either e has no mutual neighbours in G_{n,p_e} or e has exactly one mutual neighbour in G_{n,p_e}, denoted by w, and the codegree of both $\{u, w\}$ and $\{v, w\}$ is zero. In case the codegree of e in G_{n,p_e} is larger than one then $\mathcal{H}_e$ does not hold. Let $\mathcal{E}_{\{u,v\},W}$ where $W \subsetneq V(G_{n,p_e})$ be the event that W forms the set of mutual neighbours of $\{u, v\}$ in $G_{n,p_{\{u,v\}}}$. Then

$$\mathbb{P}(\mathcal{H}_e|p_e) = \mathbb{P}(\mathcal{H}_e \cap \mathcal{E}_{e,\emptyset}|p_e) + \sum_{w \in V(G_{n,p_e})\setminus e} \mathbb{P}(\mathcal{H}_e \cap \mathcal{E}_{e,\{w\}}|p_e).$$

Note that $\mathcal{E}_{e,\emptyset} \subsetneq \mathcal{H}_e$. Therefore $\mathbb{P}(\mathcal{H}_e \cap \mathcal{E}_{e,\emptyset}|p_e) = \mathbb{P}(\mathcal{E}_{e,\emptyset}|p_e)$ and the proof of Lemma 5.9 implies that $\mathbb{P}(\mathcal{E}_{e,\emptyset}|p_e) = (1+o(1))\exp(-np_e^2)$. Now to establish the value of $\mathbb{P}(\mathcal{H}_e \cap \mathcal{E}_{e,\{w\}}|p_e)$ for a fixed w. Let $M = (r, (K_3)_r)$. Let I be the union of the sets $I'_{e,M}$, $I'_{\{u,w\},M}$ and $I'_{\{v,w\},M}$ excluding the elements $((r, (K_3)_r), f)$ such that $f(V(K_3)) = \{u, v, w\}$ formally

$$I = \{((r, (K_3)_r), f) \in I'_{e,M} \cup I'_{\{u,w\},M} \cup I'_{\{v,w\},M} : f(V(K_3)) \ne \{u, v, w\}\}.$$

Note that $|I| = 3(n-3) = 3(1+O(n^{-1}))n$. For $((r, (K_3)_r), f) \in I$ let $X_{(r,(K_3)_r),f,p_e}$ be the indicator random variable for the event that $f(E((K_3)_r)) \subseteq E(G_{n,p_e})$. Let

$$X_{p_e} = \sum_{((r,(K_3)_r),f) \in I} X_{(r,(K_3)_r),f,p_e}$$

then we have that $\mathbb{E}(X_{p_e}) = 3(1+O(n^{-1}))np_e^2$. Note that

$$\mathbb{P}\left(\mathcal{H}_e \cap \mathcal{E}_{e,\{w\}}|p_e\right) = \mathbb{P}\left(\{X_{p_e} = 0\} \cap \{\{u, w\}, \{v, w\} \in E(G_{n,p_e})\}\right).$$

Note that $\mathbb{E}(X_{(r,(K_3)_r),f,p_e}) = p_e^2$ and that $1/(1-p_e^2) = 1 + O(p_0)$. Since $O(n^{-1}) = O(p_0)$ applying Lemma 2.6 gives

$$\mathbb{P}(X_{p_e} = 0) \ge \exp\left(-\frac{\mathbb{E}(X_{p_e})}{1-p_e^2}\right) \ge \exp\left(-3(1+O(p_0))(np_e^2)\right).$$

For the upper bound Theorem 2.8 will be used. First we need to calculate the value of Δ.

$$\begin{aligned}\Delta &\le \sum_{r_1,r_2\in E(K_3)} \sum_{(g_1,g_2,F)\in((r_1,H_{r_1})\oplus(r_2,H_{r_2}))} \binom{3}{|S_{g_1,g_2,F}|} n^{v(F)-|S_{g_1,g_2,F}|} p_e^{e(F)} \\ &\le \sum_{r_1,r_2\in E(K_3)} \sum_{(g_1,g_2,F)\in((r_1,H_{r_1})\oplus(r_2,H_{r_2}))} \binom{3}{|S_{g_1,g_2,F}|} n^{v(F)-|S_{g_1,g_2,F}|} p_0^{e(F)} \\ &\le \sum_{r_1,r_2\in E(K_3)} \sum_{(g_1,g_2,F)\in((r_1,H_{r_1})\oplus(r_2,H_{r_2}))} \binom{3}{|S_{g_1,g_2,F}|} n^{f_2((g_1(r_1),F))+o(1)} \\ &= o(1)\end{aligned}$$

according to Lemma 5.4, $f_2(S_{g_1,g_2,F}, F) < 0$, as we are combining strictly balanced rooted graphs of density 2. We already established that $\mathbb{E}(X_{p_e}) = 3(1 + O(p_0))np_e^2$. Applying Theorem 2.8 gives us that

$$\begin{aligned}\mathbb{P}(X_{p_e} = 0) &\le \exp\left(-3(1 + O(p_0))np_e^2 + o(1)\right) \\ &= (1 + o(1)) \exp\left(-3(1 + O(p_0))np_e^2\right).\end{aligned}$$

Since $O(np_0^3) = o(1)$

$$\mathbb{P}(X_{p_e} = 0) = (1 + o(1)) \exp\left(-3(1 + O(p_0))np_e^2\right) = (1 + o(1)) \exp\left(-3np_e^2\right).$$

Note that the events $\{X_{p_e} = 0\}$ and $\{\{u, w\}, \{v, w\} \in E(G_{n,p_e})\}$ are independent and therefore

$$\mathbb{P}(\mathcal{H}_e \cap \mathcal{E}_{e,\{w\}} | p_e) = (1 + o(1))p_e^2 \exp\left(-3np_e^2\right).$$

The statement follows as the events $\mathcal{E}_{e,w}$ and $\mathcal{E}_{e,w'}$ are mutually exclusive for any $w, w' \in V(G_{n,p_e})\backslash e$ when $w \neq w'$. ■

Lemma 5.13
Let $H = K_4^-$ then

$$\mathbb{P}(\mathcal{H}_e) = (1 + o(1)) \left(\frac{\sqrt{\pi}}{2} + \frac{\sqrt{\pi}}{12\sqrt{3}}\right) n^{-1/2}.$$

PROOF Let $p_0 = n^{-1/2} \log n$. Similarly to the previous calculations we have that

$$\mathbb{P}(\mathcal{H}_e) = \mathbb{P}(\mathcal{H}_e \cap \{p_e \le p_0\}) + \mathbb{P}(\mathcal{H}_e \cap \{p_e > p_0\}).$$

According to Lemma 5.12

$$\begin{aligned}
\mathbb{P}(\mathcal{H}_e \cap \{p_e \le p_0\}) &= \int_0^{p_0} \mathbb{P}(\mathcal{H}_e | p_e) dp_e \\
&= \int_0^{p_0} (1+o(1))(\exp(-np_e^2) + np_e^2 \exp(-3np^2)) dp_e \\
&= (1+o(1)) n^{-1/2} \int_0^{\log n} \exp(-\tau^2) + \tau^2 \exp(-3\tau^2) d\tau \\
&= (1+o(1)) n^{-1/2} \int_0^{\infty} \exp(-\tau^2) + \tau^2 \exp(-3\tau^2) d\tau \\
&= (1+o(1)) n^{-1/2} \left(\frac{\sqrt{\pi}}{2} + \frac{\sqrt{\pi}}{12\sqrt{3}} \right).
\end{aligned}$$

We also have that

$$\mathbb{P}(\mathcal{H}_e \cap \{p_e > p_0\}) \le \mathbb{P}(\mathcal{H}_e | p_0) \le 2\exp(-\log^2 n) = o(n^{-4})$$

completing the proof. ∎

Corollary 5.14

$$\mathbb{E}(e(G_{n,1}(K_4^-))) = (1+o(1)) \left(\frac{\sqrt{\pi}}{4} + \frac{\sqrt{\pi}}{24\sqrt{3}} \right) n^{3/2}.$$

5.2.4 The $K_{3,4}^-$

Now the case when $H = K_{3,4}^-$. This is also a 2-balanced graph which is not strictly 2-balanced as $d_2(K_{3,3}) = d_2(K_{3,4}^-) = 2$. Recall that in the K_4^- case we conditioned on the number of copies of $(r, (K_3)_r)$ rooted at e in G_{n,p_e} and this was done because $d_2(K_3) = d_2(K_4^-)$ and $K_3 \subsetneq K_4^-$. In case of the $K_{3,4}^-$ we will condition on the number of copies of $(r, (K_{3,3})_r)$ rooted at e in G_{n,p_e}. However in the case of the K_4^- we had that $\mathcal{H}_e$ held only in case the codegree of e was at most 1. However no such restrictions apply in the $K_{3,4}^-$ case for the number of copies of $(r, (K_{3,3})_r)$ rooted at e in G_{n,p_e}. We will show that for any pair of vertices e there are a.a.s. at most $2\log^8 n$ copies of $(r, (K_{3,4}^-)_r)$ rooted at e in G_{n,p_e} when $p_e \le n^{-1/2}\log n$, but due to this, one has to distinguish between many cases. The other difference is that the copies of $(r, (K_{3,3})_r)$ rooted at e can overlap. We will show that this happens with small probability and thus these cases can be ignored.

A key property used in the previous proof was that the event that w was a mutual neighbour of $\{u, v\}$ was independent of the event that $\{u, v\}$, $\{u, w\}$ or $\{v, w\}$ has any additional mutual neighbours. The following definition helps us express this property.

Definition 5.2

Let (R, F) be a rooted graph and G be a graph such that $v(F) \le v(G)$. Also let $S \subseteq Q \subseteq V(G)$ such that $|S| = |R|$. We say that there is a copy of (R, F) rooted at S and outside Q

in G if there exists an injective function $f : V(F) \to V(G)$ such that $f(E(f)) \subseteq E(G)$, $f(R) = S$ and $f(V(F)) \cap Q = S$.

For a set of vertices $W \subseteq V(G_{n,p_e})$ let

$$I_{S,(R,F),W} = \{f \in I_{S,(R,F)} : f(V(F) \backslash R) \cap W = \emptyset\}.$$

The sets $I_{S,M,W}$ and $I'_{S,M,W}$ are defined analogously to the sets $I_{S,M}$ and $I'_{S,M}$.

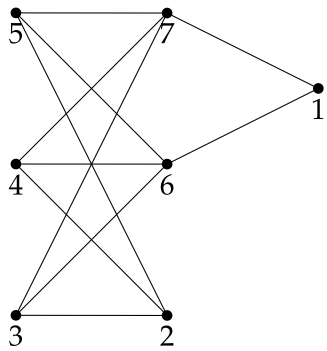


Figure 5.1: The $K_{3,4}^-$

Figure 5.1 shows a labelled copy of $K_{3,4}^-$. Let $r_1 = \{1, 6\}$, $r_2 = \{2, 3\}$ and $r_3 = \{4, 6\}$. Note that due to symmetries of the graph edge e is inserted during the $K_{3,4}^-$ elimination process iff there is no copy of $(r_1, (K_{3,4}^-)_{r_1})$, no copy of $(r_2, (K_{3,4}^-)_{r_2})$ and no copy of $(r_3, (K_{3,4}^-)_{r_3})$ rooted at e in G_{n,p_e} see Figure 5.2.

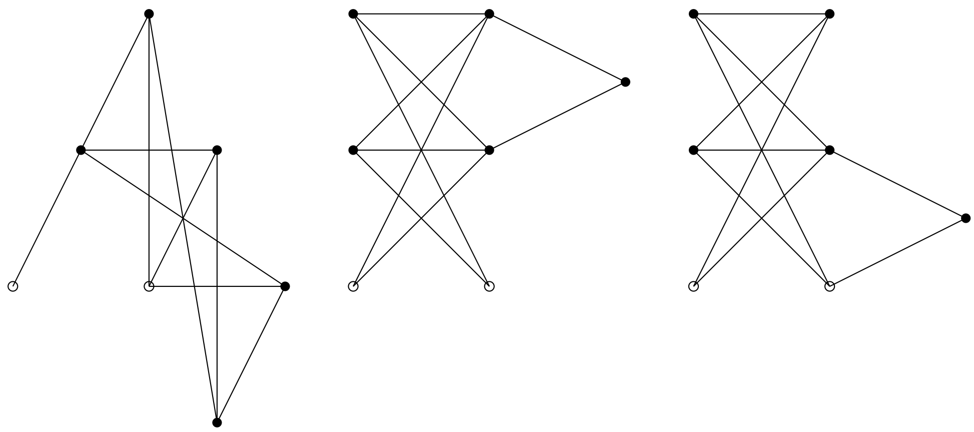


Figure 5.2: The forbidden rooted graphs

Lemma 5.15

Let $H = K_{3,4}^-$ and assume that $p_e \leq p_0 = n^{-1/2} \log n$ then

$$\mathbb{P}(\mathcal{H}_e | p_e) = (1 + o(1)) \exp\left(-\frac{n^4}{4} p_e^8 - \frac{n^5}{3} p_e^{10}\right) \exp\left(\frac{n^4}{4} p_e^8 \exp(-6np_e^2)\right) + o(n^{-4}).$$

PROOF As mentioned above we split the cases according to the number of copies of $(r, (K_{3,3})_r)$ rooted at e. Assume $e = \{u, v\}$.

Define the set B as the set of sets $A \subseteq I'_{e,(r,(K_{3,3})_r)}$ such that $a(V(K_{3,3})) \cap b(V(K_{3,3})) = e$ for every $a, b \in A$ when $a \neq b$. Let $\mathcal{M}_A$ be the event that a set $A \in B$ describes a maximal set of vertex disjoint copies of $(r, (K_{3,3})_r)$ rooted at e. Formally the event $\mathcal{M}_A$ holds if for every $a \in A$ we have that $a(E((K_{3,3})_{r_2})) \subseteq E(G_{n,p_e})$ and A is maximal in the sense that for any injective function $g \in I'_{e,(r,(K_{3,3})_r)} \setminus A$ such that $g(V(K_{3,3})) \cap a(V(K_{3,3})) = e$ for every $a \in A$ we have that $g(E((K_{3,3})_r)) \not\subseteq E(G_{n,p_e})$. Let $\mathcal{D}$ be the event that there are no overlapping copies of $(r, (K_{3,3})_r)$ rooted at e. Then

$$\mathbb{P}(\mathcal{H}_e|p_e) \geq \mathbb{P}(\mathcal{H}_e \cap \mathcal{D}|p_e) \geq \sum_{\substack{A \in B \\ |A| \leq 2\log^8 n}} \mathbb{P}(\mathcal{H}_e \cap \mathcal{D} \cap \mathcal{M}_A|p_e).$$

Fix A such that $|A| \leq 2\log^8 n$. Our goal is to determine a lower bound on $\mathbb{P}(\mathcal{H}_e \cap \mathcal{D} \cap \mathcal{M}_A|p_e)$. Let $\mathcal{E}_A$ be the event that $a(E((K_{3,3})_r)) \subseteq E(G_{n,p_e})$ for every $a \in A$ and let $\mathcal{I}_A$ be the event that $a(E(\overline{K_{3,3}})) \cap E(G_{n,p_e}) = \emptyset$ for every $a \in A$. Note that if both $\mathcal{E}_A$ and $\mathcal{I}_A$ hold then the copies in A are induced copies. Define $U = \bigcup_{a \in A} a(V(K_{3,3}))$. Now we introduce an event $\mathcal{G}$ such that $(\mathcal{G} \cap \mathcal{E}_A \cap \mathcal{I}_A|p_e) \subseteq (\mathcal{H}_e \cap \mathcal{D} \cap \mathcal{M}_A|p_e)$ and the events $\mathcal{G}$ and $\mathcal{E}_A \cap \mathcal{I}_A$ are independent, i.e. $\mathcal{G}$ depends only on the presence or absence of edges not spanned by U. Clearly for any such event $\mathcal{G}$ we have that

$$\mathbb{P}(\mathcal{H}_e \cap \mathcal{D} \cap \mathcal{M}_A|p_e) \geq \mathbb{P}(\mathcal{G} \cap \mathcal{I}_A \cap \mathcal{E}_A|p_e) = \mathbb{P}(\mathcal{G}|p_e)\mathbb{P}(\mathcal{E}_A \cap \mathcal{I}_A|p_e).$$

Let $\mathcal{G}$ be the following event:

- there is no copy of $(r, (K_3)_r)$ rooted at $a(\overline{e})$ outside U in G_{n,p_e} for any $a \in A$ and $\overline{e} \in E(\overline{K_{3,3}})$
- no copy of $(r_1, (K^-_{3,4})_{r_1})$ is rooted at e outside U in G_{n,p_e}
- no copy of $(r, (K_{3,3})_r)$ is rooted at e outside U in G_{n,p_e}
- for any $R \subsetneq V(K^-_{3,4})$, such that $|R| > 2$ and $K^-_{3,4}[R] \neq K_{3,3}$, no copy of $(R, (K^-_{3,4})_R)$ is rooted at S outside U in G_{n,p_e} for any $S \subseteq U$ with $|S| = |R|$
- for any $R \subsetneq V(K_{3,3})$ with $|R| > 2$ no copy of $(R, (K_{3,3})_R)$ is rooted at S outside U in G_{n,p_e} for any $S \subseteq U$ with $|S| = |R|$.

Note that when conditioning on $\mathcal{I}_A \cap \mathcal{E}_A$ this ensures that A contains all the copies of $(r, (K_{3,3})_r)$ rooted at e and also that no copy of $(r, (K^-_{3,4})_r)$ for any $r \in E(K^-_{3,4})$ is rooted at e.

Let $M_1 = (r, (K_3)_r)$ and $M_2 = ((r, (K_{3,3})_r), (r_1, (K_{3,4}^-)_{r_1}))$. Also for $j > 2$ let

$$\begin{aligned} M_j = \{R, (K_{3,3})_R &: R \subsetneq V(K_{3,3}) \cap |R| = j\} \\ \cup \{R, (K_{3,4}^-)_R &: R \subsetneq V(K_{3,4}^-) \cap K_{3,4}^-[R] \neq K_{3,3} \cap |R| = j\}. \end{aligned}$$

Let

$$I = \bigcup_{a \in A} \bigcup_{\overline{e} \in E(\overline{K_{3,3}})} I'_{a(\overline{e}),M_1,U} \cup I'_{e,M_2,U} \cup \bigcup_{j=3}^{6} \bigcup_{S \in \binom{U}{j}} I'_{S,M_j,U}$$

and for $((R,F),f) \in I$ let $X_{(R,F),f,p_e}$ be the indicator random variable for the event that $f(E(F)) \subseteq E(G_{n,p_e})$. Let

$$X_{p_e} = \sum_{((R,F),f) \in I} X_{(R,F),f,p_e}$$

then we have that $\mathcal{G}$ holds iff $X_{p_e} = 0$. Note that $|I'_{a(\overline{e}),M_1,U}| = n - |U| = n - 4|A| - 2 = n - O(\log^8 n)$ thus

$$\sum_{a \in A} \sum_{\overline{e} \in E(\overline{K_{3,3}})} \sum_{((r,F),f) \in I'_{a(\overline{e}),M_1,U}} \mathbb{E}(X_{(R,F),f,p_e}) = 6|A| \left(1 + O\left(\frac{\log^8 n}{n}\right)\right) np_e^2.$$

We also have that

$$\begin{aligned} \sum_{((r,F),f) \in I'_{e,M_2,U}} \mathbb{E}(X_{(R,F),f,p_e}) &= 2[n - |U|]_2 \binom{n - |U| - 2}{3} p_e^{10} + \binom{n - |U|}{4} \binom{4}{2} p_e^8 \\ &= \left(1 + O\left(\frac{\log^8 n}{n}\right)\right) \left(\frac{n^5}{3} p_e^{10} + \frac{n^4}{4} p_e^8\right). \end{aligned}$$

Fix $R \subsetneq V(K_{3,3})$ and $S \in \binom{U}{|R|}$ then

$$\begin{aligned} \sum_{((R,F),f) \in I'_{S,(R,(K_{3,3})_R),U}} \mathbb{E}(X_{(R,F),f,p_e}) &\leq \sum_{((R,F),f) \in I'_{S,(R,(K_{3,3})_R),U}} \mathbb{E}(X_{(R,F),f,p_0}) \\ &\leq n^{v(K_{3,3}) - |R|} p_0^{e(K_{3,3}) - e(K_{3,3})[R]} \\ &= n^{f_2(R,(K_{3,3})_R) + o(1)}. \end{aligned}$$

Similarly if $R \subsetneq V(K_{3,4}^-)$ and $S \in \binom{U}{|R|}$ then

$$\sum_{((R,F),f) \in I'_{S,(R,(K_{3,3})_R),U}} \mathbb{E}(X_{(R,F),f,p_e}) = n^{f_2(R,(K_{3,4}^-)_R) + o(1)}.$$

Therefore

$$\sum_{j=3}^{6} \sum_{((R,F),f)\in I'_{S,M_j,U}} \mathbb{E}(X_{(R,F),f,p_e}) \le \sum_{j=3}^{6} \sum_{((R,F),f)\in I'_{S,M_j,U}} \mathbb{E}(X_{(R,F),f,p_0})$$

$$= n^{o(1)} \left(\sum_{\substack{R\subsetneq V(K_{3,3})\\ |R|>2}} n^{f_2(R,(K_{3,3})_R)+o(1)} + \sum_{\substack{R\subsetneq V(K^-_{3,4})\\ K^-_{3,4}[R]\neq K_{3,3}\\ |R|>2}} n^{f_2(R,(K^-_{3,4})_R)+o(1)} \right)$$

Let $R \subsetneq V(K_{3,3})$ with $|R| > 2$. If $E(K_{3,3}[R]) = \emptyset$ then

$$|R| + f_2(R,(K_{3,3})_R) = f_2(\emptyset, K_{3,3}[R]) + f_2(R,(K_{3,3})_R) = f_2(\emptyset, K_{3,3}) = 1.5$$

and since $|R| > 2$ thus $f_2(R,(K_{3,3})_R) < 0$. Otherwise let $r \in E(K_{3,3}[R])$ then $f_2(r, K_{3,3}[R]) + f_2(R,(K_{3,3})_R) = f_2(r, K_{3,3}) = 0$. Since $K_{3,3}$ is strictly 2-balanced $f_2(r, K_{3,3}[R]) > 0$ and thus $f_2(R,(K_{3,3})_R) < 0$. The only difference for $K^-_{3,4}$ is that it can happen that for some $r \in E(K^-_{3,4})$ and $r \subsetneq R \subsetneq V(K^-_{3,4})$ we have that $f_2(r, K^-_{3,4}[R]) = 0$, however this is true only when $K^-_{3,4}[R] = K_{3,3}$. Thus $f_2(R,(K^-_{3,4})_R) < 0$ for every $R \subsetneq V(K^-_{3,4})$ with $|R| > 2$ and $K^-_{3,4}[R] \neq K_{3,3}$. Therefore

$$\mathbb{E}(X_{p_e}) = \left(1 + O\left(\frac{\log^8 n}{n}\right)\right)\left(6|A|np_e^2 + \frac{n^4}{4}p_e^8 + \frac{n^5}{3}p_e^{10}\right) + o(1).$$

Since $\mathbb{E}(X_{(R,F),f,p_e}) \le p_0$ for every $((R,F),f) \in I$ Lemma 2.6 gives us the following

$$\begin{aligned}
\mathbb{P}(X_{p_e} = 0) &\ge \exp\left(-(1+O(p_0))\mathbb{E}(X_{p_e})\right) \\
&= \exp\left(-(1+O(p_0))\left(6|A|np_e^2 + \frac{n^4}{4}p_e^8 + \frac{n^5}{3}p_e^{10}\right) + o(1)\right) \\
&= (1+o(1))\exp\left(-\frac{n^4}{4}p_e^8 - \frac{n^5}{3}p_e^{10}\right)\left(\exp(-6np_e^2)\right)^{|A|}.
\end{aligned}$$

Since the event $\mathcal{G}$ is independent of the edges spanned by U thus

$$\begin{aligned}
\mathbb{P}(\mathcal{H}_e \cap \mathcal{D} \cap \mathcal{M}_A | p_e) &\ge \mathbb{P}(X_{p_e} = 0)p_e^{8|A|}(1-p_e)^{\binom{4|A|+2}{2}-8|A|} \\
&\ge \mathbb{P}(X_{p_e} = 0)p_e^{8|A|}(1-p_e)^{8|A|^2+4} \\
&= (1+o(1))\mathbb{P}(X_{p_e} = 0)p_e^{8|A|}.
\end{aligned}$$

The last result needed to establish the lower bound is the number of sets $A \in B$ such that $|A| = k$ when $k \le 2\log^8 n$. Clearly

$$\binom{n^4/4}{k} \ge |\{A \in B : |A| = k\}| \ge \binom{(n-4k-2)^4/4}{k}$$

and thus $|\{A \in B : |A| = k\}| = (1+o(1))(n^{4k}/(4^k k!))$. Thus

$$\begin{aligned}\mathbb{P}(\mathcal{H}_e|p_e) &\geq \sum_{k=0}^{2\log^8 n}\left((1+o(1))\frac{n^{4k}}{4^k k!}p_e^{8k}\exp\left(-\frac{n^4}{4}p_e^8 - \frac{n^5}{3}p_e^{10}\right)\left(\exp(-6np_e^2)\right)^k\right)\\ &= (1+o(1))\exp\left(-\frac{n^4}{4}p_e^8 - \frac{n^5}{3}p_e^{10}\right)\sum_{k=0}^{2\log^8 n}\frac{1}{k!}\left(\frac{n^4}{4}p_e^8\exp(-6np_e^2)\right)^k\\ &= (1+o(1))\exp\left(-\frac{n^4}{4}p_e^8 - \frac{n^5}{3}p_e^{10}\right)\exp\left(\frac{n^4}{4}p_e^8\exp(-6np_e^2)\right).\end{aligned}$$

Now for the upper bound. Let $\mathcal{B}$ be the event that there are more than $2\log^8 n$ vertex disjoint copies of $(r,(K_{3,3})_r)$ rooted at e. Then

$$\mathbb{P}(\mathcal{H}_e|p_e) \leq \mathbb{P}(\mathcal{B}) + \sum_{\substack{A\in B\\ |A|\leq 2\log^8 n}}(\mathbb{P}(\mathcal{H}_e \cap \mathcal{M}_A|p_e).$$

We use Theorem 2.9 to bound $\mathbb{P}(\mathcal{B})$. Note that the expected number of copies of $(r,(K_{3,3})_r)$ rooted at e in G_{n,p_e} is less than the number of copies rooted at e in G_{n,p_0}. Now there are less than $n^4/4$ possible copies of $(r,(K_{3,4})_r)$ rooted at e and thus the expected number is less than $\log^8 n/4$. Therefore

$$\mathbb{P}(\mathcal{B}) \leq \left(\frac{\mathrm{e}\log^8 n}{8\log^8 n}\right)^{2\log^8 n} \leq \exp(-2\log^8 n) = o(n^{-4}).$$

Fix $A \in B$. Recall that $M_1 = (r,(K_3)_r)$ and $M_2 = ((r,(K_{3,3})_r),(r_1,(K_{3,4}^-)_{r_1}))$. Let

$$I' = \bigcup_{a\in A}\bigcup_{\overline{e}\in E(\overline{K_{3,3}})} I'_{a(\overline{e}),M_1,U} \cup I'_{e,M_2,U}$$

and for $((R,F),f) \in I'$ let $X_{(R,F),f,p_e}$ be the indicator random variable for the event that $f(E(F)) \subseteq E(G_{n,p_e})$. Let

$$X'_{p_e} = \sum_{((R,F),f)\in I'} X_{(R,F),f,p_e}$$

then we have that $(\mathcal{H}_e \cap \mathcal{M}_A|p_e) \subseteq \{X'_{p_e} = 0\} \cap (\mathcal{E}_A|p_e)$ and thus

$$\mathbb{P}(\mathcal{H}_e \cap \mathcal{M}_A|p_e) \leq \mathbb{P}(\{X'_{p_e} = 0\} \cap \mathcal{E}_A|p_e) \leq \mathbb{P}(\{X'_{p_e} = 0\}|p_e)\mathbb{P}(\mathcal{E}_A|p_e)$$

as the events $\mathcal{E}_A$ and $\{X'_{p_e} = 0\}$ are independent. Clearly $\mathbb{P}(\mathcal{E}_A|p_e) = p_e^{8|A|}$ and to bound $\mathbb{P}(\{X'_{p_e} = 0\}|p_e)$ we will use Theorem 2.8. As always the first step is to establish the value of Δ.

$$\begin{aligned}\Delta &\le \sum_{(r_1,F_1),(r_2,F_2)\in M_1\cup M_2}\ \sum_{(g_1,g_2,F)\in(r_1,F_1)\oplus(r_2,F_2)} \binom{|U|}{|S_{g_1,g_2,F}|} n^{v(F)-|S_{g_1,g_2,F}|} p_e^{e(F)} \\ &\le \sum_{(r_1,F_1),(r_2,F_2)\in M_1\cup M_2}\ \sum_{(g_1,g_2,F)\in(r_1,F_1)\oplus(r_2,F_2)} n^{o(1)} n^{v(F)-|S_{g_1,g_2,F}|} p_0^{e(F)} \\ &\le \sum_{(r_1,F_1),(r_2,F_2)\in M_1\cup M_2}\ \sum_{(g_1,g_2,F)\in(r_1,F_1)\oplus(r_2,F_2)} n^{o(1)} n^{f_2(S_{g_1,g_2,F},F)+o(1)} \\ &= o(1)\end{aligned}$$

as according to Lemma 5.4 $f_2(S_{g_1,g_2,F},F) < 0$ for every element of the sum. Note that one can calculate $\mathbb{E}(X'_{p_e})$ similarly to $\mathbb{E}(X_{p_e})$ and thus

$$\mathbb{E}(X'_{p_e}) = \left(1 + O\left(\frac{\log^8 n}{n}\right)\right)\left(6|A|np_e^2 + \frac{n^4}{4}p_e^8 + \frac{n^5}{3}p_e^{10}\right).$$

Therefore

$$\begin{aligned}\mathbb{P}(X'_{p_e} = 0|p_e) &\le (1+o(1))\exp\left(-(1+O(p_0))\mathbb{E}(X_{p_e})\right) \\ &= \exp\left(-(1+O(p_0))\left(6|A|np_e^2 + \frac{n^4}{4}p_e^8 + \frac{n^5}{3}p_e^{10}\right) + o(1)\right) \\ &= (1+o(1))\exp\left(-\frac{n^4}{4}p_e^8 - \frac{n^5}{3}p_e^{10}\right)\left(\exp(-6np_e^2)\right)^{|A|}.\end{aligned}$$

Recall that $|\{A \in B : |A| = k\}| = (1+o(1))(n^{4k}/(4^k k!))$ thus

$$\begin{aligned}\mathbb{P}(\mathcal{H}_e|p_e) &\le o(n^{-4}) + \sum_{k=0}^{2\log^8 n}(1+o(1))\frac{n^{4k}}{4^k k!}p_e^{8k}\exp\left(-\frac{n^4}{4}p_e^8 - \frac{n^5}{3}p_e^{10}\right)\left(\exp(-6np_e^2)\right)^k \\ &= (1+o(1))\exp\left(-\frac{n^4}{4}p_e^8 - \frac{n^5}{3}p_e^{10}\right)\exp\left(\frac{n^4}{4}p_e^8\exp(-6np_e^2)\right) + o(n^{-4})\end{aligned}$$

completing the proof. ∎

Lemma 5.16

Let $H = K_{3,4}^-$ *and*

$$c = \int_0^\infty \exp\left(-\frac{\tau^8}{4} - \frac{\tau^{10}}{3}\right)\exp\left(\frac{\tau^8}{4}\exp(-6\tau^2)\right)d\tau$$

then

$$\mathbb{P}(\mathcal{H}_e) = (1+o(1))cn^{-1/2}.$$

PROOF Let $p_0 = n^{-1/2}\log n$. Similarly to the previous calculations we have that

$$\mathbb{P}(\mathcal{H}_e) = \mathbb{P}(\mathcal{H}_e \cap \{p_e \le p_0\}) + \mathbb{P}(\mathcal{H}_e \cap \{p_e > p_0\}).$$

According to Lemma 5.15

$$\begin{aligned}
&\mathbb{P}(\mathcal{H}_e \cap \{p_e \le p_0\}) = \int_0^{p_0} \mathbb{P}(\mathcal{H}_e|p_e)dp_e \\
&= \int_0^{p_0} (1+o(1)) \exp\left(-\frac{n^4}{4}p_e^8 - \frac{n^5}{3}p_e^{10}\right) \exp\left(\frac{n^4}{4}p_e^8 \exp(-6np_e^2)\right) + o(n^{-4})dp_e \\
&= o(n^{-4}) + (1+o(1))n^{-1/2} \int_0^{\log n} \exp\left(-\frac{\tau^8}{4} - \frac{\tau^{10}}{3}\right) \exp\left(\frac{\tau^8}{4}\exp(-6\tau^2)\right) d\tau \\
&= o(n^{-4}) + (1+o(1))n^{-1/2} \int_0^{\infty} \exp\left(-\frac{\tau^8}{4} - \frac{\tau^{10}}{3}\right) \exp\left(\frac{\tau^8}{4}\exp(-6\tau^2)\right) d\tau.
\end{aligned}$$

Note that this integral exists as $\exp\left(\tau^8 \exp(-6\tau^2)/4\right)$ is bounded from above by a constant.

We also have that

$$\mathbb{P}(\mathcal{H}_e \cap \{p_e > p_0\}) \le \mathbb{P}(\mathcal{H}_e|p_0) \le 2\exp(-\log^{10} n) + o(n^{-4}) = o(n^{-4})$$

completing the proof. ∎

Corollary 5.17
Let $H = K_{3,4}^-$ and

$$c = \int_0^{\infty} \exp\left(-\frac{\tau^8}{4} - \frac{\tau^{10}}{3}\right) \exp\left(\frac{\tau^8}{4}\exp(-6\tau^2)\right) d\tau$$

then

$$\mathbb{E}(e(G_e(K_{3,4}^-)_{n,1})) = (1+o(1))\frac{c}{2}n^{3/2}.$$

5.3 Concentration

In this section we show that the number of edges is a.a.s. concentrated around its expected value in case H is a strictly 2-balanced graph. The proof uses the second moment method and in order to give bounds on the second moment we establish $\mathbb{P}(\mathcal{H}_{e_1} \cap \mathcal{H}_{e_2})$.

Lemma 5.18
Let H be a strictly 2-balanced graph and assume that $p_{e_1} < p_{e_2} \le p_0 = n^{-1/d_2(H)} \log n$ then

$$\mathbb{P}(\mathcal{H}_{e_1} \cap \mathcal{H}_{e_2}|p_{e_1}, p_{e_2}) = (1+o(1)) \exp\left(-\left(p_{e_1}^{e(H)-1} + p_{e_2}^{e(H)-1}\right)\frac{2e(H)n^{v_H-2}}{\text{aut}(H)}\right).$$

PROOF Recall that we have $\mathcal{H}_e$, if for any $r \in E(H)$ no copy of (r, H_r) is rooted at e in G_{n,p_e}.

Let $M = \{(r, H_r) : r \in E(H)\}$ and let $X_{(r,H_r),f,p}$ be the indicator random variable for the event that $f(E(H_r)) \subseteq E(G_{n,p})$. Obviously

$$\mathbb{P}(\mathcal{H}_e|p_e) = \mathbb{P}\left(\sum_{((r,H_r),f)\in I'_{e,M}} X_{(r,H_r),f,p_e} = 0\right).$$

Recall that $|I'_{e,M}| = (1+O(n^{-1}))2e(H)n^{v(H)-2}/\mathrm{aut}(H)$. Define

$$X_1 = \sum_{((r,H_r),f)\in I'_{e_1,M}} X_{(r,H_r),f,p_{e_1}}$$

and X_2 analogously.

We will establish a lower bound on the probability that $X_1 + X_2 = 0$ using Lemma 2.6. The Lemma requires an underlying set and a unique probability assigned to every element in the set. This is achieved by exposing the edges of $G_{n,p_{e_2}}$ in two rounds. First we consider $G_{n,p_{e_1}}$ and then add the edges in $G_{n,p'}$ where $p' = \frac{p_{e_2}-p_{e_1}}{1-p_{e_1}}$ to create $G_{n,p_{e_2}}$. In the setup of Lemma 2.6, Ω is two disjoint copies of the edge set of K_n, denoted by Γ_1 and Γ'. Let E_1 be a random subset of Γ_1 where every edge is chosen independently with probability p_{e_1}, except e_1 and e_2, which are chosen with probability 0. Similarly E' is a random subset of Γ' where every edge is chosen independently with probability p', except for e_1, which is chosen with probability 1 and e_2 which is chosen with probability 0. Note that E_1 is equivalent to $E(G_{n,p_{e_1}})$ and E' is equivalent to $E(G_{n,p'})$ when the assumption $p_{e_1} < p_{e_2}$ is taken into account as this is satisfied when $e_1 \notin E(G_{n,p_{e_1}})$, $e_1 \in E(G_{n,p_{e_2}}) = E(G_{n,p_{e_1}}) \cup E(G_{n,p'})$ and $e_2 \notin E(G_{n,p_{e_2}}) = E(G_{n,p_{e_1}}) \cup E(G_{n,p'})$. Let $((r, H_r), f) \in I'_{e_1,M}$ then $f(E(H_r)) \subseteq E(G_{n,p_{e_1}})$ is equivalent to $f(E(H_r)) \subseteq E_1$. However in the case when $((r, H_r), f) \in I'_{e_2,M}$ then for $f(E(H_r)) \subseteq E(G_{n,p_{e_2}})$, we have to consider all possible $S \subseteq E(H_r)$, such that the edges in $f(S)$ are in E_1 and the remaining edges i.e. $f(E(H_r)\backslash S)$ are in E'. For $S \subseteq E(H_r)$ define $A_{(r,H_r),f,S} = (f(E(F)\backslash S) \cap \Gamma_1)$ and $B_{(r,H_r),f,S} = (f(S) \cap \Gamma')$. Let $Y_{(r,H_r),f,S}$ be the indicator random variables for the event that $A_{(r,H_r),f,S} \subseteq E_1$ and $B_{(r,H_r),f,S} \subseteq E'$. Therefore, we have that

$$\begin{aligned}&\mathbb{P}(X_1 + X_2 = 0|p_{e_1}, p_{e_2})\\ &\qquad = \mathbb{P}\left(\sum_{((r,H_r),f)\in I'_{e_1,M}} Y_{(r,H_r),f,\emptyset} + \sum_{((r,H_r),f)\in I'_{e_2,M}} \sum_{S\subseteq E(H_r)} Y_{(r,H_r),f,S} = 0\right).\end{aligned}$$

Now we apply Lemma 2.6 to the right hand side. We have that $Y_{(r,H_r),f,S}$ is the indicator random variable that $e(H) - 1$ edges in $\Gamma_1 \cup \Gamma'$ appear in $E_1 \cup E'$. Also only one edge appears with probability one, thus

$$\max_{((r,H_r),f)\in I'_{e_1,M}\cup I'_{e_2,M}} \max_{S\subseteq E(F)} \{\mathbb{E}(Y_{(r,H_r),f,S})\} \leq p_0^{e(H)-1} \leq p_0$$

as $e(H) > 2$ for every strictly 2-balanced graph. Therefore:

$$\mathbb{P}\left(\sum_{(r,f)\in I'_{e_1,M}} Y_{(r,H_r),f,\emptyset} + \sum_{((r,H_r),f)\in I'_{e_2,M}} \sum_{S\subseteq E(H_r)} Y_{(r,H_r),f,S} = 0\right)$$

$$\geq \exp\left(-\left(\sum_{((r,H_r),f)\in I'_{e_1,M}} \frac{\mathbb{E}(Y_{(r,H_r),f,\emptyset})}{1-p_0} + \sum_{((r,H_r),f)\in I'_{e_2,M}} \sum_{S\subseteq E(H_r)} \frac{\mathbb{E}(Y_{(r,H_r),f,S})}{1-p_0}\right)\right).$$

For $((r,H_r),f) \in I'_{e_1,M}$ such that $e_2 \in f(E(H_r))$ we have that $\mathbb{E}(Y_{(r,H_r),f,\emptyset}) = 0$ and that this affects at most $O(n^{v(H)-3})$ elements in $I'_{e_1,M}$. For the remaining $((r,H_r),f) \in I'_{e_1,M}$ we have that $\mathbb{E}(Y_{(r,H_r),f,\emptyset}) = p_{e_1}^{e(H)-1}$ and thus

$$\sum_{((r,H_r),f)\in I'_{e_1,M}} \mathbb{E}(Y_{(r,H_r),f,\emptyset}) = (1+O(n^{-1}))\frac{2e(H)n^{v(H)-2}}{\mathrm{aut}(H)} p_{e_1}^{e(H)-1}.$$

Also since $n^{v(H)-2}p_0^{e(H)-1} = n^{o(1)}$ we have that

$$\sum_{((r,H_r),f)\in I'_{e_1,M}} \mathbb{E}(Y_{(r,H_r),f,\emptyset}) = \frac{2e(H)n^{v(H)-2}}{\mathrm{aut}(H)} p_{e_1}^{e(H)-1} + o(1).$$

Note that $\mathbb{E}(X_1|p_{e_1},p_{e_2}) = \sum_{((r,H_r),f)\in I'_{e_1,M}} \mathbb{E}(Y_{(r,H_r),f,\emptyset})$.

Let $((r,H_r),f) \in I'_{e_2,M}$ such that $e_1 \in f(E(H_r))$. Then

$$\sum_{S\subseteq E(H_r)} \mathbb{E}(Y_{(r,H_r),f,S}) = \sum_{f^{-1}(e_1)\subseteq S\subseteq E(H_r)} p_{e_1}^{e(H)-1-|S|} p'^{(|S|-1)} = (p_{e_1}+p')^{e(H)-2}$$

and this affects at most $O(n^{v(H)-3})$ elements in $I'_{e_2,M}$. For the remaining $(r,f) \in I'_{e_2,M}$ we have that

$$\sum_{S\subseteq E(H_r)} \mathbb{E}(Y_{(r,H_r),f,S}) = \sum_{S\subseteq E(H_r)} p_{e_1}^{e(H)-1-|S|} p'^{|S|} = (p_{e_1}+p')^{e(H)-1}.$$

Therefore:

$$\sum_{((r,H_r),f)\in I'_{e_2,M}} \sum_{S\subseteq E(H_r)} \mathbb{E}(Y_{(r,H_r),f,S})$$

$$= (1+O(n^{-1}))\frac{2e(H)n^{v(H)-2}}{\mathrm{aut}(H)}(p_{e_1}+p')^{e(H)-1} + O(n^{v(H)-3})(p_{e_1}+p')^{e(H)-2}.$$

Note that $p_{e_1} + p' = (p_{e_2} - p_{e_1}^2)/(1 - p_{e_1}) = (1 + O(p_{e_1}))(p_{e_2} - p_{e_1}^2) = (1 + O(p_{e_1}))p_{e_2}(1 + O(p_{e_1})) = p_{e_2}(1 + O(p_0))$. Thus we have that $O(n^{v(H)-3})(p_{e_2}(1 + O(p_0)))^{e(H)-2} = O(n^{v(H)-3}p_0^{e(H)-2}) = O((np_0)^{-1}) = o(1)$ as $d_2(H) > 1$ for strictly 2-balanced graphs. Therefore

$$\begin{aligned}
&\sum_{((r,H_r),f)\in I'_{e_2,M}} \sum_{S\subseteq E(H_r)} \mathbb{E}(Y_{(r,H_r),f,S}) \\
&= (1 + O(n^{-1}))\frac{2e(H)n^{v(H)-2}}{\text{aut}(H)} p_{e_2}^{e(H)-1}(1 + O(p_0)) + o(1) \\
&= \frac{2e(H)n^{v(H)-2}}{\text{aut}(H)} p_{e_2}^{e(H)-1} + o(1).
\end{aligned}$$

Note that a similar calculation where $p_{e_1} + p'$ is replaced by p_{e_2} gives us that

$$\mathbb{E}(X_2|p_{e_1}, p_{e_2}) = (1 + O(p_0))\frac{2e(H)n^{v(H)-2}}{\text{aut}(H)} p_{e_2}^{e(H)-1} + o(1).$$

The bound

$$\mathbb{P}(\mathcal{H}_{e_1} \cap \mathcal{H}_{e_2}|p_{e_1}, p_{e_2}) \geq (1 + o(1)) \exp\left(-\left(p_{e_1}^{e(H)-1} + p_{e_2}^{e(H)-1}\right)\frac{2e(H)n^{v_H-2}}{\text{aut}(H)}\right)$$

follows from the fact that $1/(1-p_0) = (1+O(p_0))$, and we have already established that

$$O(p_0)\frac{2e(H)n^{v(H)-2}}{\text{aut}(H)}(p_{e_1}^{e(H)-1} + p_{e_2}^{e(H)-1}) = o(1).$$

We use Theorem 2.7 to establish a matching upper bound. We are considering the indicator random variables $X_{(r,H_r),f,p_{e_1}}$ for $((r, H_r), f) \in I'_{e_1,M}$ and $X_{(r,H_r),f,p_{e_2}}$ for $((r, H_r), f) \in I'_{e_2,M}$, assuming $p_{e_1} < p_{e_2}$. Note that connecting $X_{(r,H_{r_1}),f_1,p}$ and $X_{(r_2,H_{r_2}),f_2,q}$ for $p, q \in \{p_{e_1}, p_{e_2}\}$, when $f_1(E(H_{r_1})) \cap f_2(E(H_{r_2})) \neq \emptyset$, forms a dependency graph for our random variables. The upper bound follows once $\delta = o(1)$ and $\Delta = o(1)$ is established.

Fix $((r_1, H_{r_1}), f_1) \in I'_{e_1,M} \cup I'_{e_2,M}$, then there are at most $O(n^{v(H)-3})$ copies of $((r_2, H_{r_2}), f_2) \in I'_{e_1,M} \cup I'_{e_2,M}$ such that $X_{(r_1,H_{r_1}),f_1,p}$ and $X_{(r_2,H_{r_2}),f_2,q}$ are connected in the dependency graph. Note that $\mathbb{P}(X_{(r,H_r),f,p_{e_1}}|p_{e_1}, p_{e_2}) < \mathbb{P}(X_{r,f,p_{e_2}}|p_{e_1}, p_{e_2}) \leq p_0^{e(H)-2}$ which implies that $\delta = O(n^{v(H)-3})p_0^{e(H)-2} = o(1)$.

Now to show that $\Delta = o(1)$. Define V_e as the set of endvertices of e_1 and e_2. For $((r_1, H_{r_1}), f_1), ((r_2, H_{r_2}), f_2) \in I'_{e_1,M} \cup I'_{e_2,M}$ such that $f_1(E(H_{r_1})) \cap f_2(E(H_{r_2})) \neq \emptyset$, we have that $f_1(H_{r_1}) \cup f_2(H_{r_2})$ is the copy of $(S_{g_1,g_2,F}, F)$ for some $(g_1, g_2, F) \in (r_1, H_{r_1}) \oplus (r_2, H_{r_2})$, rooted at some $S \subseteq V_e$ such that $|S| = |S_{g_1,g_2,F}|$. We distinguish between two cases based on whether $e_1 \in f(E(H_{r_1}) \cup f(E(H_{r_2})))$ or not.

In case $e_1 \in f(E(H_{r_1}) \cup f(E(H_{r_2})))$, assume that $e_1 \in f(E(H_{r_1}))$, which implies that $((r_1, H_{r_1}), f_1) \in I'_{e_2,M}$. There are at most $O(n^{v(H)-3})$ ways to select

$((r_1, H_{r_1}), f_1) \in I'_{e_2,M}$ such that $e_1 \in f(E(H_{r_1}))$. Since this already fixed $v(H) - 2$ vertices of F, there are $O(n^{v(H)-3} n^{v(F)-|S_{g_1,g_2,F}|-v(H)+2}) = O(n^{v(F)-|S_{g_1,g_2,F}|-1})$ ways to select $((r_1, H_{r_1}), f_1), ((r_2, H_{r_2}), f_2) \in I'_{e_1,M} \cup I'_{e_2,M}$ such that the two copies create a copy of $(S_{g_1,g_2,F}, F)$ rooted at some $S \subseteq V_e$ and $e_1 \in f(E(H_{r_1})) \cup f(E(H_{r_2}))$. The same holds when $e_1 \in f(E(H_{r_2}))$. In this case we have that $p_{e_1} \leq p \leq q \leq p_0$ thus

$$\mathbb{E}(X_{(r_1,H_{r_1}),f_1,p} X_{(r_2,H_{r_2}),f_2,q} | p_{e_1}, p_{e_2}) \leq p_0^{e(F)-1}.$$

Similarly when $e_1 \notin f(E(H_{r_1})) \cup f(E(H_{r_2}))$, then there are $O(n^{v(F)-|S_{g_1,g_2,F}|})$ ways to select $((r_1, H_{r_1}), f_1), ((r_2, H_{r_2}), f_2) \in I'_{e_1,M} \cup I'_{e_2,M}$ such that the two copies create a rooted copy of $(S_{g_1,g_2,F}, F)$. Also

$$\mathbb{E}(X_{(r_1,H_{r_1}),f_1,p} X_{(r_2,H_{r_2}),f_2,q} | p_{e_1}, p_{e_2}) \leq p_0^{e(F)}.$$

Since $O(n^{v(F)-|S_{g_1,g_2,F}|-1} p_0^{e(F)-1}) = O(n^{v(F)-|S_{g_1,g_2,F}|} p_0^{e(F)})$ thus

$$\begin{aligned}
\Delta &= \sum_{r_1,r_2 \in E(H)} \sum_{(g_1,g_2,F) \in (r_1,H_{r_1}) \oplus (r_2,H_{r_2})} O(n^{v(F)-|S_{g_1,g_2,F}|} p_0^{e(F)}) \\
&= \sum_{r_1,r_2 \in E(H)} \sum_{(g_1,g_2,F) \in (r_1,H_{r_1}) \oplus (r_2,H_{r_2})} O(n^{v(F)-|S_{g_1,g_2,F}|-e(F)/d_2(H)+o(1)}) \\
&= \sum_{r_1,r_2 \in E(H)} \sum_{(g_1,g_2,F) \in (r_1,H_{r_1}) \oplus (r_2,H_{r_2})} O(n^{f_{d_2(H)}(S_{g_1,g_2,F},F)+o(1)}) \\
&= o(1)
\end{aligned}$$

according to Lemma 5.4 completing the proof. ■

Lemma 5.19

Let H be a strictly 2-balanced graph then we have that:

$$\mathbb{P}(\mathcal{H}_{e_1} \cap \mathcal{H}_{e_2}) = (1 + o(1)) \mathbb{P}(\mathcal{H}_{e_1}) \mathbb{P}(\mathcal{H}_{e_2}).$$

PROOF Let $p_0 = n^{-1/d_2(H)} \log n$ and $\mathcal{A} = [0, p_0] \times [0, p_0]$ then similarly as before:

$$\begin{aligned}
\mathbb{P}(\mathcal{H}_{e_1} \cap \mathcal{H}_{e_2} \cap \{p_{e_1} \leq p_0\} \cap \{p_{e_2} \leq p_0\}) &= \int_{\mathcal{A}} \mathbb{P}(\mathcal{H}_{e_1} \cap \mathcal{H}_{e_2} | p_{e_1}, p_{e_2}) d(p_{e_1}, p_{e_2}) \\
&= (1 + o(1)) \int_0^{p_0} \int_0^{p_0} \mathbb{P}(\mathcal{H}_{e_1} | p_{e_1}) \mathbb{P}(\mathcal{H}_{e_2} | p_{e_2}) dp_{e_1} dp_{e_2} \\
&= (1 + o(1)) \int_0^{p_0} \mathbb{P}(\mathcal{H}_{e_1} | p_{e_1}) dp_{e_1} \int_0^{p_0} \mathbb{P}(\mathcal{H}_{e_2} | p_{e_2}) dp_{e_2} \\
&= (1 + o(1)) \mathbb{P}(\mathcal{H}_{e_1}) \mathbb{P}(\mathcal{H}_{e_2}).
\end{aligned}$$

In Lemma 5.10 we have shown that the probability of inserting e_1 if $p_{e_1} > p_0$ is $o(n^{-4}) = o(\mathbb{P}(\mathcal{H}_{e_1}) \mathbb{P}(\mathcal{H}_{e_2}))$ and similarly for e_2 completing the proof. ■

Theorem 5.20

We have a.a.s. that

$$e(G(H)_{n,1}) = (1+o(1))\frac{1}{2}\,\Gamma\left(\frac{e(H)}{e(H)-1}\right)\Big/\left(\frac{2e(H)}{\mathrm{aut}(H)}\right)^{1/(e(H)-1)} n^{2-1/d_2(H)}.$$

PROOF Let X_e be the indicator random variable for the event $\mathcal{H}_e$ and set $X = \sum_{e\in K_n} X_e$. Corollary 5.11 states that that

$$\mathbb{E}(X) = (1+o(1))\frac{1}{2}\,\Gamma\left(\frac{e(H)}{e(H)-1}\right)\Big/\left(\frac{2e(H)}{\mathrm{aut}(H)}\right)^{1/(e(H)-1)} n^{2-1/d_2(H)},$$

and therefore we only have to prove the concentration. Chebyshev's inequality implies that the statement holds if $\mathrm{Var}(X) = o((\mathbb{E}(X))^2)$. Now

$$\begin{aligned}\mathrm{Var}(X) &\le \mathbb{E}(X) + \sum_{e_1,e_2\in E(K_n)} \mathbb{E}(X_{e_1}X_{e_2}) - \mathbb{E}(X_{e_1})\mathbb{E}(X_{e_2})\\ &= o((\mathbb{E}(X))^2) + \sum_{e_1,e_2\in E(K_n)} \mathbb{P}(\mathcal{H}_{e_1}\cap\mathcal{H}_{e_2}) - \mathbb{P}(\mathcal{H}_{e_1})\mathbb{P}(\mathcal{H}_{e_2})\\ &\le o((\mathbb{E}(X))^2) + n^4 o((\mathbb{P}(\mathcal{H}_e))^2)\\ &= o((\mathbb{E}(X))^2)\end{aligned}$$

as required. ∎

5.4 Independence Number

In the following, we will show bounds on the independence number of $G(H)_{n,1}$ when H is strictly 2-balanced. Erdős [12] has shown that there exist constants c_1,c_2, such that if one removes every triangle from $G_{n,p}$, when $p = c_1/\sqrt{n}$ then a.a.s. every set of size $c_2\sqrt{n}\log n$ still contains an edge. Since this creates a subgraph of $G_e(K_3)_{n,1}$ we have that $\alpha(G_e(K_3)_{n,1}) = O(\sqrt{n}\log n)$. We can generalise this result for any strictly 2-balanced H by modifying a proof of Osthus and Taraz [31].

We will need the following result from Osthus and Taraz [31].

Lemma 5.21 ([31])

Given a graph G and $f \in E(G)$ we say that a sequence $H_1, ..., H_k$ of copies of H in G forms a (k, f, H) cluster if they all contain f and have the property that for all i with $1 \le i \le k$ H_i contains an edge which is not contained in any of the other H_j with $j < i$. Then for strictly 2-balanced H and $p = c_0 n^{-1/d_2(H)}$, where c_0 satisfies $c_0^{e(H)-1} e(H)^2 = 1/100$ we have that $G_{n,p}$ a.a.s. contains no $(\log n, f, H)$ cluster for any edge $f \in E(G_{n,p})$.

Theorem 5.22

For strictly 2-balanced H we have that a.a.s.

$$\alpha(G_e(H)_{n,1}) = O(n^{1/d_2(H)}\log n).$$

PROOF Consider $G_{n,p}$ with $p = c_0 n^{-1/d_2(H)}$, where c_0 is as in Lemma 5.21, and remove every copy of H from it. We will show that a.a.s. every vertex set of size $s = 2cn^{1/d_2(H)} \log n / c_0$, where c will be determined later, still contains an edge and thus this also holds in $G_e(H)_{n,1}$. Define $\mu = \binom{s}{2} p$.

For a set $S \subseteq V(G_{n,p})$, define an auxiliary graph $G_{S,H}$ on $G_{n,p}$. Consider the set of copies of H, such that at least one edge is contained in S. Now fix an ordering on the set and consider the copies in order. For every copy containing an edge in S which has not been covered by previous copies of H, we set it as a vertex in $G_{S,H}$. Two vertices in $G_{S,H}$ are connected if the two copies share an edge. We have that at most $e(H)v(G_{S,H})$ edges are removed from S when we delete every copy of H from the graph. To bound the number of vertices in $G_{S,H}$ we will use the fact that $v(G) \le \alpha(G) + 2\Delta(G)\gamma(G)$ where $\alpha(G)$ is the size of the maximal independent set, $\Delta(G)$ is the maximal degree and $\gamma(G)$ is the size of the largest induced matching on G. Assume that a.a.s. every $S \subseteq V(G_{n,p})$ such that $|S| = s$ spans at least $\mu/2$ edges, $\alpha(G_{S,H}) \le \mu/(8e(H))$ and $\gamma(G_{S,H}) \le \mu/(16(e(H))^2 \log n)$. Lemma 5.21 implies that a.a.s. $\Delta(G_{S,H}) \le e(H) \log n$ holds for every S such that $|S| = s$ and thus $v(G_{S,H}) \le \mu/(8e(H)) + 2e(H)(\log n)\mu/(16e(H)^2 \log n) = \mu/(4e(H))$. Then we have a.a.s. at least $\mu/2 - e(H)v(G_{S,H}) \ge \mu/4$ edges left in any set S with $|S| = s$.

Now we will show that our assumptions hold. Fix $S \subseteq V(G_{n,p})$ such that $|S| = s$ and let X_S denote the number of edges spanned by S. We have that $\mu = \binom{s}{2} p = (1 + o(1))cs \log n$. Applying the Chernoff bound gives that:

$$\mathbb{P}(X_S \le \mu/2) \le \exp(-\mu/8) \le n^{-s},$$

if $c > 8$.

Let $Y_{S,H}$ denote the maximal number of edge disjoint copies of H with at least two vertices in S. We have that $\alpha(G_{S,H}) \le Y_{S,H}$. Note that the expected number of copies of H with at least two vertices in S is

$$\mu_{S,H} \le n^{v(H)-2} s^2 p^{e(H)} = c_0^{e(H)-1} s^2 p \le \frac{\mu}{8\mathrm{e}^2 e(H)},$$

due to our choice of c_0. Then Theorem 2.9 gives us that

$$\mathbb{P}(Y_{S,H} \ge \mu/8e(H)) \le \left(\frac{\mathrm{e}\mu_{S,H}}{\mu/8e(H)} \right)^{\mu/8e(H)} \le \exp(-\mu/8(e(H))) \le n^{-s},$$

if $c > 8e(H)$.

Note that an induced matching in $G_{S,H}$ is equivalent to a set of pairs of copies of H where the two elements in a pair overlap, but any two pairs are edge disjoint. Thus we wish to determine a maximal set of edge disjoint graphs, such that every graph in this set is contained in $H \oplus H = (\emptyset, H) \oplus (\emptyset, H)$. Consider the copies of graphs F such that $(g_1, g_2, F) \in (H \oplus H)$ and at least three vertices of the copy are in S. For a given $(g_1, g_2, F) \in (H \oplus H)$ the expected number of such copies is $\mu_{S,F} \leq n^{v(F)-3} s^3 p^{e(F)}$. Note that if $T_{g_1,g_2,F}$ contains more than 2 vertices, then, since the graph is strictly 2-balanced $n^{v(T_{g_1,g_2,F})-2} p^{e(T_{g_1,g_2,F})-1} = \Omega(n^{(v(T_{g_1,g_2,F})-2)-(e(T_{g_1,g_2,F})-1)/d_2(H)}) = \Omega(n^{\varepsilon})$ for some $\varepsilon > 0$. Otherwise $T_{g_1,g_2,F}$ is an edge and $n^{v(T_{g_1,g_2,F})-2} p^{e(T_{g_1,g_2,F})-1} = 1$. Therefore the expected number of overlapping pairs of H is:

$$\begin{aligned}
\mu_\gamma = \sum_{(g_1,g_2,F) \in H \oplus H} \mu_{S,F} &\leq \sum_{(g_1,g_2,F) \in H \oplus H} n^{v(F)-3} s^3 p^{e(F)} \\
&\leq \sum_{(g_1,g_2,F) \in H \oplus H} n^{2v(H)-v(T_{g_1,g_2,F})-3} s^3 p^{2e(H)-e(T_{g_1,g_2,F})} \\
&\leq \sum_{(g_1,g_2,F) \in H \oplus H} c_0^{2(e(H)-1)} n^{1-v(T_{g_1,g_2,F})} s^3 p^{2-e(T_{g_1,g_2,F})} \\
&\leq \sum_{(g_1,g_2,F) \in H \oplus H} \frac{1}{10^4 e(H)^4} \frac{s^3 n^{-1} p}{n^{v(T_{g_1,g_2,F})-2} p^{e(T_{g_1,g_2,F})-1}}.
\end{aligned}$$

Since there are at most $e(H)^2$ elements $(g_1, g_2, F) \in H \oplus H$, such that T is an edge thus $\mu_\gamma \leq (1 + o(1)) s^3 p n^{-1}$. Similarly, as before, Theorem 2.9 gives a bound on the probability that a maximal edge disjoint subset is large. Since $d_2(H) > 1$ when H is strictly 2-balanced, we have that

$$\mathbb{P}\left(\gamma(G_{S,H}) \geq 2s\right) \leq \left(\frac{s^3 p n^{-1}}{2s}\right)^{2s} = \left(\left(\frac{sc \log n}{n}\right)^2\right)^s = o\left(\left(\frac{s}{en}\right)^s\right).$$

If $c > 32(e(H))^2$ then $2s < \mu/(16(e(H))^2 \log n)$. Set $c > 32(e(H))^2$ and our assumptions follow after applying the union bound. ∎

A lower bound, matching the upper bound up to a constant factor in the case of the triangle-elimination process follows from a Theorem of Shearer [40]. This states that every triangle free graph with average degree d, has independent set of size at least $n(d \ln d - d + 1)/(d-1)^2$ and in our case $d = \Omega(\sqrt{n})$, and so $\alpha(G(H)_{n,1}) = \Omega(n^{3/2} \log n / n) = \Omega(\sqrt{n} \log n)$. We can show a lower bound however, only when $d_2(H) > 2$ and in these cases it only matches the upper bound up to a $\log \log$ factor. Note that $d_2(H) > 2$, for every complete graph on at least 4 vertices.

Theorem 5.23

For strictly 2-balanced H such that $d_2(H) > 2$ we have that a.a.s.

$$\alpha(G_e(H)_{n,1}) = \Omega\left(n^{1/d_2(H)} \frac{\log n}{\log \log n}\right).$$

PROOF Consider $p = cn^{-1/d_2(H)}(\log\log n)^{1/(e(H)-1)}$ where c satisfies the equality $e(H)c^{e(H)-1}/\mathrm{aut}(H) = 1$. Define $p_0 = n^{-1/d_2(H)}\log n$. For a set of vertices $S \subseteq V(G_{n,p})$ let X_S be the indicator random variable for the event that S forms an independent set in $G_{n,p}$. Also let $Y_{u,v}$ be the indicator random variable for the event that $\{u,v\} \in E(G_e(H)_{n,p_0})$ and let $Y_S = \sum_{u,v\in S} Y_{u,v}$. Note that if S is independent in $G_{n,p}$ then S is also independent in $G(H)_{n,p}$ and the set S' which is created from S by removing either u or v for every pair of vertices such that $Y_{u,v} = 1$ is also independent. Also S' will be a.a.s. independent in $G_e(H)_{n,1}$ as a.a.s. no more edges are inserted. In addition, if $|S| = n^{1/d_2(H)}\frac{\log n}{\log\log n}$ and $Y_S \le |S|/2$, then $|S'| \ge |S|/2$. Fix $S_1, S_2 \subseteq V(G_{n,p})$ such that $|S_1| = |S_2| = n^{1/d_2(H)}\frac{\log n}{\log\log n}$ and condition on the event that both S_1 and S_2 are independent in $G_{n,p}$. Since this conditioning affects at most $O(n^{v(H)-3}(|S_1|+|S_2|))$ possible copies of (r, H_r) rooted at $\{u,v\}$ therefore the results of Lemma 5.9 still hold for q such that $p \le q \le p_0$. Thus for $u, v \in S_1$ and similarly for $u, v \in S_2$ we have that:

$$\begin{aligned}\mathbb{P}(Y_{u,v} = 1|\{X_{S_1}X_{S_2} = 1\}) &= (1+o(1))\int_p^{p_0} \exp\left(-\frac{2e(H)q^{e(H)-1}n^{v_H-2}}{\mathrm{aut}(H)}\right)dq \\ &\le (1+o(1))p_0\exp\left(-\frac{2e(H)p^{e(H)-1}n^{v_H-2}}{\mathrm{aut}(H)}\right) \\ &= (1+o(1))p_0\exp\left(-2\log\log n\right) \\ &= (1+o(1))n^{-1/d_2(H)}/\log n.\end{aligned}$$

Therefore $\mathbb{E}(Y_{S_1}|\{X_{S_1}X_{S_2} = 1\}) = o(|S_1|)$ and thus $\mathbb{P}(\{Y_{S_1} \ge |S|/2\}|\{X_{S_1}X_{S_2} = 1\}) = o(1)$. Due to symmetry $\mathbb{P}(\{Y_{S_2} \ge |S|/2\}|\{X_{S_1}X_{S_2} = 1\}) = o(1)$. Let Z_S be the indicator random variable that $X_S = 1$ and $Y_S \le |S|/2$. Then $\mathbb{P}(Z_{S_1}Z_{S_2} = 1) = (1+o(1))\mathbb{P}(X_{S_1}X_{S_2} = 1)$ and in the special case when $S_1 = S_2$ then $\mathbb{P}(Z_{S_1} = 1) = (1+o(1))\mathbb{P}(X_{S_1} = 1)$.

We will show that there exists an independent set in $G_e(H)_{n,1}$ of size $k = n^{1/d_2(H)}\log n/\log\log n$ by applying the second moment method to the random variable $Z_k = \sum_{S\subseteq V(G_{n,p}),|S|=k} Z_S$. Define $X_k = \sum_{S\subseteq V(G_{n,p}),|S|=k} X_S$, we have shown that $\mathbb{E}(Z_k) = (1+o(1))\mathbb{E}(X_k)$ and that $\mathbb{E}(Z_k^2) = (1+o(1))\mathbb{E}(X_k^2)$. Therefore

$$\mathbb{P}(Z_k > 0) \ge \frac{(\mathbb{E}(Z_k))^2}{\mathbb{E}(Z_k^2)} = (1+o(1))\frac{(\mathbb{E}(X_k))^2}{\mathbb{E}(X_k^2)}.$$

The statement

$$\frac{(\mathbb{E}(X_k))^2}{\mathbb{E}(X_k^2)} = 1 + o(1)$$

follows from the proof of Lemma 7.2 in Janson, Łuczak and Ruciński [26], but for completeness we include a short adaptation of their proof. Note that when $d_2(H) > 2$ then $k = o(\sqrt{n}/\log n)$. We have that $(\mathbb{E}(X_k))^2/\mathbb{E}(X_k^2) = 1 + o(1)$ iff $\mathbb{E}(X_k^2)/(\mathbb{E}(X_k))^2 = 1 + o(1)$. We have that $\mathbb{E}(X_k^2)/(\mathbb{E}(X_k))^2 > 1$ so we only have to give an upper bound. Now

$$\begin{aligned}\frac{\mathbb{E}(X_k^2)}{(\mathbb{E}(X_k))^2} &= \frac{\binom{n}{k}(1-p)^{\binom{k}{2}}\sum_{i=0}^{k}\binom{k}{i}\binom{n-k}{k-i}(1-p)^{\binom{k}{2}-\binom{i}{2}}}{\left(\binom{n}{k}(1-p)^{\binom{k}{2}}\right)^2}\\ &\leq \sum_{i=0}^{k}\frac{\binom{k}{i}\binom{n-k}{k-i}(1-p)^{-\binom{i}{2}}}{\binom{n}{k}}.\end{aligned}$$

Let

$$a_i = \frac{\binom{k}{i}\binom{n-k}{k-i}(1-p)^{-\binom{i}{2}}}{\binom{n}{k}}$$

and

$$b_i = \frac{a_{i+1}}{a_i} = \frac{(k-i)^2}{(i+1)(n-2k+i+1)}(1-p)^{-i}.$$

Note that $b_i < 1$ otherwise the following would have to hold

$$\begin{aligned}(k-i)^2(1-p)^{-i} &\geq (i+1)(n-2k+i+1)\\ k^2\exp(pk) &\geq (1+o(1))n\\ n^{2/d_2(H)+o(1)}\exp(o(\log n)) &\geq (1+o(1))n\\ n^{2/d_2(H)+o(1)} &\geq (1+o(1))n\end{aligned}$$

which leads to a contradiction as $d_2(H) > 2$. Therefore the value of a_i decreases. Note that for every $i \leq \log n$

$$a_i = \frac{\binom{k}{i}\binom{n-k}{k-i}(1-p)^{-\binom{i}{2}}}{\binom{n}{k}} \leq k^i\frac{\binom{n-k}{k-i}}{\binom{n}{k}}\exp(p\log^2 n) \leq \left(\frac{2k}{n}\right)^i.$$

Thus

$$\sum_{i=0}^{k} a_i \leq \sum_{i=0}^{k}\left(\frac{2k}{n}\right)^i + k\left(\frac{2k}{n}\right)^{\log n} = \frac{1}{1-o(1)} + o(1) = 1 + o(1)$$

completing the proof. ■

Note that this result implies that the graph created by the H-elimination process a.a.s. has a larger independence number than the graph created by the H-free process, when H is a complete graph. Therefore it will give worse lower bounds for the off-diagonal Ramsey number. This is not unexpected as the H-elimination random graph process creates a.a.s. less edges than the H-free random graph process and thus it is expected to have a larger independence number.

5.5 Subgraphs

Finally we will show that graphs with $m(F) < d_2(H)$ are a.a.s. present in the H removal process, while graphs with $m(F) > d_2(H)$ are a.a.s. not present. The existence of subgraphs in the H removal process is proven similarly to the existence of subgraphs in $G_{n,p}$, which was shown by Bollobás [7] using the second moment method. However our proof follows a simpler version by Ruciński and Vince [37].

Theorem 5.24

Every finite graph F with $m(F) < d_2(H)$ is a.a.s. present in the H-elimination process for strictly 2-balanced H.

PROOF Define $d_2'(H) = \max_{H' \subsetneq H} d_2(H')$ and let a satisfy $\max\{d_2'(H), m(F)\} < a < d_2(H)$. Instead of showing that F is present in $G_e(H)_{n,1}$ we will show that it is already present in $G_e(H)_{n,p}$ when $p = n^{-1/a}$. Let X_F denote the number of copies of F in $G_{n,p}$ then the second moment method implies that

$$\mathbb{P}(X_F = 0) \leq \frac{\text{Var}(X_F)}{(\mathbb{E}(X_F))^2}.$$

Again we use the fact that a graph is actually a rooted graph, where the roots are given by the empty set. Let $M = \{(\emptyset, F)\}$ and let $I_F' = \{f : ((\emptyset, F), f) \in I_{\emptyset,M}'\}$. Obviously $|I_F'| = (1+o(1))n^{v(F)}/\text{aut}(F)$. For $f \in I_F'$ let X_f be the indicator random variable that all of the edges in $f(E(F))$ are present in $G_e(H)_{n,p}$. We establish a lower bound on $\mathbb{P}(X_f = 1)$ by considering the appearance of an induced copy in $G_{n,p}$ where none of the edges of the induced copy are contained in a copy of H in $G_{n,p}$. Note that the second condition holds if for every $R \subsetneq V(H)$ such that R spans an edge in H, and every $S \subsetneq f(V(F))$ such that $|R| = |S|$ there is no copy of (R, H_R) rooted at S and outside $f(V(F))$ in $G_{n,p}$. Applying Theorem 2.6 gives us that

$$\mathbb{P}(X_f = 1) \geq p^{e(F)}(1-p)^{\binom{v(F)}{2}} \exp\left(- \sum_{\substack{R \subsetneq V(H) \\ e(H[R])>0}} \sum_{\substack{S \subsetneq f(V(H)) \\ |S|=|R|}} \frac{n^{v(H)-|R|}p^{e(H)-e(H[R])}}{1-p} \right).$$

Note that with our choice of p we have for every $R \subsetneq V(H)$ such that R spans an edge in H, that $n^{v(H[R])-2}p^{e(H[R])-1} \geq 1$ and $n^{v(H)-2}p^{e(H)-1} = o(1)$. Therefore

$$\frac{n^{v(H)-|R|}p^{e(H)-e(H[R])}}{1-p} = O\left(\frac{n^{v(H)-2}p^{e(H)-1}}{n^{(|R|-2)}p^{(e(H[R])-1)}} \right) = O\left(n^{v(H)-2}p^{e(H)-1} \right) = o(1),$$

and thus $\mathbb{P}(X_f = 1) \geq (1 + o(1))p^{e(F)}$. Obviously $\mathbb{P}(X_f = 1) \leq p^{e(F)}$. Now $\mathbb{E}(X) = \sum_{f \in I'} \mathbb{E}(X_f) = \Theta(n^{v(F)} p^{e(F)})$. On the other hand:

$$\begin{aligned} \mathrm{Var}(X) &= \sum_{f_1, f_2 \in I'_F} \mathbb{E}(X_{f_1} X_{f_2}) - \mathbb{E}(X_{f_1})\mathbb{E}(X_{f_2}) \\ &\leq \sum_{\substack{T \subseteq F \\ e(T) > 0}} n^{2v(F) - v(T)} p^{2e(F) - e(T)} \\ &= o(n^{2v(F)} p^{2e(F)}) \end{aligned}$$

Which is $o((\mathbb{E}(X_F))^2)$, completing the proof. ∎

The fact that no copies of graphs denser than $d_2(H)$ are present follows trivially when comparing to $G_{n,p}$, with $p = n^{-1/d_2(H)} \log n$.

Observation 5.25

Every finite graph F with $m(F) > d_2(H)$ is a.a.s. not present in the H-removal process when H is strictly 2-balanced.

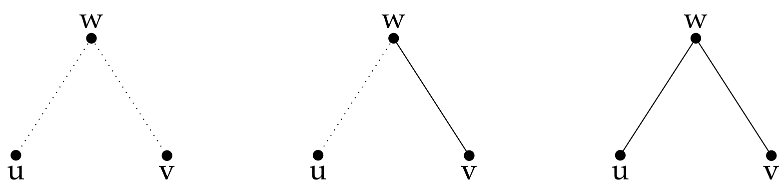
PROOF Follows from a Theorem of Erdős and Rényi [14] that no copy of F with $m(F) > d_2(H)$ is contained in $G_{n,p}$ when $p = n^{-1/d_2(H)} \log n$. ∎

Chapter 6

Subgraphs of the triangle-free random graph process

In Chapter 4 we discussed results on the H-free random graph process. The main focus was on determining the number of edges and the independence number of the H-free random graph process. In this chapter we will focus on the subgraphs created by the process. Wolfovitz [50] has shown for the triangle-free process and Bohman and Keevash [6] have shown for the H-free random graph process, when H is strictly 2-balanced, that the random variables tracked described in detail in Chapter 4 imply that the H-free random graph process contains every finite graph F as a subgraph when $m(F) \leq d_2(H)$. It also implies that as long as the random variables are tracked no denser subgraphs appear. However it is still possible for denser subgraphs to appear during the later stages of the process. In this chapter we show that very dense graphs do not appear in the triangle-free random graph process i.e. there exists a constant c such that no copy of any finite graph with $d(F) > c$ is present in $G_f(K_3)_n$. This is joint work with Stefanie Gerke [20].

In this chapter we will use the slightly simpler notation used by Bohman [4]. A non-edge is open at step i if inserting it into $G_f(K_3)_{n,i}$ would not create a copy of a triangle, otherwise it is closed. The set of open pairs in step i is denoted by $O(i)$ and the set of closed pairs is denoted by $C(i)$. In the original version of the proof the number of open pairs and for any pair of non-adjacent vertices u, v, the number of open, partial and complete vertices were tracked.



w open w partial w complete

Figure 6.1: open/partial/complete vertices

w is open/partial/complete with respect to $\{u, v\}$.
Dotted lines indicate open pairs, continuous lines indicate edges

A vertex w is open with respect to $\{u, v\}$ if both pairs $\{u, w\}$ and $\{v, w\}$ are open. A vertex w is partial with respect to $\{u, v\}$ if exactly one of $\{u, w\}$ and $\{v, w\}$

is open and the other is an edge. Finally w is complete with respect to $\{u, v\}$ if it is a mutual neighbour of $\{u, v\}$, see Figure 6.1. Although these random variables can be constructed from the random variables described in Chapter 4 due to the simplicity of the triangle slightly tighter bounds exist.

Our proof requires only the number of open pairs at step i, denoted by $Q(i)$, and the set of partial vertices with respect to $\{u, v\}$ at step i denoted by $Y_{u,v}(i)$. The following bounds were proven by Bohman [4] for the first $\mu n^{3/2}\sqrt{\log n}$ steps. (Bohman sets $\mu = 1/32$, however no effort was made to optimise the value.) For the remainder of this chapter we set $\mu = 1/32$ and $m = \mu n^{3/2}\sqrt{\log n}$.

Definition 6.1

Let $\mathcal{H}(i)$ be the event that the following bounds hold for all $j \leq i$ and for all pairs $\{u, v\} \notin E(j)$:

$$|Q(j) - n^2 q(t(j))| \leq n^2 g_q(t(j))$$
$$\Big| |Y_{u,v}(j)| - \sqrt{n} y(t(j)) \Big| \leq \sqrt{n} g_y(t(j))$$
$$|Z_{u,v}(j)| \leq \log^2 n$$

where

$$t(i) = i/n^{3/2}$$
$$q(t) = \exp(-4t^2)/2$$
$$y(t) = 4t\exp(-4t^2)$$
$$g_q(t) = \begin{cases} \exp(41t^2 + 40t)n^{-1/6} & : t \leq 1 \\ \frac{\exp(41t^2+40t)}{t} n^{-1/6} & : t > 1 \end{cases}$$
$$g_y(t) = \exp(41t^2 + 40t)n^{-1/6}.$$

Theorem 6.1 ([4])

The event $\mathcal{H}(m)$ holds a.a.s..

Let $W \subseteq V(G_f(K_3)_{n,i})$ and define $e_W(i)$ as the number of edges spanned by W after step i.

Lemma 6.2

Fix k. Let $\mathcal{S}_k(i)$ be the event that there exists $W \subseteq V(G_f(K_3)_{n,i})$ with $|W| = k$ and $e_W(i) \geq 3k$. Then a.a.s. $\mathcal{S}_k(m)$ does not hold.

Note that since $\overline{\mathcal{S}_k(m)}$ is a decreasing property it implies $\overline{\mathcal{S}_k(i)}$ for every $i < m$.

A sharper result showing that a.a.s. $\frac{e_W(m)}{|W|} \leq 2$ for any finite set of vertices W can be found in [6] and [50], however for completeness a short proof of the result has been included.

PROOF Fix k vertices in V and denote this set by W. Let $\mathcal{A}_i$ be the event that an edge is added between two vertices in W at step i.

Then since $q(t(i)) > q(t(m))$ for $i < m$ we have

$$\mathbb{P}(\mathcal{A}_{i+1}|\mathcal{H}(i)) \leq \frac{k^2}{Q(i)} \leq \frac{k^2}{n^2(q(t(i)) - g_q(t(i)))} \leq \frac{2k^2}{n^2 q(t(i))}$$

$$\leq \frac{2k^2}{n^2 q(t(m))} \leq \frac{4k^2}{n^2 \exp(-4\mu^2 \log n)} = \frac{4k^2}{n^{2-4\mu^2}}.$$

Note that these estimates hold irrespective of the events $\mathcal{A}_j$ occurring when $j \neq i$. Since $\mathbb{P}(\mathcal{A}_{i+1} \cap \mathcal{H}(i)) \geq \mathbb{P}(A_{i+1} \cap \mathcal{H}(m))$ and $\mathcal{H}(m)$ holds a.a.s. thus

$$(1 + o(1))\mathbb{P}(\mathcal{A}_{i+1}|\mathcal{H}(i)) \geq \mathbb{P}(\mathcal{A}_{i+1}|\mathcal{H}(m))$$

Thus conditioning on $\mathcal{H}(m)$, we have that the random variable $e_W(m)$ is dominated by Bin $\left(m, \frac{(1+o(1))4k^2}{n^{2-4\mu^2}}\right)$ therefore

$$\mathbb{P}(e_W(m) \geq 3k|\mathcal{H}(m)) \leq \binom{m}{3k}\left((1 + o(1))\frac{4k^2}{n^{2-4\mu^2}}\right)^{3k} \leq \left(\frac{(1 + o(1))\mathrm{e}m4k^2}{3kn^{2-4\mu^2}}\right)^{3k}$$

$$\leq \left(\frac{(1 + o(1))\mathrm{e}\mu n^{3/2}\sqrt{\log n}4k}{3n^{2-4\mu^2}}\right)^{3k} = o\left(\frac{1}{n^{3k/2-20k\mu^2}}\right).$$

Since there are $\binom{n}{k}$ ways to select k vertices, it follows from the union bound that

$$\mathbb{P}(\mathcal{S}_k(m)|\mathcal{H}(m)) \leq \binom{n}{k} o\left(\frac{1}{n^{3k/2-20k\mu^2}}\right) = o\left(\frac{n^k}{n^{3k/2-20k\mu^2}}\right) = o(1)$$

as μ^2 is sufficiently small. ∎

The above proof shows that no copy of a dense graph appears in the process while the first m edges are taken. We will now show that when m edges have been taken at least one edge of any placement of a dense graph F is closed.

Theorem 6.3

Let $\mathcal{T}$ be the event that there exists a copy of a finite graph F satisfying $10v(F)/\mu^2 \leq e(F)$ in the triangle-free graph process. Then a.a.s. $\mathcal{T}$ does not hold.

PROOF Fix a set of vertices $W \subset V(G_f(K_3)_{n,i})$ with $|W| = v(F)$, and a set of pairs of vertices $E_F \subseteq W \times W$ such that if the pairs in E_F were inserted as edges they would form a copy of F on W. Let $\mathcal{C}_F(i)$ be the event that at least one pair in E_F is closed after step i and $\mathcal{O}_F(i)$ be the event that none is closed after step i. For the following assume we are in the event $\mathcal{O}_F(i)$.

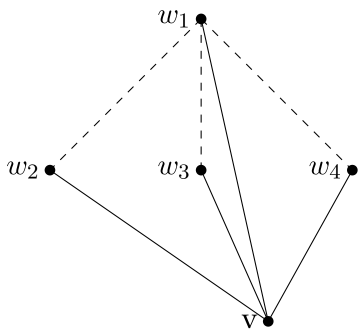


Figure 6.2: Closing multiple open pairs

The edge $\{v, w_1\}$ closes $\{w_1, w_2\}$,$\{w_1, w_3\}$ and $\{w_1, w_4\}$

Note that a pair $\{u, v\}$ is closed at step i if and only if there is a partial vertex $w \in Y_{u,v}(i)$ and the missing edge is chosen. Thus the probability of closing a pair $s \in O_i$ is $|Y_s(i)|/Q(i)$. The problem is that an edge can close several pairs of vertices in E_F. The subset of $W \times W$ closed by $\{w_j, v\} \in O_i$, with $v \notin W$ is at most $w_j \times (N_i(v) \cap W)$ (see Figure 6.2), where $N_i(v)$ denotes the neighbourhood of v in $G_f(K_3)_{n,i}$.

Let D_i be the set of vertices not in W that have more than 6 neighbours in W at time i. Excluding the pairs with both vertices in W and the pairs with a vertex in D_i the remaining pairs close at most 6 pairs in $W \times W$ and in particular at most 6 pairs in E_F. Therefore $\sum_{f \in E_F(W) \setminus E_i} |Y_f(i) \setminus (D_i \cup W)|$ counts any pair that closes a pair in E_F at most 6 times.

Assume we are in $\overline{\mathcal{S}_{2v(F)}(i)}$, then the set D_i can have size at most $v(F)$ otherwise $W \cup D_i$ forms a set of $2v(F)$ vertices that span more then $6v(F)$ edges. Let $\mathcal{B}(i) = \overline{\mathcal{S}_{v(F)}(i)} \cap \overline{\mathcal{S}_{2v(F)}(i)} \cap \mathcal{H}(i)$. Hence

$$\begin{aligned}\mathbb{P}(\mathcal{C}_F(i+1)|[\mathcal{O}_F(i) \cap \mathcal{B}(i)]) &\geq \frac{\sum_{f \in E_F \setminus E_i} |Y_f(i) \setminus (D_i \cup W)|}{6\,Q(i)} \\ &\geq \frac{\sum_{f \in (E_F) \setminus E_i} (|Y_f(i)| - 2v(F))}{6\,Q(i)}.\end{aligned}$$

Since we are in the event $\overline{\mathcal{S}_{v(F)}(i)}$ there are at most $3v(F)$ edges in E_F also $|E_F| \geq 10v(F)/\mu^2$ and because $\mathcal{O}_F(i)$ holds every non-adjacent pair in E_F is open therefore the sum is over at least $(10/\mu^2 - 3)v(F) \geq 9v(F)/\mu^2$ open pairs. Thus

$$\mathbb{P}(\mathcal{C}_F(i+1)|[O_F(i) \cap \mathcal{B}(i)]) \geq \frac{9v(F)}{\mu^2} \frac{\sqrt{n}(y(t(i)) - g_y(t(i))) - 2v(F)}{6n^2(q(t(i)) + g_q(t(i)))}.$$

If n is large enough then $q(t(i)) + g_q(t(i)) \leq 2q(t(i))$, and for $m \geq i \geq n^{4/3}$ we have

$$y(t(i)) - g_y(t(i)) \geq \frac{y(t(i))}{2} \geq 2t(n^{4/3}) \exp(-4t^2(m)) = 2n^{-\frac{1}{6} - 4\mu^2},$$

therefore since $v(F)$ is a constant:

$$\frac{\sqrt{n}y(t(i))}{2} - 2v(F) \geq \frac{7}{15}\sqrt{n}y(t(i)),$$

and so:

$$\begin{aligned}
\mathbb{P}(\mathcal{C}_F(i+1)|[\mathcal{O}_F(i) \cap \mathcal{B}(i)]) &\geq \frac{9v(F)}{\mu^2}\frac{7\sqrt{n}y(t(i))/15}{6n^2(q(t(i))+g_q(t(i)))} \\
&\geq \frac{7v(F)}{\mu^2}\frac{\sqrt{n}y(t(i))}{20n^2q(t(i))} = \frac{7v(F)}{\mu^2}\frac{4t(i)\exp(-4t^2(i))}{10n^{3/2}\exp(-4t^2(i))} \\
&= \frac{14v(F)i}{5\mu^2n^3}.
\end{aligned}$$

It follows that for $m \geq i \geq n^{4/3}$ and sufficiently large n,

$$\mathbb{P}(\mathcal{O}_F(i+1)|[\mathcal{O}_F(i) \cap \mathcal{B}(i)]) \leq 1 - \frac{14v(F)i}{5\mu^2n^3} \leq \exp\left(-\frac{14v(F)i}{5\mu^2n^3}\right).$$

Using $\mathcal{O}_F(i) \subset \mathcal{O}_F(i+1)$ and that $\mathcal{B}(i) \subset \mathcal{B}(i+1)$ thus for sufficiently large n

$$\begin{aligned}
\mathbb{P}(\mathcal{O}_F(m) \cap \mathcal{B}(m)) &= \prod_{i=0}^{m-1} P(\mathcal{O}_F(i+1) \cap \mathcal{B}(i+1)|[\mathcal{O}_F(i) \cap \mathcal{B}(i)]) \\
&\leq \prod_{i=0}^{m-1} P(\mathcal{O}_F(i+1)|[\mathcal{O}_F(i) \cap \mathcal{B}(i)]) \\
&\leq \prod_{i=\lceil n^{4/3}\rceil}^{m-1} \exp\left(-\frac{14v(F)i}{5\mu^2n^3}\right) = \exp\left(\sum_{i=\lceil n^{4/3}\rceil}^{m-1} -\frac{14v(F)i}{5\mu^2n^3}\right) \\
&= \exp\left(-\frac{14v(F)}{5\mu^2n^3}\left(\frac{m(m-1)}{2} - \frac{\lceil n^{4/3}\rceil(\lceil n^{4/3}\rceil-1)}{2}\right)\right) \\
&\leq \exp\left(-\frac{4v(F)}{3}\frac{m^2}{\mu^2n^3}\right) = \exp\left(-\frac{4v(F)}{3}\log n\right) = n^{-4v(F)/3}.
\end{aligned}$$

As there are $\binom{n}{v(F)}\frac{v(F)!}{aut(F)}$ possible placements of F applying the union bound gives

$$P(\mathcal{T} \cap \mathcal{B}(m)) \leq \binom{n}{v(F)}v(F)!\, n^{-4k/3} \leq n^{v(F)-4v(F)/3} = o(1).$$

■

Note that there is an alternative way to estimate $\mathbb{P}(\mathcal{C}_F(i+1)|[\mathcal{O}_F(i) \cap \mathcal{B}(i)])$. As mentioned earlier the vertices with more than one neighbour in W cause a problem. However the total number of such vertices is small as each of them is a complete vertex with respect to $\{u, v\}$ for some $u, v \in W$. Bohman [4] showed that the number of complete vertices with respect to any pair of vertices $\{u, v\}$ is at most $\log^2 n$ and this implies that

$$\mathbb{P}(\mathcal{C}_F(i+1)|[\mathcal{O}_F(i) \cap \mathcal{B}(i)]) \geq \frac{\sum_{f \in E_F(W)\setminus E_i} |Y_f(i)| - v(F)^2 \log^2 n}{Q(i)}.$$

This was one of the observations in Warnke [48] where he generalised the result for strictly 2-balanced H.

Theorem 6.4 ([48])
For every strictly 2-balanced graph H there exist constants $c, d > 0$ such that a.a.s. no copy of any graph F with $1 \leq v(F) \leq n^d$ and $m(F) > c$ is present in the H-free random graph process.

Chapter 7

The $K_{3,4}^{-}$-free process

7.1 The differential equations

In this section we are considering the $K_{3,4}^{-}$-free random graph process. We will show that the process contains a.a.s. $\Omega(n^{3/2}\sqrt{\log\log n})$ edges. This improves the previous lower bound by Osthus and Taraz [31] who have shown that the process contains $\Omega(n^{3/2})$ edges.

Recall that the graph $K_{3,4}^{-}$ is 2-balanced, but not strictly 2-balanced as $d_2(K_{3,4}^{-}) = d_2(K_{3,3}) = 2$ and $K_{3,3}$ is the only proper subgraph of $K_{3,4}^{-}$ which has 2-density of 2 the remaining subgraphs have a smaller 2-density.

As in Chapter 4 we call a pair of non-adjacent vertices which can be connected by an edge in $G_f(K_{3,4}^{-})_{n,i}$ without creating a copy of $K_{3,4}^{-}$ an open pair at step i. The set of open pairs at step i are denoted by $O(i)$ and $Q(i) = |O(i)|$. Otherwise a pair of non-adjacent vertices is called a closed pair and the set of closed pairs is denoted by $C(i)$. Recall that we are interested in the rooted graphs $(r_1, (K_{3,4}^{-})_{r_1})$, $(r_2, (K_{3,4}^{-})_{r_2})$ and $(r_3, (K_{3,4}^{-})_{r_3})$, see Section 5.2.4, as the presence of any of these graphs rooted at a vertex implies that it is closed.

The differential equation method will be used for the proof. Our aim in this section is to introduce the random variables and the differential equations implied by the expected changes in these random variables needed for Lemma 4.11. This should not be considered a rigorous proof as for simplicity we only estimate the expected change. In the following section we will show that these estimates are fairly accurate during the early stages of the process.

We start by considering the random variables. As in the previous results for H-free random graph processes we wish to track the number of open pairs. Recall that in the diamond-free process for an open pair $\{u, v\}$ the codegree of $\{u, v\}$ had a significant affect on the number of edges which would close $\{u, v\}$. In case of the $K_{3,4}^{-}$ one can observe a similar phenomenon, however instead of the codegree of $\{u, v\}$ it is the number of copies of $(r, (K_{3,3})_r)$ rooted at $\{u, v\}$ which have a significant affect on the number of edges which would close $\{u, v\}$. Note that unlike in the case of the diamond-free graph process where the codegree of any open pair is at most one, in this case no upper bound exists. Let $O_k(i) \subseteq O(i)$ be the set of open pairs $\{u, v\}$ where the number of copies of $(r, (K_{3,3})_r)$ rooted at $\{u, v\}$ is exactly k. Define $Q_k(i) = |O_k(i)|$. Obviously $Q(i) = \sum_{k\geq 0} Q_k(i)$.

Let (R, Γ) be a fixed rooted graph where R forms an independent set. Later we will see that we can restrict our analysis to graphs Γ which contain at most one copy of $K_{3,3}$ where this copy is induced. Let $\phi : R \to V(G_f(K_{3,4}^{-})_{n,i})$ be an injective mapping and $J \subseteq \Gamma$ be a fixed spanning subgraph of Γ. Similarly to the strictly two balanced case we are interested in the random variables $\Xi_{\phi,J,\Gamma}(i)$. Recall that these random variables were the maps $g : V(\Gamma) \to V(G_f(K_{3,4}^-)_{n,i})$ such that $g|_R = \phi$, $g(E(J)) \subseteq E(G_f(K_{3,4}^-)_{n,i})$ and for every $f \in E(\Gamma)\backslash E(J)$ we have that $g(f) \in O(i)$.

Similarly to the diamond-free case we would like to partition this based on the number of copies of $(r,(K_{3,3})_r)$ rooted at the open pairs of this copy. Let $\kappa : E(\Gamma)\backslash E(J) \to \mathbb{N}$, where $\mathbb{N}$ denotes the set of non-negative integers. This gives us sets $\Xi_{\phi,J,\Gamma,\kappa}(i)$ such that $g \in \Xi_{\phi,J,\Gamma,\kappa}(i)$ if $g \in \Xi_{\phi,J,\Gamma}(i)$ and for every $f \in E(\Gamma)\backslash E(J)$ we have that $g(f) \in O_{\kappa(f)+\gamma}(i)$ where γ is the number of copies of $(r,(K_{3,3})_r)$ rooted at f in J. Introducing γ allows us to differentiate between the edges of $G_f(K_{3,4}^-)$ based on whether they are spanned by $g(V(\Gamma))$ or not. Note that since Γ contains at most one copy of $K_{3,3}$ the value of γ is either 0 or 1.

We partition $\Xi_{\phi,J,\Gamma,\kappa}(i)$ even further. In order to motivate this consider the case when $(R,\Gamma) = (r,(K_{3,3})_r)$. Note that for $\phi(r)$ to be open it is vital that the codegree of every non-edge in the copy of $(K_{3,3})_r$, except for that of r has to be exactly 2. Let $C \subseteq (E(\overline{\Gamma}_R))$ be a subset of the non-edges of Γ, but not in R and for convenience let $\overline{C} = E(\overline{\Gamma}_R)\backslash C$. Then $g \in \Xi_{\phi,J,\Gamma,\kappa,C}(i)$ if $g \in \Xi_{\phi,J,\Gamma,\kappa}(i)$ and the codegree of $g(\overline{c})$ in $G_f(K_{3,4}^-)_{n,i}$ matches the codegree of $\overline{c}$ in J for every $\overline{c} \in \overline{C}$ and the codegree of $g(c)$ in $G_f(K_{3,4}^-)_{n,i}$ is larger than the codegree of c in J for every $c \in C$.

In summary the set $\Xi_{\phi,J,\Gamma,\kappa,C}(i)$ denotes the set of injective maps $g : V(\Gamma) \to V(G_f(K_{3,4}^-)_{n,i})$ such that the following conditions hold:

- $g|_R = \phi$
- $g(E(J)) \subseteq E(G_f(K_{3,4}^-)_{n,i})$
- for every $f \in E(\Gamma)\backslash E(J)$, $g(f) \in O_{\kappa(f)+\gamma}(i)$ where γ is the number of copies of $(r,(K_{3,3})_r)$ rooted at f in J
- for every $c \in C$ the codegree of $g(c)$ in $G_f(K_{3,4}^-)_{n,i}$ is larger then the codegree of c in J.
- for every $c \in \overline{C}$ the codegree of $g(c)$ in $G_f(K_{3,4}^-)_{n,i}$ matches the codegree of c in J.

Define $X_{\phi,J,\Gamma,\kappa,C}(i) = |\Xi_{\phi,J,\Gamma,\kappa,C}(i)|$.

In order to express small changes in the function κ let

$$\kappa_{e,k}(f) = \begin{cases} \kappa(f) & \text{if } f \neq e \\ k & \text{if } f = e \end{cases}$$

and this notation is also used when $\kappa(e)$ is undefined. Also let

$$\kappa_e^-(f) = \begin{cases} \kappa(f) & \text{if } f \neq e \\ \kappa(f) - 1 & \text{if } f = e \end{cases}$$

and let $\kappa_e(f)$ be the same as κ after removing e from its domain. In some cases when $|E(\Gamma)\backslash E(J)| = 1$ we replace κ with k when $\kappa(e) = k$ e.g. $X_{\phi,(K_3)_{\{r,e\}},(K_3)_r,\kappa_{e,k},\emptyset}(i) = X_{\phi,(K_3)_{\{r,e\}},(K_3)_r,k,\emptyset}(i)$. In many cases when $|E(\Gamma)\backslash E(J)| = 1$ we are not only interested in the individual random variables $X_{\phi,J,\Gamma,k,C}(i)$ but also in their sum over k, which gives us all the possibilities when this last edge is open, therefore define:

$$X_{\phi,J,\Gamma,C}(i) = \sum_{k\geq 0} X_{\phi,J,\Gamma,k,C}(i).$$

Similarly to the other H-free processes where H has 2-density 2, we define $t(i) = i/n^{3/2}$. In the following we will write t instead of $t(i)$. Next we set up our differential equations. As mentioned earlier in this section we will ignore small errors thus approximate

$$Q_k(i) \approx q_k(t)n^2$$
$$X_{\phi,J,\Gamma,\kappa,C}(i) \approx x_{R,J,\Gamma,\kappa,C}(t)n^{f_2(R,J)}$$

where $f_2(R,J) = v(J) - |R| - e(J)/2$. We will also assume that for $\{u,v\} \in O_k(i)$ the copies of $(r,(K_{3,3})_r)$ rooted at $\{u,v\}$ are non-edge disjoint.

One expects that in most cases we will have a similar result to the one shown for the H-free process when H is strictly 2-balanced namely

$$x_{R,J,\Gamma}(t) = \sum_{\kappa:E(\Gamma)\setminus E(J)\to\mathbb{N}} \sum_{C\subseteq\overline{\Gamma}_R} x_{R,J,\Gamma,\kappa,C}(t) = (2t)^{e(J)}(2q(t))^{e(\Gamma)-e(J)}.$$

Next we examine the changes of the random variables. Let $Q_k^+(i), Q_k^-(i)$ represent the positive and negative change in the random variable formally $Q_k^+(i) = |O_k(i+1)\setminus O_k(i)|$ and $Q_k^-(i) = |O_k(i)\setminus O_k(i+1)|$. For $\{u,v\} \in O_k(i)$ consider the number of edges which when inserted into the graph would remove $\{u,v\}$ from $O_k(i)$. Trivially inserting the edge $\{u,v\}$ would remove $\{u,v\}$ from $O_k(i)$. The remaining edges can be split into two parts based on whether $\{u,v\}$ is in $O_{k+1}(i+1)$ or $C(i+1)$. Denote the number of open pairs which when added would cause $\{u,v\}$ to end up in $O_{k+1}(i)$ with $A_{uv}(i)$ and let $C_{uv}(i)$ denote the number of open pairs which when added to the graph would cause $\{u,v\}$ to end up in $C(i+1)$. The random variables $A_{uv}(i)$ will be needed to determine $Q_k^+(i)$ and together with $C_{uv}(i)$ it helps to determine $Q_k^-(i)$.

We have that

$$\begin{aligned} A_{uv}(i) &\approx \frac{1}{\mathrm{aut}(K_{3,3})} \sum_{\substack{r,e\in E(K_{3,3})\\ r\neq e}} \sum_{\phi:r\to\{u,v\}} \sum_{k\geq 0} X_{\phi,(K_{3,3})_{\{r,e\}},(K_{3,3})_r,k,\emptyset}(i) \qquad (7.1)\\ &= \frac{1}{\mathrm{aut}(K_{3,3})} \sum_{\substack{r,e\in E(K_{3,3})\\ r\neq e}} \sum_{\phi:r\to\{u,v\}} X_{\phi,(K_{3,3})_{\{r,e\}},(K_{3,3})_r,\emptyset}(i)\\ &\approx \frac{2}{72} \sum_{\substack{r,e\in E(K_{3,3})\\ r\neq e}} x_{r,(K_{3,3})_{\{r,e\}},(K_{3,3})_r,\emptyset}(t)\sqrt{n} \end{aligned}$$

since $f_2(r,(K_{3,3})_{r,e}) = 6 - 2 - 7/2 = 1/2$ for any choice of r and e. Define

$$a(t) = \frac{2}{72} \sum_{\substack{r,e\in E(K_{3,3})\\ r\neq e}} x_{r,(K_{3,3})_{\{r,e\}},(K_{3,3})_r,\emptyset}(t).$$

Now $C_{uv}(i)$ can be split into further 3 parts. The first of these is when $\{u,v\}$ is closed by adding the last edge to a copy of $(r_1,(K_{3,4}^-)_{r_1})$ rooted at $\{u,v\}$. The second is when an edge is added which creates a copy of $(r_2,(K_{3,4}^-)_{r_2})$ or a copy of $(r_3,(K_{3,4}^-)_{r_3})$ rooted at $\{u,v\}$ and in particular a copy of $(r,(K_{3,3})_r)$ rooted at $\{u,v\}$. Finally when it creates a mutual neighbour for a pair of vertices assigned to a non-edge of an already existing copy of $(r,(K_{3,3})_r)$ rooted at $\{u,v\}$. Denote these in order $C_{uv,1}(i)$, $C_{uv,2}(i)$ and $C_{uv,3}(i)$.

We start with determining $C_{uv,1}(i)$. Note that this either requires a copy of $K_{3,3}$ to be already present in the graph or we create a copy of $K_{3,3}$ in the last step. However for a copy of $K_{3,3}$ to be created it is required that all of the vertex pairs assigned to the non-edges of $K_{3,3}$ have no mutual neighbours other than the ones required to create the copy of $K_{3,3}$. Let L be the set of non-edges contained in the copy of $K_{3,3}$ found in $(K_{3,4}^-)$ formally $L = \{\{2,6\},\{2,7\},\{6,7\},\{3,4\},\{3,5\},\{4,5\}\} \subseteq E(\overline{(K_{3,4}^-)})$ as in Figure 5.1. Note that for any C such that $L \cap C \neq \emptyset$ we have that $X_{\phi,(K_{3,4}^-)_{\{r_1,e\}},(K_{3,4}^-)_{r_1},k,C}(i) = 0$. Therefore

$$
\begin{aligned}
C_{uv,1}(i) &\approx \sum_{e\in E((K_{3,4}^-)_{r_1})} \sum_{L\subseteq \overline{C}\subseteq E(\overline{(K_{3,4}^-)})} \sum_{\phi:r_1\to\{u,v\}} \frac{X_{\phi,(K_{3,4}^-)_{\{r_1,e\}},(K_{3,4}^-)_{r_1},C}(i)}{\mathrm{aut}(r_1,H_{r_1})} \\
&\approx \frac{2}{6} \sum_{e\in E((K_{3,4}^-)_{r_1})} \sum_{L\subseteq \overline{C}\subseteq E(\overline{(K_{3,4}^-)})} x_{r_1,(K_{3,4}^-)_{\{r_1,e\}},(K_{3,4}^-)_{r_1},C}(t)\sqrt{n} \\
&= c_1(t)\sqrt{n}
\end{aligned}
$$

where $\mathrm{aut}(r,H_r)$ is the number of automorphisms of H_r which transfer the elements of r to r e.g. $\mathrm{aut}(r_1,(K_{3,4}^-)_{r_1}) = 6$. $C_{uv,2}(i)$ resembles $A_{uv}(i)$ except here $C \neq \emptyset$

$$
\begin{aligned}
C_{uv,2}(i) &\approx \frac{1}{\mathrm{aut}(K_{3,3})} \sum_{\substack{r,e\in E(K_{3,3})\\ r\neq e}} \sum_{\phi:r\to\{u,v\}} \sum_{\substack{C\subseteq E(\overline{K_{3,3}})\\ C\neq\emptyset}} X_{\phi,(K_{3,3})_{\{r,e\}},(K_{3,3})_r,C}(i) \\
&\approx \frac{1}{36} \sum_{\substack{r,e\in E(K_{3,3})\\ r\neq e}} \sum_{\substack{C\subseteq E(\overline{K_{3,3}})\\ C\neq\emptyset}} x_{r,(K_{3,3})_{\{r,e\}},(K_{3,3})_r,C}(t)\sqrt{n} \\
&= c_2(t)\sqrt{n}.
\end{aligned}
$$

Finally fix $r' \in E(K_{3,3})$ and $\phi: r' \to \{u,v\}$ then since we assume that the copies of $(r,(K_{3,3})_r)$ rooted at $\{u,v\}$ are non-edge disjoint we have that

$$
\begin{aligned}
C_{uv,3}(i) &\approx \frac{1}{4} \sum_{g\in \Xi_{\phi,(K_{3,3})_{r'},(K_{3,3})_{r'},\kappa_0,\emptyset}(i)} \sum_{\overline{e}\in\overline{K_{3,3}}} \sum_{\substack{r,e\in E(K_3)\\ r\neq e}} \sum_{\phi:r\to g(\overline{e})} \frac{X_{\phi,(K_3)_{\{r,e\}},(K_3)_r,\emptyset}(i)}{\mathrm{aut}(K_3)} \\
&\approx 6\frac{X_{\phi,(K_{3,3})_{r'},(K_{3,3})_{r'},\kappa_0,\emptyset}(i)}{4} \sum_{\substack{r,e\in E(K_3)\\ r\neq e}} \frac{x_{r,(K_3)_{\{r,e\}},(K_3)_r,\emptyset}(t)}{3}\sqrt{n} \\
&= c_{3,X_{\phi,(K_{3,3})_{r'},(K_{3,3})_{r'},\kappa_0,\emptyset}(i)/4}(t)\sqrt{n}
\end{aligned}
$$

where κ_0 is the function $\emptyset \to \mathbb{N}$. Note that if $\{u,v\} \in O_k(i)$ then according to our assumptions we have that $k = X_{\phi,(K_{3,3})_{r'},(K_{3,3})_{r'},(K_{3,3})_{r'},\kappa_0,\emptyset}(i)/4$ and thus $C_{uv,3}(i) \approx c_{3,k}(t)\sqrt{n}$.

In order to set up our differential equations we examine the one step changes for each random variable. Start with the random variable $Q_k(i)$. Let

$$Q_k(i+1) - Q_k(i) = Q_k^+(i) - Q_k^-(i)$$

where $Q_k^+(i), Q_k^-(i) \geq 0$. For $k = 0$ we have that $Q_k^+(i) = 0$ for all i and for $k > 0$ the following holds:

$$\mathbb{E}(Q_k^+(i)) \approx \frac{1}{Q(i)} \sum_{f \in O_{k-1}(i)} A_f(i) \approx \frac{q_{k-1}(t)}{q(t)} a(t)\sqrt{n}. \tag{7.2}$$

We also have that

$$\begin{aligned} E(Q_k^-(i)) &\approx \frac{1}{Q(i)} \sum_{f \in O_k(i)} 1 + A_f(i) + C_f(i) \\ &\approx \frac{1}{Q(i)} \sum_{f \in O_k(i)} A_f(i) + C_{f,1}(i) + C_{f,2}(i) + C_{f,3}(i) \\ &\approx \frac{q_k(t)}{q(t)}(a(t) + c_1(t) + c_2(t) + c_{3,k}(t))\sqrt{n}. \end{aligned} \tag{7.3}$$

Similarly for $X_{\phi,J,\Gamma,\kappa,C}(i)$

$$X_{\phi,J,\Gamma,\kappa,C}(i+1) - X_{\phi,J,\Gamma,\kappa,C}(i) = X^+_{\phi,J,\Gamma,\kappa,C}(i) - X^-_{\phi,J,\Gamma,\kappa,C}(i)$$

with $X^+_{\phi,J,\Gamma,\kappa,C}(i) \geq 0$ and $X^-_{\phi,J,\Gamma,\kappa,C}(i) \geq 0$ where $X^+_{\phi,J,\Gamma,\kappa,C}(i)$ and $X^-_{\phi,J,\Gamma,\kappa,C}(i)$ are defined analogously to $Q_k^+(i)$ and $Q_k^-(i)$.

We will distinguish between the cases when J contains a copy of $K_{3,3}$ or not. We start with the case when J contains no copy of $K_{3,3}$. Note that the value of $X_{\phi,J,\Gamma,\kappa,C}(i)$ increases if the last edge of a copy (r, J) with respect to ϕ is inserted, a last copy of $(r,(K_{3,3})_r)$ needed to satisfy κ is added and finally if the last non-edge not in C receives a mutual neighbour, assuming of course that the remaining parameters remain unchanged. Therefore

$$\begin{aligned} \mathbb{E}(X^+_{\phi,J,\Gamma,\kappa,C}(i)) \approx{}& \frac{1}{Q(i)} \sum_{e \in E(J)} \sum_{k \geq 0} X_{\phi,J_e,\Gamma,\kappa_{e,k},C}(i) \\ &+ \frac{1}{Q(i)} \sum_{f \in E(\Gamma)\setminus E(J)} \sum_{g \in \Xi_{\phi,J,\Gamma,\kappa_f^-,C}(i)} A_{g(f)}(i) \\ &+ \frac{1}{Q(i)} \sum_{c \in C} \sum_{g \in \Xi_{\phi,J,\Gamma,\kappa,C\setminus c}(i)} \sum_{\substack{r,e \in E(K_3) \\ r \neq e}} \sum_{\phi: r \to g(c)} \frac{X_{\phi,(K_3)_{\{r,e\}},(K_3)_r,\emptyset}(i)}{\mathrm{aut}(K_3)} \end{aligned}$$

$$\approx \frac{1}{q(t)} \sum_{e\in E(J)} \sum_{k\geq 0} x_{R,J_e,\Gamma,\kappa_{e,k},C}(t)\sqrt{n}n^{f_2(R,J)}$$
$$+ \frac{1}{q(t)} \sum_{f\in E(\Gamma)\setminus E(J)} x_{R,J,\Gamma,\kappa_f^-,C}(t)a(t)\sqrt{n}n^{f_2(R,J)}$$
$$+ \frac{1}{q(t)} \sum_{c\in C} x_{\phi,J,\Gamma,\kappa,C\setminus c}(t) \sum_{\substack{r,e\in E(K_3)\\ r\neq e}} \frac{x_{\phi,(K_3)_{\{r,e\}},(K_3)_r,\emptyset}(t)}{3}\sqrt{n}n^{f_2(R,J)}.$$

The only ways to decrease its value other than adding an edge are to close an open pair required by $E(\Gamma)\setminus E(J)$ or add a copy of $(r, (K_{3,3})_r)$ rooted at one of these open pairs or to create a mutual neighbour for a pair of vertices which are required not to have one. Thus

$$\mathbb{E}(X^-_{\phi,J,\Gamma,\kappa,C}(i)) \approx \frac{1}{Q(i)} \sum_{f\in E(\Gamma)\setminus E(J)} \sum_{g\in \Xi_{\phi,J,\Gamma,\kappa,C}(i)} \big(1 + C_{g(f)}(i) + A_{g(f)}(i)\big)$$
$$+ \frac{1}{Q(i)} \sum_{c\in \overline{C}} \sum_{g\in \Xi_{\phi,J,\Gamma,\kappa,C}(i)} \sum_{\substack{r,e\in E(K_3)\\ r\neq e}} \sum_{\phi:r\to g(c)} \frac{X_{\phi,(K_3)_{\{r,e\}},(K_3)_r,\emptyset}(i)}{\mathrm{aut}(K_3)}$$
$$\approx \frac{x_{R,J,\Gamma,\kappa,C}(t)}{q(t)} \left(\sum_{f\in E(\Gamma)\setminus E(J)} (a(t) + c_1(t) + c_2(t) + c_{3,\kappa(f)}(t)) \right.$$
$$\left. + \sum_{c\in \overline{C}} \sum_{\substack{r,e\in E(K_3)\\ r\neq e}} \frac{x_{\phi,(K_3)_{\{r,e\}},(K_3)_r,\emptyset}(t)}{3} \right) n^{f_2(R,J)+1/2}.$$

Something similar is true for the case when J contains a copy of $K_{3,3}$. The only difference is in the number of edges which can create a mutual neighbour for a pair of vertices which are required not to have one. Let $L(J)$ be the set of non-edges found in the induced copy of $K_{3,3}$ and note that $L \subseteq \overline{C}$. Also note that once the copy of $K_{3,3}$ has been completed the codegree of the vertex-pairs in $G_f(K^-_{3,4})_{n,i}$ corresponding to the non-edges in $L(J)$ are fixed i.e. they will not change any more during the process. Thus if $K_{3,3} \subseteq J$ then we have that:

$$\mathbb{E}(X^-_{\phi,J,\Gamma,\kappa,C}(i)) \approx \frac{1}{Q(i)} \sum_{f\in E(\Gamma)\setminus E(J)} \sum_{g\in \Xi_{\phi,J,\Gamma,\kappa,C}(i)} \big(1 + C_{g(f)}(i) + A_{g(f)}(i)\big)$$
$$+ \frac{1}{Q(i)} \sum_{c\in \overline{C}\setminus L(J)} \sum_{g\in \Xi_{\phi,J,\Gamma,\kappa,C}(i)} \sum_{\substack{r,e\in E(K_3)\\ r\neq e}} \sum_{\phi:r\to g(c)} \frac{X_{\phi,(K_3)_{\{r,e\}},(K_3)_r,\emptyset}(i)}{\mathrm{aut}(K_3)}$$

$$\approx \frac{x_{R,J,\Gamma,\kappa,C}(t)}{q(t)} \left(\sum_{f \in E(\Gamma)\setminus E(J)} (a(t) + c_1(t) + c_2(t) + c_{3,\kappa(f)}(t)) \right.$$
$$\left. + \sum_{c \in \overline{C} \setminus L(J)} \sum_{\substack{r,e \in E(K_3) \\ r \neq e}} \frac{x_{\phi,(K_3)_{\{r,e\}},(K_3)_r,\emptyset}(t)}{3} \right) n^{f_2(R,J)+1/2}.$$

Note that if there is no copy of $K_{3,3}$ in J then $L(J) = \emptyset$. These heuristics suggest the following system of differential equations:

$$q(t)q_k'(t) = q_{k-1}(t)a(t) - q_k(t)(a(t) + c_1(t) + c_2(t) + c_{3,k}(t))$$
$$q(t)x'_{R,J,\Gamma,\kappa,C}(t) = \sum_{e \in E(J)} \sum_{k \geq 0} x_{R,J_e,\Gamma,\kappa_{e,k},C}(t) + \sum_{f \in E(\Gamma)\setminus E(J)} x_{R,J,\Gamma,\kappa_f^-,C}(t)a(t)$$
$$+ \sum_{c \in C} x_{\phi,J,\Gamma,\kappa,C \setminus c}(t) \sum_{\substack{r,e \in E(K_3) \\ r \neq e}} \frac{x_{\phi,(K_3)_{\{r,e\}},(K_3)_r,\emptyset}(t)}{3}$$
$$- x_{R,J,\Gamma,\kappa,C}(t) \sum_{f \in E(\Gamma)\setminus E(J)} (a(t) + c_1(t) + c_2(t) + c_{3,\kappa(f)}(t))$$
$$- x_{\phi,J,\Gamma,\kappa,C}(t) \sum_{c \in \overline{C} \setminus L(J)} \sum_{\substack{r,e \in E(K_3) \\ r \neq e}} \frac{x_{\phi,(K_3)_{\{r,e\}},(K_3)_r,\emptyset}(t)}{3}.$$

Since the process starts from the empty graph we have that $q_0(0) = 1/2$ and $q_k(0) = 0$ for $k > 0$. Also $x_{R,J,\Gamma,\kappa,C}(0) = 1$ if $E(J) = 0$, $C = \emptyset$ and κ is the constant 0 function. Otherwise $x_{R,J,\Gamma,\kappa,C}(0) = 0$.

In order to find a solution for this system of differential equations we will assume that similarly to all of the previous cases the functions $x_{R,J,\Gamma,\kappa,C}(t)$ can be constructed by multiplying different terms. In the first step assume that we have found functions $q_k(t)$, which satisfy the differential equations. We have seen that the number of triangles in the diamond-free process differs from the number of triangles in $G_{n,i}$. Similarly the number of copies of $K_{3,3}$ found in the $K_{3,4}^-$ process will diverge from the number of copies of $K_{3,3}$ found in $G_{n,i}$. We will first consider cases where J does not contain a copy of $K_{3,3}$.

In the strictly 2-balanced case we have seen that the solution for the differential equations was $x_{R,J,\Gamma}(t) = (2t)^{e(J)}(2q(t))^{e(\Gamma)-e(J)}$. In the diamond-free process we have seen that when partitioning the open pairs one just has to replace the $q(t)$ terms with the appropriate $q_k(t)$ terms. This indicates the following solution

$$x_{\phi,J,\Gamma,\kappa}(t) = \sum_{C \subseteq \overline{\Gamma}_R} x_{\phi,J,\Gamma,\kappa,C}(t) = (2t)^{e(J)} \prod_{f \in E(\Gamma)\setminus E(J)} (2q_{\kappa(f)}).$$

Finally we will rely on the fact that this process is similar to $G_{n,i}$ and thus also to $G_{n,p}$ when $i \approx \binom{n}{2}p$ and since $i = tn^{3/2}$ thus $p \approx \frac{2t}{\sqrt{n}}$. Note that the codegree of a pair of vertices in $G_{n,p}$ follows a Poisson distribution with parameter $\lambda = np^2 = (2t)^2$.

Based on these intuitions this suggests the following solutions

$$x_{R,J,\Gamma,\kappa,C}(t) = (2t)^{e(J)} \prod_{f\in E(\Gamma)\setminus E(J)} (2q_{\kappa(f)}) \left(\exp(-4t^2)\right)^{|\overline{C}|} \left(1-\exp(-4t^2)\right)^{|C|}$$

and one can easily verify that, based on our assumptions, these solutions satisfy the differential equations.

Now consider the function $x_{\phi_0,K_{3,3},K_{3,3},\kappa_0,C}(t)$ and note that this is zero unless $C=\emptyset$ in which case we have that:

$$\begin{aligned} q(t)x'_{\phi_0,K_{3,3},K_{3,3},\kappa_0,\emptyset}(t) &= \sum_{e\in E(K_{3,3})} \sum_{k\geq 0} x_{\phi_0,(K_{3,3})_e,K_{3,3},\kappa_{e,j},\emptyset}(t) \\ &= \sum_{k\geq 0} 9(2t)^8 2q_k(t) \left(\exp(-(2t)^2)\right)^6 \\ &= 18(2t)^8 q(t) \left(\exp(-(2t)^2)\right)^6 \end{aligned}$$

and thus

$$\begin{aligned} x_{\phi_0,K_{3,3},K_{3,3},\kappa_0,\emptyset}(t) &= \int_0^t 18(2\tau)^8 \exp(-6(2\tau)^2)d\tau \\ &= \frac{35\int_0^t \mathrm{e}^{-24\tau^2}d\tau}{384} - \mathrm{e}^{-24t^2}\left(96t^7 + \frac{7}{2}t^5 + \frac{35}{24}t^3 + \frac{35}{384}t\right). \end{aligned}$$

Define the correction factor

$$\xi(t) = \frac{\int_0^t 18(2\tau)^8 \exp(-6(2\tau)^2)d\tau}{(2t)^9 \exp(-6(2t)^2)}$$

where the numerator is the correct form for $x_{\phi_0,K_{3,3},K_{3,3},\kappa_0,\emptyset}(t)$, while the denominator is the form which we would get based on our previous calculations i.e. if we ignored the fact that it is a copy of $K_{3,3}$. Let $\gamma(J)$ be the indicator function that J contains a copy of $K_{3,3}$. Since Γ contains at most one copy of $K_{3,3}$, and this copy is induced we have that:

$$x_{R,J,\Gamma,\kappa,C}(t) = (2t)^{e(J)} \prod_{f\in E(\Gamma)\setminus E(J)} (2q_{\kappa(f)}) \left(\exp(-4t^2)\right)^{|\overline{C}|} \left(1-\exp(-4t^2)\right)^{|C|} \xi^{\gamma(J)}.$$

Note that this also satisfies the system of differential equations assuming that $q_k(t)$ satisfies them.

Now we will determine $q_k(t)$. Recall that:

$$a(t) = \frac{1}{36} \sum_{\substack{r,e\in E(K_{3,3})\\ r\neq e}} x_{r,(K_{3,3})_{\{r,e\}},(K_{3,3})_r,\emptyset}(t) = 2(2t)^7 2q(t)\exp(-(2t)^2)^6$$

$$c_1(t) = \frac{1}{3} \sum_{e\in E((K^-_{3,4})_{r_1})} \sum_{L\subseteq\overline{C}\subseteq E(\overline{(K^-_{3,4})})} x_{r_1,(K^-_{3,4})_{\{r_1,e\}},(K^-_{3,4})_{r_1},C}(t) =$$

$$= \frac{1}{3}(2t)^9 2q(t)\exp(-(2t)^2)^6(9+\xi(t))$$

$$c_2(t) = \frac{1}{36} \sum_{\substack{r,e \in E(K_{3,3}) \\ r \neq e}} \sum_{\substack{C \subseteq E(\overline{K_{3,3}}) \\ C \neq \emptyset}} x_{r,(K_{3,3})_{\{r,e\}},(K_{3,3})_r,C}(t)$$
$$= 2(2t)^7 2q(t)(1 - \exp(-(2t)^2)^6)$$
$$c_{3,k}(t) = 6k \sum_{\substack{r,e \in E(K_3) \\ r \neq e}} \frac{x_{r,(K_3)_{\{r,e\}},(K_3)_r}(t)}{3} = 6k2(2t)(2q(t)).$$

In order to solve the differential equations we again use the connection to $G_{n,p}$ when $p = 2t/\sqrt{n}$. In this graph the number of copies of $(r, (K_{3,3})_r)$ rooted at $\{u, v\}$ is Poisson distributed with parameter $(2t)^8/4$ and thus the probability that there are exactly k of them is

$$\frac{(2t)^{8k}}{4^k k!} \exp\left(-\frac{(2t)^8}{4}\right).$$

Note that $(2t)^8/4 = \int (c_2(\tau) + a(\tau))/q(\tau)d\tau$ Also the codegree of a vertex pair is Poisson distributed and thus the probability that a vertex pair has codegree zero is $\exp(-(2t)^2)$. Now if we consider multiple vertex pairs and assume that their codegrees are independent thus the probability that $6k$ vertex pairs have codegree zero is

$$\exp\left(-(2t)^2\right)^{6k}.$$

Note that $24kt^2 = \int c_{3,k}(\tau)/q(\tau)d\tau$. This leads us to guess that the the last term needed is

$$\exp\left(-\int_0^t \frac{c_1(\tau}{q(\tau)}d\tau\right).$$

Define

$$F(t) = \int_0^t \frac{c_1(\tau)}{q(\tau)}d\tau = \int_0^t \frac{1}{3}(2\tau)^9 2\exp(-(2\tau)^2)^6(9 + \xi(\tau))d\tau$$

and note that

$$F(t) = \int_0^t 6(2\tau)^9 \exp(-(2\tau)^2)^6 d\tau + \int_0^t \int_0^\tau 18(2\rho)^8 \exp(-6(2\rho)^2)d\rho d\tau$$
$$= \frac{1}{216} - \exp(-(2t)^2)^6 \left(64t^8 + \frac{32}{3}t^6 + \frac{4}{3}t^4 + \frac{1}{9}t^2 + \frac{1}{216}\right)$$
$$+ \frac{35}{384}t \int_0^t \exp(-(2\tau)^2)^6 d\tau + \left(\frac{114}{18432}\right)\left(\exp(-24t^2) - 1\right)$$
$$+ \exp(-(2t)^2)^6 \left(2t^6 + \frac{93}{32}t^4 + \frac{33}{64}t^2\right).$$

We will only need that when $t \leq 1$ then $0 \leq F(t) \leq 1$ and that for $t > 1$ there exists a small constant c such that $ct \leq F(t) \leq t$.

Multiplying these terms together, and including a $1/2$ term to satisfy the initial conditions, gives us that:

$$q_k(t) = \frac{1}{2}\frac{(2t)^{8k}}{4^k k!} \exp\left(-\frac{(2t)^8}{4}\right) \left(\exp(-(2t)^2)\right)^{6k} \exp\left(-F(t)\right).$$

Note that summing over $k \geq 0$ gives us

$$q(t) = \frac{1}{2}\exp\left(\frac{(2t)^8}{4}\exp(-6(2t)^2)\right)\exp\left(-\frac{(2t)^8}{4}\right)\exp\left(-F(t)\right).$$

It is easy to verify that $q_k(t)$ is in fact the solution. Derive $q_k(t)$ according to the product rule. Deriving the first term gives us $a(t)q_{k-1}(t)/q(t)$, the second term gives us $-(a(t)+c_2(t))q_k(t)/q(t)$, the third term gives $-c_{3,k}(t)q_k(t)/q(t)$ and the final term gives $-c_1(t)q_k(t)/q(t)$. Also note that the only terms in $q(t)$ which tend to 0 are $\exp\left(-(2t)^8/4\right)$ and $\exp(-F(t))$ and that the $\exp\left(-(2t)^8/4\right)$ term tends to zero much faster. Note that the process contains $q(t)n^2$ open pairs. Assuming that the differential equations hold until the end of the process we have that the process terminates once $q(t) = n^{-2}$. Note that this happens when $t = \Theta(\log^{1/8} n)$ suggesting that the process terminates with $\Theta(n^{3/2}\log^{1/8} n)$ edges. This differs from the strictly 2-balanced case where it is conjectured [6] that the process terminates after $\Theta(n^{2-1/d_2(H)}\log^{1/(e(H)-1)} n)$ steps, as the exponents of the $\log$ factors do not match. We will only show that a.a.s. the process runs for $\Omega(n^{3/2}\sqrt{\log\log n})$ steps.

Theorem 7.1

The $K_{3,4}^-$-free random graph process. a.a.s. contains $\Omega(n^{3/2}\sqrt{\log\log n})$ edges.

7.2 Differential equation proof

7.2.1 The Setup

In this section we verify the conditions of Lemma 4.11. In the setup of the Lemma 4.11 set $m = \mu n^{3/2}\sqrt{\log\log n}$, where $\mu = 1/200$. No attempt was made to optimise the value of μ. Also let $s = n^{3/2}$. We make several additional restrictions on the random variables described in the previous section, first we are only interested in the random variables $X_{\phi,J,\Gamma,\kappa,C}(i)$ when Γ is one of $(K_3)_r$, $(K_{3,3})_r$ or $(K_{3,4}^-)_{r_1}$ and J is a proper spanning subgraph of Γ i.e. $e(\Gamma) - e(J) \geq 1$. The differential equations imply that since there are only a few open pairs f with many copies of $(r, (K_{3,3})_r)$ rooted at f, they will have no significant impact on the outcome of the process. Set $k_{max} = \log n/22\log\log n$. However we still need an estimate for $\sum_{k>k_{max}} Q_k(i)$ and similarly for $\sum_{\kappa:\max\{\kappa\}>k_{max}} X_{\phi,J,\Gamma,\kappa,C}(i)$. It would be hard to track these sums as even determining their expected change is difficult. Instead of $\sum_{k>k_{max}} Q_k(i)$ we track the number of open pairs {u,v} which have been in $O_k(j)$ for some $k > k_{max}$ and $j \leq i$, formally

$$Y_o(i) = \left|\bigcup_{j\leq i}\bigcup_{k>k_{max}} O_k(j)\right|.$$

Similarly for every random variable $X_{\phi,J,\Gamma,\kappa,C}(i)$ and every $f \in E(\Gamma)\backslash E(J)$ we would like to track $\sum_{\kappa:\max\{\kappa\}>k_{max}} X_{\phi,J,\Gamma,\kappa,C}(i)$ therefore define

$$Z_{\phi,J,\Gamma,\kappa,C,S}(i) = \left| \bigcup_{j\le i} \bigcup_{\substack{\kappa':E(\Gamma)\backslash E(J)\to\mathbb{N} \\ \min_{f\in S}\kappa'(f)>k_{max} \\ \kappa'|_{E(\Gamma)\backslash(E(J)\cup S)}=\kappa}} \Xi_{\phi,J,\Gamma,\kappa,C}(j) \right|.$$

Let $\mathcal{V} = \{q, x, y, z\}$. The elements of I_j for $j \in \mathcal{V}$ are as follows:

- $\mathcal{I}_q$ is the set of positive integers from 0 to k_{max}
- $\mathcal{I}_x$ is the set of quintuples $(\phi, J, \Gamma, \kappa, C)$ such that (r, Γ) is one of $(r,(K_3)_r)$, $(r,(K_{3,3})_r)$, or $(r_r,(K_{3,4}^-)_{r_1})$ such that $\phi : r \to V(G_f(K_{3,4}^-)_{n,i})$ is an injective function, J is a proper spanning subgraph of Γ, $C \subseteq E(\overline{\Gamma}_R)$ and $\kappa : E(\Gamma)\backslash E(J) \to [k_{max}]$
- $\mathcal{I}_y = \{o\}$
- $\mathcal{I}_z$ is the set of sextuplets $(\phi, J, \Gamma, \kappa, C, S)$ such that (r, Γ) is one of $(r,(K_3)_r)$, $(r,(K_{3,3})_r)$, or $(r_r,(K_{3,4}^-)_{r_1})$ such that $\phi : r \to V(G_f(K_{3,4}^-)_{n,i})$ is an injective function, J is a proper spanning subgraph of Γ, $C \subseteq E(\overline{\Gamma}_R)$, $S \subseteq E(\Gamma)\backslash E(J)$ and $\kappa : E(\Gamma)\backslash(E(J) \cup S) \to [k_{max}]$.

We are interested in the set of random variables $Q_k(i)$ for $k \in \mathcal{I}_o$, $X_{\phi,J,\Gamma,\kappa,C}(i)$ when $(\phi, J, \Gamma, \kappa, C) \in \mathcal{I}_x$, Y_o and $Z_{\phi,J,\Gamma,\kappa,C,S}$ when $(\phi, J, \Gamma, \kappa, C, S) \in \mathcal{I}_z$. Let $\mathcal{I} = \bigcup_{j\in\mathcal{V}} \mathcal{I}_j$. In order to comply with the notation of Lemma 4.11 for $\sigma \in \mathcal{I}_q$ we have that $Q_\sigma(i)$ plays the role of $X_\sigma(i)$ and similarly for $\sigma \in \mathcal{I}_y$ and $\sigma \in \mathcal{I}_z$ we have that $Y_\sigma(i)$ and $Z_\sigma(i)$ play the role of $X_\sigma(i)$. We will use this notation only when referring to all random variables i.e. to X_σ for $\sigma \in \mathcal{I}$.

We do not track every random variable through all of the m steps. If Γ is either $(K_{3,3})_r$ or $(K_{3,4}^-)_{r_1}$ we stop tracking the random variables $X_{\phi,J,\Gamma,\kappa,C}(i)$ and $Z_{\phi,J,\Gamma,\kappa,C,S}(i)$ for every $S \in E(\Gamma)\backslash E(J)$ if $\phi(r)$ or respectively $\phi(r_1)$ becomes closed or an edge. In case $\Gamma = (K_3)_r$ we stop tracking the variables when $\phi(r)$ becomes a non-edge in a copy of $K_{3,3}$ in $G_f(K_{3,4}^-)_{n,i}$.

Define

$$\begin{aligned}
W &= 420 \\
\varepsilon &= 1/66 \\
u_\sigma &= \log^2 n \\
\tau_\sigma &= \log n \\
S_k &= S_o = n^2 \\
S_{r,J,\Gamma,\kappa,C} &= S_{r,J,\Gamma,\kappa,C,S} = n^{f_2(r,J)} \\
\beta_k &= \frac{1}{k!}
\end{aligned}$$

$$\begin{aligned}
\beta_{r,J,\Gamma,\kappa,C} &= \prod_{f\in E(\Gamma)\setminus E(J)} \beta_{\kappa(f)} \\
\lambda_\sigma &= \beta_\sigma n^\varepsilon \\
\beta_o &= 1 \\
\beta_{r,J,\Gamma,\kappa,C,S} &= \prod_{f\in E(\Gamma)\setminus(E(J)\cup S)} \beta_{\kappa(f)} \\
s_\sigma &= n^{(1/22)-\varepsilon} \\
f_k(t) &= \beta_k 2q(t)\exp(24Wk_{max}(t^2+t)) \\
f_{r,J,\Gamma,\kappa,C}(t) &= (2q(t))^{e(\Gamma)-e(J)}\beta_{r,J,\Gamma,\kappa,C}\exp(24Wk_{max}(t^2+t)) \\
f_o &= (2q(t))\exp(24Wk_{max}(t^2+t)) \\
f_{r,J,\Gamma,\kappa,C,S}(t) &= (2q(t))^{e(\Gamma)-e(J)}\beta_{r,J,\Gamma,\kappa,C,S}\exp(24Wk_{max}(t^2+t)).
\end{aligned}$$

Also define $f(t) = \sum_{k\geq 0} f_k(t) = \mathrm{e}2q(t)\exp(W(24k_{max}(t^2+t))$ and note that $f_k(t) = f(t)/(\mathrm{e}k!)$. If we can show that the conditions of Lemma 4.11 hold then we will have that for every $\sigma \in \mathcal{I}$

$$X_\sigma = \left(x_\sigma \pm \frac{f_\sigma}{s_\sigma}\right) S_\sigma$$

where $q_k(t)$ and $x_{r,J,\Gamma,\kappa,C}(t)$ are as in the previous chapter and $y_o(t) = 0$ and $z_{r,J,\Gamma,\kappa,C,S}(t) = 0$.

Next we define the event $\mathcal{H}_i$. The event $\mathcal{H}_i$ is the union of several events, most of which concern the number of overlapping rooted graphs.

Definition 7.1

For a set of graphs F_i where $i = 1..k$ define the set $F_1 \otimes F_2 \otimes ... \otimes F_k$ as the set of $k+1$ tuples $(g_1, g_2, ..., g_k, F)$ where $g_i : V(F_i) \to V(F)$ are injective functions for $i = 1..k$ such that the following hold

- $V(F) = \bigcup_{i=1}^k g_i(V(F_i))$
- $E(F) = \bigcup_{i=1}^k g_i(E(F_i))$
- *for every $i, j \in [k]$ with $i \neq j$ we have that $g_i(E(F_i)) \neq g_j(E(F_j))$.*

The first event is $\mathcal{B}_i$. Consider the rooted graphs (r', F) which are formed of two copies of $(r, (K_{3,3})_r)$, with shared roots and an overlapping non-edge. Formally $(r', F) = (g_1(r), F)$ where $(g_1, g_2, F) \in ((K_{3,3})_r) \otimes ((K_{3,3})_r)$ such that $g_1(r) = g_2(r)$, and there exists $e_1, e_2 \in E(\overline{K_{3,3}})$ such that $g_1(e_1) = g_2(e_2)$. Let $\mathcal{B}_i$ be the event that for every $j \leq i$ we have that for any such rooted graph the total number of copies of F in $G_f(K_{3,4}^-)_{n,j}$ such that (r', F) is rooted at an open pair is at most $f(t)n^\varepsilon$.

Let $\mathcal{C}_i$ be the event that for every $j \leq i$ we have that

- for any pair of vertices $u, v \in V(G_f(K_{3,4}^-)_{n,j})$ with $u \neq v$ and for any rooted graph $(r, F) \in \{(r_1, (K_{3,4}^-)_{r_1}), (r, (K_3)_r), (r, (K_{3,3})_r)\}$ we have that the number of edges which when inserted into $G_f(K_{3,4}^-)_{n,j}$ would create multiple copies of (r, F) rooted at u, v is at most $f(t)\sqrt{n}/n^{1/11}$

- for any four vertices $u_1, u_2, v_1, v_2 \in V(G_f(K_{3,4}^-)_{n,i})$, with $u_1 \neq v_1$, $u_2 \neq v_2$ and for $(r_1, F_1), (r_2, F_2) \in \{(r_1, (K_{3,4}^-)_{r_1}), (r, (K_3)_r), (r, (K_{3,3})_r)\}$ the number of edges which when inserted into $G_f(K_{3,4}^-)_{n,j}$ would create both a copy of (r_1, F_1) at $\{u_1, v_1\}$ and (r_2, F_2) at u_2, v_2 is at most $f(t)\sqrt{n}/n^{1/11}$.

The event $\mathcal{D}_i$ holds if for every $j \leq i$ the following hold. Select a rooted graph $(r, \Gamma) \in \{(r, (K_3)_r), (r, (K_{3,3})_r), (r_1, (K_{3,4}^-)_{r_1})\}$ and let (r, J) be a proper spanning subgraph of (r, Γ). Also let $(r_2, F_2) \in \{(r, (K_3)_r), (r, (K_{3,3})_r), (r_1, (K_{3,4}^-)_{r_1})\}$. Then for any $(g_1, g_2, F) \in (J \otimes F_2)$ such that $g_2(r_2) \subseteq g_1(V(J))$ and $g_1(E(\Gamma)\backslash E(J)) \cap g_2(E(F_2)) = \emptyset$ we have that for any pair of distinct vertices $u, v \in V(G_f(K_{3,4}^-)_{n,j})$ the number of edges which when inserted into the graph would create a copy of $(g_1(r), F)$ rooted at $\{u, v\}$ in $G_f(K_{3,4}^-)_{n,j}$ is $O(n^{f_2(r,J)} n^{\varepsilon})$.

Finally let $\mathcal{E}_i$ be the event that for every $j \leq i$ the Boundedness hypothesis holds. Define

$$\mathcal{H}_i = \mathcal{B}_i \cap \mathcal{C}_i \cap \mathcal{D}_i \cap \mathcal{E}_i.$$

Note that in $\mathcal{C}_i$ and $\mathcal{D}_i$ we are considering non-edges not open pairs. This is clearly an upper bound on the number of open pairs which satisfy the conditions.

Now we can state the Theorem

Definition 7.2

Let $\mathcal{G}_m$ be the event that for every $\sigma \in \mathcal{I}$ and every $i \leq m$ we have that

$$X_\sigma(i) = \left(x_\sigma(t) \pm \frac{f_\sigma(t)}{s_\sigma} \right) S_\sigma.$$

Theorem 7.2

We have that $\mathcal{G}_m$ holds a.a.s. for $m = \mu n^{3/2}\sqrt{\log \log n}$.

PROOF We will use Lemma 4.11 for the proof. The trend hypothesis is verified in Corollary 7.14. Note that the boundedness hypothesis follows from the condition on the high probability events. The initial conditions follow from Lemma 7.15. We show that the number of variables is bounded in Lemma 7.16. The fact that $\mathcal{H}_i$ is a high probability event follows from Corollary 7.30. Finally the remaining technical assumptions are verified in Lemma 7.31. ∎

PROOF (THEOREM 7.1) Theorem 7.2 implies that there are still $n^{2-o(1)}$ open pairs after m steps. ∎

7.2.2 Basic Observations

Before continuing we first consider how two copies of $(r, (K_{3,3})_r)$ rooted at the same or different open pairs in $G_f(K_{3,4}^-)_{n,i}$ can overlap.

We start by considering two copies rooted at the same open pair.

Lemma 7.3
Let $\{u, v\} \in O(i)$ then any edge adjacent to u or v can be found in at most one copy of $(r, (K_{3,3})_r)$ rooted at $\{u, v\}$.

PROOF Without loss of generality assume that it is adjacent to u and let w denote the other end of the edge. Note that $v \neq w$ and that v, w must have exactly 2 mutual neighbours as two are needed for $\{u, w\}$ to be in a copy of $(r, (K_{3,3})_r)$ rooted at $\{u, v\}$, but more than two would imply that the presence of a copy of $(r, (K_{3,3})_r)$ rooted at $\{u, v\}$ including the $\{u, w\}$ edge would also imply the presence of $(r_3, (K^-_{3,4})_{r_3})$ rooted at $\{u, v\}$. Denote the vertices connected to both u and w with w_1 and w_2. A similar argument gives us that w_1 and w_2 can have at most one mutual neighbour other then v and w. This restricts us to one possibility of $(r, (K_{3,3})_r)$ containing $\{u, w\}$ as an edge. ∎

Lemma 7.4
For $\{u, v\} \in O(i)$ any edge is found in at most two copies of $(r, (K_{3,3})_r)$ rooted at $\{u, v\}$.

PROOF In Lemma 7.3 we have seen that if the edge is adjacent to u or v then we can have at most one copy. Now assume that $\{w_1, w_2\}$ is not adjacent to either u or v. In order for $\{w_1, w_2\}$ to be an edge in a copy of $(r, (K_{3,3})_r)$ rooted at $\{u, v\}$ the copy must also contain either $\{u, w_1\}$ or $\{u, w_2\}$ as an edge. Since there are at most one of each of these thus at most two copies containing $\{w_1, w_2\}$. ∎

Lemma 7.5
Let $f \in O(i)$. Assume that there are two copies of $(r, (K_{3,3})_r)$ rooted at f which share a non-edge in addition to f. Then we have that the two copies create the following construction rooted at f:

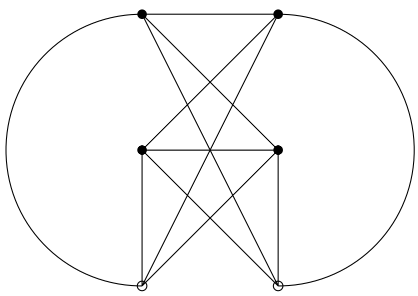


Figure 7.1: Overlap at an open pair

PROOF Let two copies of $(r, (K_{3,3})_r)$ rooted at f share $\{w_1, w_2\}$ as a non-edge and let $f = \{u, v\}$. There are two cases based on whether $\{w_1, w_2\} \cap \{u, v\} = \emptyset$ or not.

In case $\{w_1, w_2\} \cap \{u, v\} \neq \emptyset$ then without loss of generality we may assume that $v = w_1$ and thus any copies which share $\{w_1, w_2\}$ as a non-edge, must also contain $\{u, w_2\}$ as an edge. However no two copies may share an edge adjacent to $\{u, v\}$ according to Lemma 7.3. Otherwise since f is open and $\{w_1, w_2\}$ is a non-edge in a copy of $(r, (K_{3,3})_r)$ rooted at f thus $\{w_1, w_2\}$ has exactly 2 mutual neighbours other than u and v. Denote these vertices by w_3 and w_4. Thus both copies must be on the same set of vertices and the edges spanned by w_1, w_2, w_3, w_4 can be found in both copies. Since for $i = 1..4$ the edge $\{u, w_i\}$ and $\{v, w_i\}$ can be found in at most one copy of $(r, (K_{3,3})_r)$ rooted at f hence none of the remaining edges may overlap. The result follows. ∎

7.2.3 Approximations

In the following we introduce several approximations which we will rely on heavily while verifying the trend hypothesis for Lemma 4.11.

The first of these is that when $0 \leq a \leq 1/2$ then

$$\frac{1}{1 \pm a} = 1 \pm 2a.$$

Also according to Stirling's formula

$$\begin{aligned} k_{max}! &= (1 + o(1))\sqrt{2\pi k_{max}} \left(\frac{k_{max}}{e}\right)^{k_{max}} \\ &= n^{o(1)} \exp\left(\frac{\log n}{22 \log\log n}(1 + o(1)) \log\log n\right) = n^{1/22+o(1)} \end{aligned}$$

and note that $1/k_{max}! = o(1/s_\sigma)$ and also $(k_{max}!)^{10} = o(n^{1/2-\varepsilon})$. We also have that if $\mathcal{G}_i$ holds then for every $\sigma \in \mathcal{I}_q \cup \mathcal{I}_x$ we have that $X_\sigma = (1 + o(1))x_\sigma S_\sigma$, due to the fact that in these cases $x_\sigma = \omega(\beta_\sigma n^{-\varepsilon})$ and $f_\sigma \leq \beta_\sigma n^{\varepsilon}$.

Lemma 7.6

We have that

$$q_k(t) \leq \left(\frac{(2t)^8}{4} \exp(-6(2t)^2)\right)^k \frac{q(t)}{k!} \leq \frac{q(t)}{k!}$$

PROOF The first inequality follows from the fact that $\exp((2t)^8 \exp(-(2t)^2)^6/4) \geq 1$ and the second follows from the fact that $(2t)^8 \exp(-(2t)^2)^6/4 < 1/4$. ∎

Since $e(J) \leq 9$ we have that

$$x_{r,J,\Gamma,\kappa,C}(t) \leq (2t)^{e(J)} \prod_{f \in E(\Gamma)\setminus E(J)} 2q_k(t) \leq ((2t)^9 + 1)(2q(t))^{e(\Gamma)-e(J)}\beta_{r,J,\Gamma,\kappa,C}.$$

Also note that

$$(2t)^7 \exp(-(2t)^2)^6 < \frac{1}{4} \tag{7.4}$$

$$(2t)^8 \exp(-(2t)^2)^6 < \frac{1}{4} \tag{7.5}$$

$$((2t)^9 + 1)(2t)^7 \exp(-(2t)^2)^6 < \frac{1}{4} \tag{7.6}$$

$$((2t)^9 + 1)(2t)^8 \exp(-(2t)^2)^6 < \frac{1}{4}. \tag{7.7}$$

Lemma 7.7

Assuming that $\mathcal{G}_i$ and $\mathcal{H}_i$ hold then

$$Q(i) = \left(q(t) \pm \frac{3f(t)}{s_\sigma}\right) n^2.$$

PROOF Note that Taylor's theorem using the Lagrange form of the remainder implies that

$$e^x = \sum_{k=0}^{k_{max}} \frac{x^k}{k!} \pm e^x \frac{x^{k+1}}{(k+1)!}.$$

Note that when $x \leq 1/16$ then $e^x \leq 2$. Set $x = (2t)^8 \exp(-(2t)^2)^6/4$ and (7.5) implies that $(2t)^8 \exp(-(2t)^2)^6/4 \leq 1/16$. Therefore

$$\exp\left(\frac{(2t)^8}{4} \exp(-(2t)^2)^6\right) = \sum_{k=0}^{k_{max}} \frac{(2t)^{8k}}{4^k k!} \exp(-(2t)^2)^{6k}$$
$$\pm 2 \frac{(2t)^{8(k+1)}}{4^{k+1}(k+1)!} \exp(-(2t)^2)^{6(k+1)}.$$

This implies that

$$q(t) = \sum_{k=0}^{\infty} q_k(t) = \exp\left(-\frac{(2t)^8}{4}\right) \exp(-F(t) \sum_{k=0}^{\infty} \frac{(2t)^{8k}}{4^k k!} \exp(-(2t)^2)^{6k}$$
$$= \exp\left(-\frac{(2t)^8}{4}\right) \exp(-F(t) \sum_{k=0}^{k_{max}} \frac{(2t)^{8k}}{4^k k!} \exp(-(2t)^2)^{6k}$$
$$\pm 2 \exp\left(-\frac{(2t)^8}{4}\right) \exp(-F(t) \frac{(2t)^{k+1}}{4^{k+1}(k+1)!} \exp(-(2t)^2)^{6(k+1)}$$
$$= \sum_{k=1}^{k_{max}} q_k(t) \pm 2q_{k+1}(t) = \sum_{k=1}^{k_{max}} q_k(t) \pm 2 \max_{t \geq 0} q_{k+1}(t)$$

as we are only interested in cases when $t \geq 0$. Thus we have that

$$Q(i) = \sum_{k=0}^{\infty} Q_k(i) = \sum_{k=0}^{k_{max}} Q_k(i) \pm Y_o(i) = \sum_{k=0}^{k_{max}} \left(q_k(t) \pm \frac{f_k(t)}{s_\sigma}\right) n^2 \pm \frac{f(t)}{s_\sigma} n^2$$
$$= \left(q(t) \pm \frac{f(t)}{s_\sigma}\right) n^2 \pm 2 \max_{t \geq 0}(q_{k_{max}+1}(t)) n^2 \pm \frac{f(t)}{s_\sigma} n^2.$$

Now all we have to show is that $2\max_{t\geq 0}(q_{k_{max}+1}(t)) \leq f(t)/s_\sigma$ in fact we will show that $q_{k_{max}+1}(t) = o(1/s_\sigma)$. We have for every k that

$$2q_k(t) \leq 2\frac{q(t)}{k!} = O\left(\frac{1}{k!}\right).$$

Subbing in $k = k_{max} + 1$ gives us that

$$2q_{k_{max}+1}(t) \leq n^{-1/22+o(1)} = o\left(\frac{1}{s_\sigma}\right)$$

completing the proof. ∎

Recall that γ was the indicator function whether a graph contains a copy of $K_{3,3}$.

Lemma 7.8

Select $(r,\Gamma) \in \{(r,(K_3)_r),(r,(K_{3,3})_r),(r,(K_{3,4}^-)_{r_1})\}$ and J a spanning subgraph of Γ such that $e(J) \geq 1$. Fix $\phi : r \to V(G_f(K_{3,4}^-)_{n,i})$ with $\phi(r) = \{u,v\}$ such that $X_{\phi,J,\Gamma,\kappa,C}(i)$ is still tracked. Assuming that $\mathcal{G}_i$ and $\mathcal{H}_i$ holds then for any $e \in E(J)$ we have that

$$\sum_{k\geq 0} X_{\phi,J_e,\Gamma,\kappa_{e,k},C}(i) = \left(\frac{2q(t)x_{\phi,J_e,\Gamma,\kappa,C}(t)}{(2t)^{1-\gamma(J)}(\xi(t))^{\gamma(J)}} \pm \frac{3\mathrm{e}2q(t)f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma}\right) n^{f_2(r,J)}\sqrt{n}.$$

PROOF Let $e(t) = 2q(t)f_{r,J,\Gamma,\kappa,C}(t)$ then we have that

$$\begin{aligned}\sum_{k\geq 0} X_{\phi,J_e,\Gamma,\kappa_{e,k},C}(i) &= \sum_{k=0}^{k_{max}} X_{\phi,J_e,\Gamma,\kappa_{e,k},C}(i) \pm Z_{\phi,J_e,\Gamma,\kappa,C,e}(i)\\ &= \sum_{k=0}^{k_{max}} \left(x_{r,J_e,\Gamma,\kappa_{e,k},C}(t) \pm \frac{1}{k!}\frac{e(t)}{s_\sigma}\right) n^{f_2(r,J_e)} \pm \left(\frac{e(t)}{s_\sigma}\right) n^{f_2(r,J_e)}\\ &= \left(\frac{2q(t)x_{r,J,\Gamma,\kappa,C}(t)}{(2t)^{1-\gamma(J)}(\xi(t))^{\gamma(J)}} \pm \zeta \pm \frac{(2\mathrm{e})e(t)}{s_\sigma}\right) n^{f_2(r,J_e)}\end{aligned}$$

where $\zeta = 2\max_{t\geq 0} x_{r,J_e,\Gamma,\kappa_{e,k_{max}+1},C}(t)$ and similarly as before all we have to show is that $2\max_{t\geq 0} x_{r,J_e,\Gamma,\kappa_{e,k_{max}+1},C}(t) \leq \beta_{r,J,\Gamma,\kappa,C}/s_\sigma$. Note that

$$2x_{r,J_e,\Gamma,\kappa_{e,k_{max}+1},C}(t) \leq 2((2t)^9+1)\frac{1}{k_{max}!}\beta_{r,J,\Gamma,\kappa,C}.$$

Also $t = O(\sqrt{\log\log n})$ and thus

$$2x_{r,J_e,\Gamma,\kappa_{e,k_{max}+1},C}(t) \leq n^{-1/22+o(1)}\beta_{r,J,\Gamma,\kappa,C} \leq \frac{\beta_{r,J,\Gamma,\kappa,C}}{s_\sigma}$$

completing the proof. ∎

Note that when $e(\Gamma) - e(J) = 1$ then we have that $f_{r,J,\Gamma,k,C}(t) = f_k(t)$ and thus $\sum_{k\geq 0} f_{r,J,\Gamma,k,C}(t) = f(t)$.

7.2.4 Trend hypothesis

In the following Lemmas we verify that the trend hypothesis holds.

Lemma 7.9

For all $0 \leq i \leq m$ *and* $k \in \mathcal{I}_q$*, whenever* $\mathcal{G}_i \cap \mathcal{H}_i$ *holds there exists a smooth function* $h_k(t)$ *such that*

$$\mathbb{E}(Q_k^{\pm}(i)|\mathcal{F}_i) = \left(q_k^{\pm} \pm \frac{h_k(t)}{s_\sigma}\right)\sqrt{n}$$

and

$$f_k(t) \geq 2\int_0^t h_k(\tau)d\tau + \beta_\sigma.$$

PROOF We first consider $Q_k^+(i)$. Clearly if $k = 0$ then $Q_k^+(i) = 0$ so we are only interested in the cases when $k > 0$. In (7.2) we used the estimate

$$\mathbb{E}(Q_k^+(i)) \approx \frac{1}{Q(i)} \sum_{f \in O_{k-1}(i)} A_f(i).$$

This estimate assumed that inserting an edge creates at most one copy of $(r, (K_{3,3})_r)$ rooted at f. However this is not the case, but according to Lemma 7.4 we have that for any given open pair of vertices inserting any open pair can create at most 2 copies. Since we are in the event $\mathcal{C}_i$ there are at most $f(t)\sqrt{n}/n^{1/11}$ such open pairs. Since we are in $\mathcal{G}_i$ we have that $Q_{k-2}(i) \leq (k-1)Q_{k-1}(i)$ thus

$$\begin{aligned}\mathbb{E}(Q_k^+(i)) &= \frac{1}{Q(i)}\left(\sum_{f \in O_{k-1}(i)} (A_f(i)) \pm Q_{k-2}(i)\frac{f(t)}{n^{1/11}}\sqrt{n}\right)\\ &= \frac{1}{Q(i)}\left(\sum_{f \in O_{k-1}(i)} (A_f(i)) \pm \frac{(k-1)f(t)}{n^{1/11}}Q_{k-1}(i)\sqrt{n}\right)\\ &= \frac{1}{Q(i)}\left(\sum_{f \in O_{k-1}(i)} \left(A_f(i) \pm \frac{f(t)}{s_\sigma}\sqrt{n}\right)\right).\end{aligned}$$

We already established the following estimate in (7.1):

$$A_f(i) \approx \frac{1}{\mathrm{aut}(K_{3,3})} \sum_{\substack{r,e \in E(K_{3,3}) \\ r \neq e}} \sum_{\phi: r \to \{u,v\}} \sum_{k \geq 0} X_{\phi,(K_{3,3})_{\{r,e\}},(K_{3,3})_r,k,\emptyset}(i).$$

However some of these edges might create multiple copies of $(r, (K_{3,3})_r)$ rooted at f and thus should not be counted. Note that assuming $\mathcal{C}_i$ there are at most $f(t)\sqrt{n}/n^{1/11}$ of these and each of them is counted at most twice. Also inserting some of these edges might not just create a copy of $(r, (K_{3,3})_r)$ rooted at f it could also close f. Note that this happens if inserting an edge not only creates a copy of $(r, (K_{3,3})_r)$ rooted at f it also creates one of the following

- a copy of $(r_1, (K_{3,4}^-)_{r_1})$ rooted at f
- a second copy of $(r, (K_{3,3})_r)$ rooted at f
- a copy of $(r, (K_3)_r)$ rooted at a non-edge of the already present $k-1$ copies of $(r, (K_{3,3})_r)$ rooted at f.

Since we are in the event $\mathcal{C}_i$ there are at most $(6(k-1)+2)f(t)\sqrt{n}/n^{1/11}$ such open pairs. Therefore by applying Lemma 7.8

$$\begin{aligned} A_f(i) &= 2\left((2t)^7 2q(t)\exp(-6(2t)^2) \pm \frac{3f(t)}{s_\sigma}\right)\sqrt{n} \pm \frac{(6k-3)f(t)}{n^{1/11}}\sqrt{n} \\ &= \left(2(2t)^7 2q(t)\exp(-6(2t)^2) \pm \frac{6f(t)}{s_\sigma}\right)\sqrt{n} \pm \frac{f(t)}{s_\sigma}\sqrt{n} \\ &= \left(2(2t)^7 2q(t)\exp(-6(2t)^2) \pm \frac{7f(t)}{s_\sigma}\right)\sqrt{n}. \end{aligned}$$

Recall that we already had an $f(t)/s_\sigma$ error term and by applying Lemma 7.7 we have that

$$\mathbb{E}(Q_k^+(i)) = \frac{\sqrt{n}}{q(t) \pm \frac{3f(t)}{s_\sigma}}\left(q_{k-1}(t) \pm \frac{f_{k-1}(t)}{s_\sigma}\right)\left(2(2t)^7 2q(t)\exp(-6(2t)^2) \pm \frac{8f(t)}{s_\sigma}\right).$$

Note that for our range of t we have that $0 \le 3f(t)/(q(t)s_\sigma) \le 1/2$ and thus

$$\begin{aligned} &\frac{1}{q(t) \pm \frac{3f(t)}{s_\sigma}}\left(q_{k-1}(t) \pm \frac{f_{k-1}(t)}{s_\sigma}\right)\left(2(2t)^7 2q(t)\exp(-6(2t)^2) \pm \frac{8f(t)}{s_\sigma}\right) \\ &= \frac{1}{q(t)}\frac{1}{1 \pm \frac{3f(t)}{q(t)s_\sigma}}\left(q_{k-1}(t) \pm \frac{f_{k-1}(t)}{s_\sigma}\right)\left(2(2t)^7 2q(t)\exp(-6(2t)^2) \pm \frac{8f(t)}{s_\sigma}\right) \\ &= \frac{1}{q(t)}\left(1 \pm \frac{6f(t)}{q(t)s_\sigma}\right)\left(q_{k-1}(t) \pm \frac{f_{k-1}(t)}{s_\sigma}\right)\left(2(2t)^7 2q(t)\exp(-6(2t)^2) \pm \frac{8f(t)}{s_\sigma}\right) \\ &= \left(2(2t)^7 2q(t)\exp(-6(2t)^2)\frac{q_{k-1}(t)}{q(t)} \pm \frac{h_k^+(t)}{s_\sigma}\right) \\ &= \left(4(2t)^7\exp(-6(2t)^2)q_{k-1}(t) \pm \frac{h_k^+(t)}{s_\sigma}\right) \\ &= q_k^+(t) \pm \frac{h_k^+(t)}{s_\sigma} \end{aligned}$$

where

$$\begin{aligned} h_k^+(t) = &\frac{6f(t)q_{k-1}(t)2(2t)^7 2q(t)\exp(-6(2t)^2)}{q^2(t)} + \frac{f_{k-1}(t)2(2t)^7 2q(t)\exp(-6(2t)^2)}{q(t)} \\ &+ \frac{q_{k-1}(t)8f(t)}{q(t)} + o(1/k!). \end{aligned}$$

Using (7.4) and Lemma 7.6 gives us that:

$$h_k^+(t) \le 6\frac{q_{k-1}(t)}{q(t)}f(t) + f_{k-1}(t) + 8\frac{q_{k-1}(t)}{q(t)}f(t) = \frac{14}{(k-1)!}f(t) + f_{k-1}(t).$$

We also have that $f(t)/(k-1)! \le \mathrm{e} f_{k-1}(t)$ thus

$$h_k^+(t) \le (14\mathrm{e}+1)f_{k-1}(t) = (14\mathrm{e}+1)kf_k(t) \le 43kf_k(t).$$

Next we consider $\mathbb{E}(Q_k^-(i))$. (7.3) implies that

$$\mathbb{E}(Q_k^-(i)) \approx \frac{1}{Q(i)} \sum_{f\in O_k(i)} A_f(i) + C_f(i).$$

Next we partitioned the set of open pairs which contribute to $C_f(i)$ into three parts, $C_{f,1}(i)$, $C_{f,2}(i)$ and $C_{f,3}(i)$ depending on the structure this last edge completes. Note that inserting an open pair can possibly complete multiple structures, which would remove it from $O_k(i)$, at the same time, however since we are in the event $\mathcal{C}_i$ there are at most $(6k+2)^2 f(t)\sqrt{n}/n^{1/11} \le \log^2 n f(t)\sqrt{n}/n^{1/11}$ such pairs. Thus

$$\begin{aligned}\mathbb{E}(Q_k^-(i)) &= \frac{1}{Q(i)}\left(\sum_{f\in O_k(i)} A_f(i) + C_{f,1}(i) + C_{f,2}(i) + C_{f,3}(i) \pm \log^2 n \frac{f(t)}{n^{1/11}}\sqrt{n}\right)\\ &= \frac{1}{Q(i)}\left(\sum_{f\in O_k(i)} A_f(i) + C_{f,1}(i) + C_{f,2}(i) + C_{f,3}(i) \pm \frac{f(t)}{s_\sigma}\sqrt{n}\right).\end{aligned}$$

We will examine the terms individually. Note that a similar argument as before gives us that

$$\frac{1}{Q(i)}\sum_{f\in O_k(i)} A_f(i) = \left(4(2t)^7\exp(-6(2t)^2)q_k(t) \pm \frac{43f_k(t)}{s_\sigma}\right)\sqrt{n}.$$

Note that there is a k factor difference to our previous calculations which is due to the fact that this time we sum over the open pairs in $O_k(i)$ instead of the open pairs in $O_{k-1}(i)$ and thus would end up with an error term of $43(k+1)f_{k+1}(t) = 43f_k(t)$. We continue with the $C_{f,1}(i)$ term. Note the additional $f(t)\sqrt{n}/s_\sigma$ term at the end which is for completing multiple copies.

$$\begin{aligned}C_{f,1}(i) &= \sum_{e\in E((K_{3,4}^-)_{r_1})}\ \sum_{L\subseteq \overline{C}\subseteq E(\overline{(K_{3,4}^-)})}\ \sum_{\phi:r_1\to\{u,v\}} \frac{X_{\phi,(K_{3,4}^-)_{\{r_1,e\}},(K_{3,4}^-)_{r_1},C}(i)}{\mathrm{aut}(r_1,H_{r_1})} \pm \frac{f(t)}{s_\sigma}\sqrt{n}\\ &= \left(\frac{1}{3}(2t)^9 2q(t)\exp(-(2t)^2)^6(9+\xi(t)) \pm \frac{2\cdot 10\cdot 2^{10}\cdot 3+1}{s_\sigma}f(t)\right)\sqrt{n}\end{aligned}$$

where the 2 in the error term comes from the two options for ϕ, the 10 is for the 10 edges and 2^{10} is from the 10 non-edges which have to be considered. Therefore:

$$\begin{aligned}\frac{1}{Q(i)}\sum_{f\in O_k(i)} C_{f,1}(i) &= \frac{1}{q(t)\pm\frac{3f(t)}{s_\sigma}}\left(q_k(t)\pm\frac{f_k(t)}{s_\sigma}\right)\\ &\left(\frac{1}{3}(2t)^9 2q(t)\exp(-(2t)^2)^6(9+\xi(t))\pm\frac{2^{16}}{s_\sigma}f(t)\right)\sqrt{n}\\ &= \frac{1}{q(t)}\left(1\pm\frac{6f(t)}{q(t)s_\sigma}\right)\left(q_k(t)\pm\frac{f_k(t)}{s_\sigma}\right)\\ &\left(\frac{1}{3}(2t)^9 2q(t)\exp(-(2t)^2)^6(9+\xi(t))\pm\frac{2^{16}}{s_\sigma}f(t)\right)\sqrt{n}\\ &= \left(q_k(t)\frac{2}{3}(2t)^9\exp(-(2t)^2)^6(9+\xi(t))\pm\frac{h_1^-(t)}{s_\sigma}\right)\sqrt{n}\end{aligned}$$

where

$$\begin{aligned}h_1^-(t) &= 6\frac{q_k(t)}{q^2(t)}f(t)\frac{1}{3}(2t)^9 2q(t)\exp(-(2t)^2)^6(9+\xi(t))\\ &+\frac{f_k(t)}{q(t)}\frac{1}{3}(2t)^9 2q(t)\exp(-(2t)^2)^6(9+\xi(t)) + 2^{16}\frac{q_k(t)}{q(t)}f(t)+o(1/k!).\end{aligned}$$

We have that $2(2t)^9\exp(-(2t)^2)^6(9+\xi(t)) < 3$ for every value of t thus:

$$\begin{aligned}h_1^-(t) &\le 6\frac{f(t)}{k!}+f_k(t)+2^{16}\frac{f(t)}{k!}\\ &\le 18f_k(t)+f_k(t)+3\cdot 2^{16}f_k(t)\le 2^{18}f_k(t).\end{aligned}$$

Next we consider $C_{f,2}(i)$:

$$\begin{aligned}C_{f,2}(i) &= \frac{1}{\mathrm{aut}(K_{3,3})}\sum_{\substack{r,e\in E(K_{3,3})\\ r\neq e}}\ \sum_{\phi:r\to\{u,v\}}\ \sum_{\substack{C\subseteq E(\overline{K_{3,3}})\\ C\neq\emptyset}} X_{\phi,(K_{3,3})_{\{r,e\}},(K_{3,3})_r,C}(i)\pm\frac{f(t)}{n^{1/11}}\sqrt{n}\\ &= \left(2(2t)^7 2q(t)(1-(\exp(-(2t)^2))^6)\pm\frac{2(2^6-1)3+1}{s_\sigma}f(t)\right)\sqrt{n}.\end{aligned}$$

Therefore:

$$\begin{aligned}\frac{1}{Q(i)}\sum_{f\in O_k(i)} C_{f,2}(i) &= \frac{1}{q(t)\pm\frac{3f(t)}{s_\sigma}}\left(q_k(t)\pm\frac{f_k(t)}{s_\sigma}\right)\\ &\left(2(2t)^7 2q(t)(1-(\exp(-(2t)^2))^6)\pm\frac{2^9 f(t)}{s_\sigma}\right)\sqrt{n}\\ &= \frac{1}{q(t)}\left(1\pm\frac{6f(t)}{q(t)s_\sigma}\right)\left(q_k(t)\pm\frac{f_k(t)}{s_\sigma}\right)\\ &\left(2(2t)^7 2q(t)(1-(\exp(-(2t)^2))^6)\pm\frac{2^9 f(t)}{s_\sigma}\right)\sqrt{n}\\ &= \left(q_k(t)2(2t)^7 2(1-(\exp(-(2t)^2))^6)\pm\frac{h_2^-(t)}{s_\sigma}\right)\sqrt{n}\end{aligned}$$

where

$$\begin{aligned}h_2^-(t) &= 6\frac{q_k(t)}{q^2(t)}f(t)2(2t)^7 2q(t)(1-(\exp(-(2t)^2))^6)\\&+\frac{f_k(t)}{q(t)}2(2t)^7 2q(t)(1-(\exp(-(2t)^2))^6)+2^9\frac{q_k(t)}{q(t)}f(t)+o(1/k!)\\&\le 72(2t)^7 f_k(t)+4(2t)^7 f_k(t)+3\cdot 2^9 f_k(t)\\&\le (76(2t)^7+2^{11})f_k(t).\end{aligned}$$

Finally consider $C_{f,3}(i)$. Now we need the fact that $f \in O_k(i)$. Assume for a moment that all of the k copies of $(r,(K_{3,3})_r)$ rooted at f are vertex disjoint. As usual we have to be careful with an edge completing multiple copies of $(r_2,(K_{3,4}^-)_{r_2})$ and $(r_3,(K_{3,4}^-)_{r_3})$ rooted at f. For each pair of non-edges belonging to a copy of $(r,(K_{3,3})_r)$ rooted at f there are at most $f(t)\sqrt{n}/n^{1/11}$ edges which would create a copy of $(r,(K_3)_r)$ rooted at both of them as we are in the event $\mathcal{C}_i$. Thus there are at most $36k^2 f(t)\sqrt{n}/n^{1/22} \le f(t)\sqrt{n}/s_\sigma$ such edges. Therefore

$$\begin{aligned}C_{f,3}(i) &= \frac{1}{4}\sum_{g\in\Xi_{\phi,(K_{3,3})_{r'},\kappa_0,\emptyset}(i)}\sum_{\bar{e}\in\overline{K_{3,3}}}\sum_{\substack{r,e\in E(K_3)\\ r\neq e}}\sum_{\phi:r\to g(\bar{e})}\frac{X_{\phi,(K_3)_{\{r,e\}},(K_3)_r,\emptyset}(i)}{\mathrm{aut}(K_3)}\pm\frac{f(t)}{s_\sigma}\sqrt{n}\\&= 6k\left(2(2t)2q(t)\pm\frac{6f(t)}{s_\sigma}\right)\sqrt{n}\pm\frac{f(t)}{s_\sigma}\sqrt{n}\\&=\left(12k(2t)2q(t)\pm\frac{(36k+1)f(t)}{s_\sigma}\right)\sqrt{n}.\end{aligned}$$

Since we are in the event $\mathcal{B}_i$ there are at most $f(t)n^\varepsilon$ not non-edge disjoint copies in the graph hence:

$$\begin{aligned}\sum_{f\in O_k(i)}C_{f,3}(i) &= Q_k(i)\left(12k(2t)2q(t)\pm\frac{(36k+1)f(t)}{s_\sigma}\right)\sqrt{n}\\&\quad\pm f(t)n^\varepsilon\left(2(2t)2q(t)\pm\frac{6f(t)}{s_\sigma}\right)\sqrt{n}\\&\quad=(Q_k(i)\pm f(t)n^\varepsilon)\left(12k(2t)2q(t)\pm\frac{(36k+1)f(t)}{s_\sigma}\right)\sqrt{n}\\&\quad=\left(Q_k(i)\pm\frac{f_k(t)}{s_\sigma}n^2\right)\left(12k(2t)2q(t)\pm\frac{(36k+1)f(t)}{s_\sigma}\right)\sqrt{n}\end{aligned}$$

as $n^\varepsilon = o(n^2/(k_{max}!s_\sigma))$. Therefore

$$\begin{aligned}\frac{1}{Q(i)}\sum_{f\in O_k(i)}C_{f,3}(i) &= \frac{1}{q(t)}\left(1\pm\frac{6f(t)}{s_\sigma}\right)\left(q_k(t)\pm\frac{2f_k(t)}{s_\sigma}\right)\\&\quad\left(12k(2t)2q(t)\pm\frac{(36k+1)f(t)}{s_\sigma}\right)\sqrt{n}\\&=\left(12k(2t)2q_k(t)\pm\frac{h_3^-(t)}{s_\sigma}\right)\sqrt{n}\end{aligned}$$

where

$$\begin{aligned}h_3^-(t) &= 6\frac{q_k(t)}{q^2(t)}f(t)24k(2t)q(t) + 2\frac{f_k(t)}{q(t)}24k(2t)q(t) + (36k+1)\frac{q_k(t)}{q(t)}f(t) + o(1/k!)\\ &= (960kt + 108k + 3)f_k(t).\end{aligned}$$

Recall that $Q_k(i) = (1+o(1))q_k(t)$ and $Q(i) = (1+o(1))q(t)$ for $0 \le t \le t(m)$ when $\mathcal{G}_i$ holds thus

$$\frac{1}{Q(i)}Q_k(i)\frac{f(t)}{s_\sigma}\sqrt{n} \le \frac{(1+o(1))q_k(t)}{q(t)}\frac{f(t)}{s_\sigma}\sqrt{n} \le \frac{3f_k(t)}{s_\sigma}\sqrt{n}.$$

Define $h_k^-(t) = 43f_k(t) + h_1^-(t) + h_2^-(t) + h_3^-(t) + 12f_k(t)$ then we have shown that

$$\mathbb{E}(Q_k^-(i)) = \left(q_k^-(t) \pm \frac{h_k^-(t)}{s_\sigma}\right)\sqrt{n}.$$

Also let $h_k(t) = h_k^+(t) + h_k^-(t)$ and if we show that $f_k(t) \ge 2\int_0^t h_\sigma(\tau) + \beta_k$ then we have completed the proof. Note that this is implied if $f_k(0) = \beta_k$, which holds by definition and $f_k'(t) \ge 2h_k(t)$. Note that

$$\begin{aligned}h_k(t) &\le (43k + 43 + 2^{18} + 76(2t)^7 + 2^{11} + 960kt + 108k + 2 + 3)f_k(t)\\ &\le (960t + 200)k_{max}f_k(t).\end{aligned}$$

We also have that

$$\begin{aligned}f_k'(t) &= q'(t)\exp(24W(t^2+t))/k! + 24Wk_{\max}(2t+1)f_k(t)\\ &= (1+o(1))24Wk_{\max}(2t+1)f_k(t)\end{aligned}$$

as $q'(t)\exp(24W(t^2+t)) = O(t^9+t+1) = O((\log\log n)^5) = o(k_{max})$. Therefore $f_k'(t) > 2h_k(t)$ as $W > 40$. ■

Lemma 7.10

Whenever $\mathcal{G}_i \cap \mathcal{H}_i$ holds there exists a smooth function $h_o(t)$ such that

$$\mathbb{E}(Y_o^\pm(i)|\mathcal{F}_i) = \left(0 \pm \frac{h_o(t)}{s_\sigma}\right)\sqrt{n}$$

and

$$f_o(t) \ge 2\int_0^t h_o(\tau)d\tau + \beta_o.$$

PROOF Note that $Y_o^-(i) = 0$ so we only need the expected increase. Since we are in the event $\mathcal{G}_i$ we have that $Q_k(i)$ is concentrated around $q_k(t)n^2$and $X_\sigma(i)$ is concentrated around $x_\sigma(t)S_\sigma$. Also note that

$$\frac{q_{k_{max}-1}}{q(t)}\frac{\sqrt{n}}{n^{1/11}} = o\left(\frac{q_{k_{max}}}{q(t)}\sqrt{n}\right)$$

therefore

$$\mathbb{E}(Y_o^+) = (1+o(1))\frac{1}{q(t)}\left(q_{k_{max}}((2t)^7)2q(t)\exp(-(2t)^2)^6\right)\sqrt{n}$$
$$\leq \frac{q(t)}{k_{\max}!}\sqrt{n} \leq n^{-1/22+o(1)}\sqrt{n} = o\left(\frac{f_o'(t)}{s_\sigma}\right)\sqrt{n}$$

which in addition to the fact that $f_o(0) = \beta_o$ completes the proof. ∎

Next we concentrate on the random variables $X_{\phi,J,\Gamma,\kappa,C}(i)$.

Lemma 7.11

We have that for every $1 \leq i \leq m$ assuming $\mathcal{G}_i \cap \mathcal{H}_i$ hold then:

$$\mathbb{E}(X^+_{\phi,J,\Gamma,\kappa,C}(i)) = \frac{1}{Q(i)}\sum_{e\in E(J)}\sum_{k\geq 0} X_{\phi,J_e,\Gamma,\kappa_{e,k},C}(i)$$
$$+\frac{1}{Q(i)}\sum_{f\in E(\Gamma)\backslash E(J)}\sum_{g\in\Xi_{\phi,J,\Gamma,\kappa_f^-,C}(i)} A_{g(f)}(i)$$
$$+\frac{1}{Q(i)}\sum_{c\in C}\sum_{g\in\Xi_{\phi,J,\Gamma,\kappa,C\backslash c}(i)}\sum_{\substack{r,e\in E(K_3)\\ r\neq e}}\sum_{\phi:r\to g(c)}\frac{X_{\phi,(K_3)_{\{r,e\}},(K_3)_r,\emptyset}(i)}{\mathrm{aut}(K_3)}$$
$$\pm\frac{1}{Q(i)}\left(\frac{q(t)f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma}\right)n^{f_2(r,J)}\sqrt{n}.$$

PROOF First note that we are considering the mappings $g \in \Xi_{\phi,J,\Gamma,\kappa,C}(i+1)$ such that $g \notin \Xi_{\phi,J,\Gamma,\kappa,C}(i)$. The majority of these g come form one of the following sets $\Xi_{\phi,J_e,\Gamma,\kappa_{e,k},C}(i)$ for some $e \in E(J)$, $\Xi_{\phi,J,\Gamma,\kappa_f^-,C}(i)$ for some $f \in E(\Gamma)\backslash E(J)$ or $\Xi_{\phi,J,\Gamma,\kappa,C\backslash c}(i)$ for some $c \in C$.

However these are not the only sets g could have originated from, as it is possible for the same edge to create multiple constructions e.g. there are some open pairs, which when inserted as edges would transfer g from $\Xi_{\phi,J,\Gamma,\kappa_f^-,C\backslash c}(i)$ for $f \in E(\Gamma)\backslash E(J)$ and $c \in C$ to $\Xi_{\phi,J,\Gamma,\kappa,C}(i)$ as inserting this open pair would create a copy of $(r,(K_{3,3})_r)$ rooted at $g(f)$ and a copy of $(r,(K_3)_r)$ rooted at $g(c)$. Note that if g was transferred from $\Xi_{\phi,J',\Gamma,\kappa',C'}(i)$ to $\Xi_{\phi,J,\Gamma,\kappa,C}(i+1)$ during step i then $e(J)-e(J') \leq 1$ and according to Lemma 7.4 $\kappa(f)-\kappa(f') \leq 2$ for every $f \in E(\Gamma)\backslash E(J)$.

Start with the case when $J' \neq J$ and fix $e \in E(J)$. Since we are in the event $\mathcal{D}_i$ there are at most $O(n^{f_2(r,J)}n^{\varepsilon})$ elements such that inserting $g(e)$ in step i would transfer g to $\Xi_{\phi,J,\Gamma,\kappa,C}(i+1)$. Note that $k_{max}! = n^{-1/22+o(1)}$ and that $e(\Gamma)-e(J) \leq 10$. Hence

$$n^{\varepsilon} = n^{1/66} = o(\sqrt{n}(n^{-1/22+o(1)})^{10}) = o(q(t)f_{r,J,\Gamma,\kappa,C}(t)\sqrt{n}/s_\sigma)$$

thus this has little affect on the expected value.

For the remaining $\Xi_{\phi,J,\Gamma,\kappa',C'}(i)$ such that $g \in \Xi_{\phi,J,\Gamma,\kappa',C'}(i)$ can be transferred to $\Xi_{\phi,J,\Gamma,\kappa,C}(i+1)$ we have that in order for g to be transferred to $\Xi_{\phi,J,\Gamma,\kappa,C}(i+1)$ at least one of the following has to happen

- inserting an edge has to create more than one copy of $(r,(K_{3,3})_r)$ rooted at $g(f)$ for some $f \in E(\Gamma)\backslash E(J)$
- inserting an edge has to create a copy of $(r,(K_{3,3})_r)$ rooted both at $g(f)$ and $g(f')$ for some distinct $f, f' \in E(\Gamma)\backslash E(J)$
- inserting an edge has to create a copy of $(r,(K_{3,3})_r)$ rooted at $g(f)$ and a copy of $(r,(K_3)_r)$ rooted at $g(c)$ for some $f \in E(\Gamma)\backslash E(J)$ and $c \in C$
- inserting an edge has to create a copy of $(r,(K_3)_r)$ rooted at $g(c)$ and $g(c')$ for some distinct $c, c' \in C$.

Since we are in the event $\mathcal{C}_i$ for a fixed g there are $O(f(t)\sqrt{n}/n^{1/11})$ open pairs which would satisfy any of the above conditions. We also have that $x_{r,J,\Gamma,\kappa',C'}(t)f(t) \leq 2^{10}(t^9+1)q(t)^{e(\Gamma)-e(J)}\beta_{r,J,\Gamma,\kappa',C'}f(t) \leq (\log\log n)^5 \log^{20} nq(t)f_{r,J,\Gamma,\kappa,C}(t)$ and thus this adds an error term of

$$O\left(\log^{21} n \frac{q(t)f_{r,J,\Gamma,\kappa,C}(t)n^{f_2(r,J)}\sqrt{n}}{n^{1/11}}\right) = o\left(\frac{q(t)f_{r,J,\Gamma,\kappa,C}(t)n^{f_2(r,J)}\sqrt{n}}{s_\sigma}\right).$$

On the other hand we overestimate the change as not every combination we count would actually result in a new element in $\Xi_{\phi,J,\Gamma,\kappa,C}(i+1)$. Similarly as before this is due to the fact that when inserting an edge transfers g from $\Xi_{\phi,J',\Gamma,\kappa',C'}(i)$ to $\Xi_{\phi,J,\Gamma,\kappa,C}(i+1)$ more than one of J', κ' and C' can change. Similarly as before we have to take into account the case that for some $g \in \Xi_{\phi,J_e,\Gamma,\kappa_{e,k},C}(i)$ inserting the edge $g(e)$ could affect the status of $g(f)$ for some $f \in E(\Gamma)\backslash E(J)$. In the previous section we considered the case when the value of $\kappa(f)$ was increased, now we also have to consider the case when it is closed. However since we are in $\mathcal{D}_i$ there are $O(n^{f_2(r,J)}n^{\varepsilon}) = o(q(t)f_{r,J,\Gamma,\kappa,C}(t)n^{f_2(r,J)}\sqrt{n}/s_\sigma)$ such copies. Now fix $g \in \Xi_{\phi,J,\Gamma,\kappa_f^-,C}(i)$ then there are edges counted in $A_{g(f)}(i)$ such that inserting that edge does not result in $g \in \Xi_{\phi,J,\Gamma,\kappa,C}(i+1)$ as inserting the edge might transfer $g(f)$ to $O_{\kappa(f)+1}(i+1)$ or change the status of $g(f')$ for some $f' \in E(\Gamma)\backslash(E(J)\cup\{f\})$ or the status of some $\overline{c} \in \overline{C}$. Since we are in the event $\mathcal{C}_i$ for every g there are $O(f(t)\sqrt{n}/n^{1/11})$ non-edges which when inserted any of these would happen. Therefore in total

$$O(x_{\phi,J,\Gamma,\kappa_f^-,C}(t)n^{f_2(r,J)}f(t)\sqrt{n}/n^{1/11}) = o(q(t)f_{r,J,\Gamma,\kappa,C}(t)n^{f_2(r,J)}\sqrt{n}/s_\sigma).$$

Finally a similar argument gives that the same bound holds when $g \in \Xi_{\phi,J,\Gamma,\kappa,C\setminus c}(i)$ for some $c \in C$. Therefore the total affect on the expected value is at most

$$\frac{1}{Q(i)} \frac{q(t) f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma} n^{f_2(r,J)} \sqrt{n}.$$

■

Lemma 7.12

For all $0 \leq i \leq m$ and $(\phi, J, \Gamma, \kappa, C) \in \mathcal{I}_x$, whenever $\mathcal{G}_i \cap \mathcal{H}_i$ holds there exists a smooth function $h_{r,J,\Gamma,\kappa,C}(t)$ such that

$$\mathbb{E}(X^{\pm}_{\phi,J,\Gamma,\kappa,C}(i)|\mathcal{F}_i) = \left(x^{\pm}_{r,J,\Gamma,\kappa,C}(t) \pm \frac{h_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma}\right) \frac{n^{f_2(r,J)}}{n^{3/2}}$$

and

$$f_{r,J,\Gamma,\kappa,C}(t) \geq 2 \int_o^t h_{\phi,J,\Gamma,\kappa,C}(\tau) d\tau + \beta_{r,J,\Gamma,\kappa,C}.$$

PROOF In Lemma 7.11 we have established that

$$\begin{aligned}
\mathbb{E}(X^{+}_{\phi,J,\Gamma,\kappa,C}(i)) &= \frac{1}{Q(i)} \sum_{e \in E(J)} \sum_{k \geq 0} X_{\phi,J_e,\Gamma,\kappa_{e,k},C}(i) \\
&+ \frac{1}{Q(i)} \sum_{f \in E(\Gamma)\setminus E(J)} \sum_{g \in \Xi_{\phi,J,\Gamma,\kappa^-_f,C}(i)} A_{g(f)}(i) \\
&+ \frac{1}{Q(i)} \sum_{c \in C} \sum_{g \in \Xi_{\phi,J,\Gamma,\kappa,C\setminus c}(i)} \sum_{\phi : r \to g(c)} \sum_{\substack{r,e \in E(K_3) \\ r \neq e}} \frac{X_{\phi,(K_3)_{\{r,e\}},(K_3)_r,\emptyset}(i)}{\mathrm{aut}(K_3)} \\
&\pm \frac{1}{Q(i)} \left(\frac{q(t) f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma}\right) n^{f_2(r,J)} \sqrt{n}
\end{aligned}$$

and now we will analyse the individual terms one by one. Assume $K_{3,3} \not\subseteq J$ then we have that

$$\begin{aligned}
&\frac{1}{Q(i)} \left(\sum_{e \in E(J)} \sum_{k \geq 0} X_{\phi,J_e,\Gamma,\kappa_{e,k},C}(i) \right) \\
&= \frac{1}{q(t)} \left(1 \pm \frac{6 f(t)}{q(t) s_\sigma}\right) e(J) \left(\frac{2q(t) x_{r,\phi,J,\Gamma,\kappa,C}}{2t} \pm \frac{3\mathrm{e}2q(t) f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma}\right) \frac{n^{f_2(r,J)}}{n^{3/2}} \\
&= \left(e(J) \frac{x_{r,J,\Gamma,\kappa,C}(t)}{t} \pm \frac{h^+_{1,1}(t)}{s_\sigma}\right) \frac{n^{f_2(r,J)}}{n^{3/2}}
\end{aligned}$$

where

$$h^+_{1,1}(t) = 6e(J) \frac{x_{r,J,\Gamma,\kappa,C}(t)}{t q(t)} f(t) + 6e(J)\mathrm{e} f_{r,J,\Gamma,\kappa,C}(t) + o(\beta_{r,J,\Gamma,\kappa,C}).$$

Note that when $K_{3,3} \not\subseteq J$ then

$$x_{r,J,\Gamma,\kappa,C}(t) \leq (2t)^{e(J)} \prod_{f \in E(\Gamma) \setminus E(J))} (2q_{\kappa(f)}(t)) \leq (2t)^{e(J)} (2q(t))^{e(\Gamma)-e(J)} \beta_{r,J,\Gamma,\kappa,C}$$

and that in our case $e(J) \geq 1$ and $e(\Gamma) - e(J) \geq 1$ as we only track variables with at least one open edge. Therefore

$$\begin{aligned} \frac{x_{r,J,\Gamma,\kappa,C}(t)}{tq(t)} &\leq 4(2t)^{e(J)-1}(2q(t))^{e(\Gamma)-e(J)-1}\beta_{r,J,\Gamma,\kappa,C} \\ &\leq 4((2t)^8 + 1)(2q(t))^{e(\Gamma)-e(J)-1}\beta_{r,J,\Gamma,\kappa,C}. \end{aligned}$$

We also have that $f(t)(2q(t))^{e(\Gamma)-e(J)-1}\beta_{r,J,\Gamma,\kappa,C} \leq \mathrm{e} f_{r,J,\Gamma,\kappa,C}(t)$ and thus

$$h^+_{1,1}(t) \leq 24e(J)\mathrm{e}((2t)^8 + 2)f_{r,J,\Gamma,\kappa,C}(t) \leq 600((2t)^8 + 2)f_{r,J,\Gamma,\kappa,C}(t).$$

On the other hand if $K_{3,3} \subseteq J$ then since we only track one such random variable we have that

$$\begin{aligned} &\frac{1}{Q(i)} \sum_{e \in E(J)} \sum_{k \geq 0} X_{\phi,J_e,\Gamma,\kappa_{e,k},C}(i) \\ &= \frac{1}{q(t)} \left(1 \pm \frac{6f(t)}{q(t)s_\sigma}\right) e(J) \left(\frac{2q(t)x_{r,J,\Gamma,\kappa,C}(t)}{2t\xi(t)} \pm \frac{3\mathrm{e}2q(t)f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma}\right) \frac{n^{f_2(r,J)}}{n^{3/2}} \\ &= \left(e(J)\frac{x_{r,J,\Gamma,\kappa,C}(t)}{t\xi(t)} \pm \frac{h^+_{1,2}(t)}{s_\sigma}\right) \frac{n^{f_2(r,J)}}{n^{3/2}} \end{aligned}$$

where

$$\begin{aligned} h^+_{1,2}(t) &= 6e(J)\frac{x_{r,J,\Gamma,\kappa,C}(t)}{2tq(t)\xi(t)} f(t) + 6e(J)\mathrm{e} f_{r,J,\Gamma,\kappa,C}(t) + o(\beta_{r,J,\Gamma,\kappa,C}) \\ &\leq 600((2t)^8 + 2)f_{r,J,\Gamma,\kappa,C}(t) \end{aligned}$$

due to a similar argument as before. Let $h^+_1 = 600((2t)^8 + 2)f_{r,J,\Gamma,\kappa,C}(t)$ then clearly both $h^+_{1,1}$ and $h^+_{1,2}$ are less than h^+_1. The second term is

$$\begin{aligned} &\frac{1}{Q(i)} \sum_{f \in E(\Gamma) \setminus E(J)} \sum_{g \in \Xi_{\phi,J,\Gamma,\kappa^-_f,C}(i)} A_{g(f)}(i) \\ &= \sum_{f \in E(\Gamma) \setminus E(J)} \frac{1}{q(t)} \left(1 \pm \frac{6f(t)}{q(t)s_\sigma}\right) \left(\frac{q_{\kappa(f)-1}(t)x_{r,J,\Gamma,\kappa,C}(t)}{q_{\kappa(f)}(t)} \pm \frac{\kappa(f)f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma}\right) \\ &\left(2(2t)^7 2q(t)\exp(-6(2t)^2) \pm \frac{3f(t)}{s_\sigma}\right) \frac{n^{f_2(r,J)}}{n^{3/2}} \\ &= \left(\sum_{f \in E(\Gamma) \setminus E(J)} 2(2t)^7 2\exp(-6(2t)^2) \frac{q_{\kappa(f)-1}(t)x_{r,J,\Gamma,\kappa,C}(t)}{q_{\kappa(f)}(t)} \pm \frac{h^+_2(t)}{s_\sigma}\right) \frac{n^{f_2(r,J)}}{n^{3/2}} \end{aligned}$$

where

$$h_2^+(t) = \sum_{f\in E(\Gamma)\setminus E(J)} 6\frac{f(t)}{q(t)}2(2t)^7 2\exp(-6(2t)^2)\frac{q_{\kappa(f)-1}(t)x_{r,J,\Gamma,\kappa,C}(t)}{q_{\kappa(f)}(t)}$$
$$+\kappa(f)f_{r,J,\Gamma,\kappa,C}(t)2(2t)^7 2q(t)\exp(-6(2t)^2) + \frac{1}{q(t)}\frac{q_{\kappa(f)-1}(t)x_{r,J,\Gamma,\kappa,C}(t)}{q_{\kappa(f)}(t)}3f(t)$$
$$+ o(\beta_{r,J,\Gamma,\kappa,C}).$$

Note that

$$\frac{q_{\kappa(f)-1}(t)x_{r,J,\Gamma,\kappa,C}(t)}{q_{\kappa(f)}(t)} \leq \kappa(f)((2t)^9+1)\left(\frac{(2t)^8}{4}\exp(-(2t)^6)\right)^k q(t)^{e(\Gamma)-e(J)}\beta_{r,J,\Gamma,\kappa,C}.$$

Now if $\kappa(f) = 1$ then $(2t)^9 + 1 \leq k_{max}$ otherwise we have an $\exp(-6(2t)^2)$ factor and $(2t)^8((2t)^9+1)\exp(-6(2t)^2) < 1/4$. Therefore

$$\begin{aligned} h_2^+(t) &\leq \sum_{f\in E(\Gamma)\setminus E(J)} (12\mathrm{e}\kappa(f) + \kappa(f) + 6\mathrm{e}k_{max})f_{r,J,\Gamma,\kappa,C}(t) \\ &\leq e(\Gamma)(55k_{max})f_{r,J,\Gamma,\kappa,C}(t) \\ &\leq 550k_{max}f_{r,J,\Gamma,\kappa,C}(t). \end{aligned}$$

Note that for the third term we only consider random variables which are still tracked. This is due to the fact that for a fixed $c \in C$ and $g \in \Xi_{\phi,J,\Gamma,\kappa,C\setminus c}(i)$ if $g(c)$ is a non-edge in a copy of $K_{3,3}$ then $c \notin \overline{C}$. The third term is

$$\frac{1}{Q(i)}\sum_{c\in C}\sum_{g\in\Xi_{\phi,J,\Gamma,\kappa,C\setminus c}(i)}\sum_{\substack{r,e\in E(K_3)\\ r\neq e}}\sum_{\phi:r\to g(c)}\frac{X_{\phi,(K_3)_{\{r,e\}},(K_3)_r,\emptyset}(i)}{\mathrm{aut}(K_3)}$$
$$= |C|\frac{1}{q(t)}\left(1\pm\frac{6f(t)}{q(t)s\sigma}\right)\left(\frac{x_{\phi,J,\Gamma,\kappa,C}(t)\exp(-(2t)^2)}{1-\exp(-(2t)^2)}\pm\frac{f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma}\right)$$
$$\left(2(2t)2q(t)\pm\frac{3f(t)}{s_\sigma}\right)\frac{n^{f_2(r,J)}}{n^{3/2}}$$
$$= \left(|C|\frac{x_{\phi,J,\Gamma,\kappa,C}(t)\exp(-(2t)^2)}{1-\exp(-(2t)^2)}2(2t)2 \pm \frac{h_3^+(t)}{s_\sigma}\right)\frac{n^{f_2(r,J)}}{n^{3/2}}$$

where

$$h_3^+(t) = |C|\frac{6f(t)}{q(t)}\frac{x_{\phi,J,\Gamma,\kappa,C}(t)\exp(-(2t)^2)}{1-\exp(-(2t)^2)}2(2t)2$$
$$+|C|f_{r,J,\Gamma,\kappa,C}(t)2(2t)2 + 3|C|\frac{f(t)}{q(t)}\frac{x_{\phi,J,\Gamma,\kappa,C}(t)\exp(-(2t)^2)}{1-\exp(-(2t)^2)}.$$

Note that when $|C| \geq 1$ then $x_{\phi,J,\Gamma,\kappa,C}(t) \leq (2t)^{e(J)} q(t)^{e(\Gamma)-e(J)}(1-\exp(-(2t)^2))$ and that $(2t)^a e(-4t^2) \leq 15$ when $a \leq 10$ thus

$$\begin{aligned} h_3^+(t) &\leq |C|(4320t + 8t + 270) f_{r,J,\Gamma,\kappa,C}(t) \\ &\leq |C|(4328t + 270) f_{r,J,\Gamma,\kappa,C}(t) \\ &\leq (47608t + 2970) f_{r,J,\Gamma,\kappa,C}(t) \end{aligned}$$

as $|C| \leq 11$.

Finally

$$\begin{aligned} \frac{1}{Q(i)} \left(\frac{q(t) f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma} \right) n^{f_2(r,J)} \sqrt{n} &\leq \frac{2}{q(t)} \left(\frac{q(t) f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma} \right) \frac{n^{f_2(r,J)}}{n^{3/2}} \\ &\leq \frac{2 f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma} \frac{n^{f_2(r,J)}}{n^{3/2}}. \end{aligned}$$

Define $h^+(t) = h_1^+(t) + h_2^+(t) + h_3^+(t) + 2f_{r,J,\Gamma,\kappa,C}(t)$ and note that

$$h^+(t) \leq (600(2t)^8 + 550k_{max} + 47608t + 2970) f_{r,J,\Gamma,\kappa,C}(t)$$

and that we have shown that

$$X^+_{\phi,J,\Gamma,\kappa,C}(i) = \left(x^+_{\phi,J,\Gamma,\kappa,C}(t) \pm \frac{h^+(t)}{s_\sigma} \right) \frac{n^{f_2(r,J)}}{n^{3/2}}.$$

Now for $X^-_{\phi,J,\Gamma,\kappa,C}(i)$. We have that

$$\begin{aligned} \mathbb{E}(X^-_{\phi,J,\Gamma,\kappa,C}(i)) = &\frac{1}{Q(i)} \sum_{f \in E(\Gamma) \setminus E(J)} \sum_{g \in \Xi_{\phi,J,\Gamma,\kappa,C}(i)} \left(C_{g(f)}(i) + A_{g(f)}(i) \right) \\ &+ \frac{1}{Q(i)} \sum_{c \in \overline{C}} \sum_{g \in \Xi_{\phi,J,\Gamma,\kappa,C}(i)} \sum_{\phi: r \to g(c)} \sum_{\substack{r,e \in E(K_3) \\ r \neq e}} \frac{X_{\phi,(K_3)_{\{r,e\}},(K_3)_r,\emptyset}(i)}{\mathrm{aut}(K_3)} \\ &\pm \frac{1}{Q(i)} \sum_{g \in \Xi_{\phi,J,\Gamma,\kappa,C}(i)} \left(\frac{f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma} \right) \sqrt{n} n^{f_2 r,J} \end{aligned}$$

where the last term is the usual error term for counting the same open pair multiple times. Note that from this point on we will not consider the terms resulting from overlaps as we are in the event $\mathcal{H}_i$, we have that the total error caused by them is at most $f_{r,J,\Gamma,\kappa,C}(t)\sqrt{n} n^{f_2(r,J)}/n^{1/11}$ and that we have to consider at most $\log^2 n$ pairs, thus the last term already takes them into account.

Fix $f \in E(\Gamma)\backslash E(J)$. Now we examine each term in the sum separately. The first term is

$$\begin{aligned}
&\frac{1}{Q(i)} \sum_{g \in \Xi_{\phi,J,\Gamma,\kappa,C}(i)} A_{g(f)}(i) \\
&= \frac{1}{q(t)} \left(1 \pm \frac{6f(t)}{q(t)s_\sigma}\right) \left(x_{r,J,\Gamma,\kappa,C}(t) \pm \frac{f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma}\right) \\
&\left(2(2t)2q(t) \pm \frac{6f(t)}{s_\sigma}\right) \frac{n^{f_2(r,J)}}{n^{3/2}} \\
&= \left(\frac{1}{q(t)} x_{r,J,\Gamma,\kappa,C}(t)2(2t)2q(t) \pm \frac{h^-_{a,f}(t)}{s_\sigma}\right) \frac{n^{f_2(r,J)}}{n^{3/2}}
\end{aligned}$$

where

$$\begin{aligned}
h^-_{a,f}(t) &= \frac{6f(t)}{q(t)} x_{r,J,\Gamma,\kappa,C}(t)2(2t)2 + f_{r,J,\Gamma,\kappa,C}(t)2(2t)2 \\
&+ x_{r,J,\Gamma,\kappa,C}(t)\frac{6f(t)}{q(t)} + o(\beta_{r,J,\Gamma,\kappa,C}(t)) \\
&\leq (144((2t)^9+1)(2t) + 8t + 36((2t)^9+1))f_{r,J,\Gamma,\kappa,C}(t).
\end{aligned}$$

The second one is

$$\begin{aligned}
&\frac{1}{Q(i)} \sum_{g \in \Xi_{\phi,J,\Gamma,\kappa,C}(i)} C_{g(f),1}(i) \\
&= \frac{1}{q(t)} \left(1 \pm \frac{6f(t)}{q(t)s_\sigma}\right) \left(x_{r,J,\Gamma,\kappa,C}(t) \pm \frac{f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma}\right) \\
&\left(\frac{1}{3}(2t)^9 2q(t)\exp(-(2t)^2)^6(9+\xi(t)) \pm \frac{2^{16}}{s_\sigma} f(t)\right) \frac{n^{f_2(r,J)}}{n^{3/2}} \\
&= \left(x_{r,J,\Gamma,\kappa,C}(t)\left(\frac{1}{3}(2t)^9 2\exp(-(2t)^2)^6(9+\xi(t))\right) \pm \frac{h^-_{1,f}(t)}{s_\sigma}\right) \frac{n^{f_2(r,J)}}{n^{3/2}}
\end{aligned}$$

where

$$\begin{aligned}
h^-_{1,f}(t) &= \frac{6f(t)}{q(t)} x_{r,J,\Gamma,\kappa,C}(t) \frac{1}{3}(2t)^9 2\exp(-(2t)^2)^6(9+\xi(t)) \\
&+ \frac{1}{3}(2t)^9 2\exp(-(2t)^2)^6(9+\xi(t)) f_{r,J,\Gamma,\kappa,C}(t) \\
&+ \frac{1}{q(t)} x_{r,J,\Gamma,\kappa,C}(t) 2^{16} f(t) + o(\beta_{r,J,\Gamma,\kappa,C}) \\
&\leq (36 + 1 + 2^{17}\mathrm{e}((2t)^9+1)) f_{r,J,\Gamma,\kappa,C}(t) \\
&\leq 2^{20}((2t)^9+1) f_{r,J,\Gamma,\kappa,C}(t).
\end{aligned}$$

The third term is

$$\begin{aligned}
&\frac{1}{Q(i)}\sum_{g\in\Xi_{\phi,J,\Gamma,\kappa,C}(i)} C_{g(f),2}(i)\\
&= \frac{1}{q(t)}\left(1 \pm \frac{6f(t)}{q(t)s_\sigma}\right)\left(x_{r,J,\Gamma,\kappa,C}(t) \pm \frac{f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma}\right)\\
&\left(2(2t)^7 2q(t)(1-(\exp(-(2t)^2))^6) \pm \frac{2^9 f(t)}{s_\sigma}\right) \pm \frac{f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma}\frac{n^{f_2(r,J)}}{n^{3/2}}\\
&= \left(\frac{1}{q(t)}x_{r,J,\Gamma,\kappa,C}(t)2(2t)^7 2q(t)(1-(\exp(-(2t)^2))^6) \pm \frac{h^-_{2,f}(t)}{s_\sigma}\right)\frac{n^{f_2(r,J)}}{n^{3/2}}
\end{aligned}$$

where

$$\begin{aligned}
h^-_{2,f}(t) &= 6\frac{f(t)}{q(t)}x_{r,J,\Gamma,\kappa,C}(t)2(2t)^7 2(1-(\exp(-(2t)^2))^6)\\
&+ f_{r,J,\Gamma,\kappa,C}(t)2(2t)^7 2(1-(\exp(-(2t)^2))^6)\\
&+ 2^9\frac{f(t)}{q(t)}x_{r,J,\Gamma,\kappa,C}(t) + o(\beta_{r,J,\Gamma,\kappa,C})\\
&\leq (2^{11}((2t)^{16}+1))f_{r,J,\Gamma,\kappa,C}(t).
\end{aligned}$$

The fourth term is:

$$\begin{aligned}
&\frac{1}{Q(i)}\sum_{g\in\Xi_{\phi,J,\Gamma,\kappa,C}(i)} C_{g(f),3}(i)\\
&= \frac{1}{q(t)}\left(1 \pm \frac{6f(t)}{q(t)s_\sigma}\right)\left(x_{r,J,\Gamma,\kappa,C}(t) \pm \frac{f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma}\right)\\
&\left(6\kappa(f)2(2t)(2q(t)) \pm \frac{18\kappa(f)f(t)}{s_\sigma}\right)\frac{n^{f_2(r,J)}}{n^{3/2}}\\
&= \left(\frac{1}{q(t)}x_{r,J,\Gamma,\kappa,C}(t)6\kappa(f)2(2t)(2q(t)) \pm \frac{h^-_{3,f}(t)}{s_\sigma}\right)\frac{n^{f_2(r,J)}}{n^{3/2}}
\end{aligned}$$

where

$$\begin{aligned}
h^-_{3,f}(t) &= 6\frac{f(t)}{q(t)}x_{r,J,\Gamma,\kappa,C}(t)6\kappa(f)4(2t)\\
&+ f_{r,J,\Gamma,\kappa,C}(t)6\kappa(f)4(2t) + 18\kappa(f)\frac{f(t)}{q(t)}x_{r,J,\Gamma,\kappa,C}(t) + o(\beta_{r,J,\Gamma,\kappa,C}(t))
\end{aligned}$$

Note that if $\kappa(f) \geq 1$ then

$$x_{\phi,J,\Gamma,\kappa,C}(t)f(t) \leq 2\mathrm{e}((2t)^9+1)(2t)^8\exp(-(2t)^2)^6 f_{r,J,\Gamma,\kappa,C}(t)/4$$

and (7.7) implies that

$$\begin{aligned}
h^-_{3,f}(t) &\leq (972\kappa(f) + 96\kappa(f)t)f_{r,J,\Gamma,\kappa,C}\\
&\leq (972k_{max} + 96k_{max}t)f_{r,J,\Gamma,\kappa,C}.
\end{aligned}$$

Finally fix $c \in \overline{C}$ and note that for every $g \in \Xi_{\phi,J,\Gamma,\kappa,C}(i)$ we might not be counting a constant number of these, when the edges which would create the mutual neighbour are also required for the completion of Γ. Therefore

$$\begin{aligned}
&\frac{1}{Q(i)} \sum_{g\in\Xi_{\phi,J,\Gamma,\kappa,C}(i)} \sum_{\phi:r\to g(c)} \sum_{\substack{r,e\in E(K_3)\\ r\neq e}} \frac{X_{\phi,(K_3)_{\{r,e\}},(K_3)_r,\emptyset}(i)}{\mathrm{aut}(K_3)} \\
&= \frac{1}{q(t)}\left(1 \pm \frac{6f(t)}{q(t)s_\sigma}\right)\left(x_{r,J,\Gamma,\kappa,C}(t) \pm \frac{f_{r,J,\Gamma,\kappa,C}(t)}{s_\sigma}\right)\left(2(2t)2q(t) \pm \frac{6f(t)}{s_\sigma}\right)\frac{n^{f_2(r,J)}}{n^{3/2}} \\
&= \left(4(2t)x_{r,J,\Gamma,\kappa,C}(t) \pm \frac{h_c^-(t)}{s_\sigma}\right)\frac{n^{f_2(r,J)}}{n^{3/2}}
\end{aligned}$$

where

$$\begin{aligned}
h_c^-(t) &= \frac{6f(t)}{q(t)}x_{r,J,\Gamma,\kappa,C}(t)4(2t) + f_{r,J,\Gamma,\kappa,C}(t)4(2t) + 6\frac{f(t)}{q(t)}x_{r,\phi,J,\Gamma,\kappa,C}(t) \\
&\le (144((2t)^9+1)(2t) + 8t + 36((2t)^9+1))f_{r,J,\Gamma,\kappa,C}(t).
\end{aligned}$$

Define $h^-(t) = \sum_{f\in E(\Gamma)\setminus E(J)} h_{a,f}^- + h_{1,f}^- + h_{2,f}^- + h_{3,f}^- + \sum_{c\in\overline{C}} h_c^- + 2f_{f,J,\Gamma,\kappa,C}(t)$ and note that

$$h^-(t) \le 1000k_{max}t + 10000k_{max}$$

and that we have shown that

$$\mathbb{E}(X^-_{\phi,J,\Gamma,\kappa,C}(i)) = \left(x^-_{\phi,J,\Gamma,\kappa,C}(t) \pm \frac{h^-(t)}{s_\sigma}\right)\frac{n^{f_2(r,J)}}{n^{3/2}}.$$

Define $h(t) = 1000k_{max}t + 10000k_{max}$ and note that $h^+(t), h^-(t) \le h(t)$. We also have that

$$2h(t) \le f'_{r,J,\Gamma,\kappa,C}(t) = (1+o(1))(24W(2k_{max}t + k_{max}))$$

as $W = 420$ completing the proof. ■

Now for the last set of random variables $Z_{\phi,J,\Gamma,\kappa,C,S}(i)$.

Lemma 7.13

For all $0 \le i \le m$ and $(\phi, J, \Gamma, \kappa, C, S) \in \mathcal{I}_z$, whenever $\mathcal{G}_i \cap \mathcal{H}_i$ holds there exists a smooth function $h_{r,J,\Gamma,\kappa,C,S}(t)$ such that

$$\mathbb{E}(Z^{\pm}_{\phi,J,\Gamma,\kappa,C,S}(i)|\mathcal{F}_i) = \left(0 \pm \frac{h_{r,J,\Gamma,\kappa,C,S}(t)}{s_\sigma}\right)\frac{n^{f_2(r,J)}}{n^{3/2}}$$

and

$$f_{r,J,\Gamma,\kappa,C,S}(t) \ge 2\int_o^t h_{\phi,J,\Gamma,\kappa,C,S}(\tau)d\tau + \beta_{r,J,\Gamma,\kappa,C,S}.$$

PROOF Note that the expected decrease in this random variable is 0. Similarly to the previous Lemma we have an increase from 3 cases. The first is the expected change when an edge is added:

$$\sum_{e\in E(J)}(1+o(1))\frac{1}{q(t)}\sum_{k\geq 0}\frac{f_{r,J_e,\Gamma,\kappa_{e,k},C,S}(t)}{s_\sigma}\frac{n^{f_2(r,J)}}{n^{3/2}}.$$

We also have that for fixed $e\in E(J)$ the expected change is

$$\begin{aligned}\frac{1}{q(t)}\sum_{k\geq 0}\frac{f_{r,J_e,\Gamma,\kappa_{e,k},C,S}(t)}{s_\sigma} &\leq \frac{1}{q(t)}\left(\sum_{k=0}^{k_{max}}\frac{2q(t)f_{r,J_e,\Gamma,\kappa_{e,k},C,S}(t)}{k!s_\sigma}+\frac{f_{r,J_e,\Gamma,\kappa,C,S\cup e}(t)}{s_\sigma}\right)\\ &\leq \frac{1}{q(t)}\left(\sum_{k=0}^{k_{max}}\frac{2q(t)f_{r,J,\Gamma,\kappa_{e,k},C,S}(t)}{k!s_\sigma}+\frac{2q(t)f_{r,J,\Gamma,\kappa,C,S}(t)}{s_\sigma}\right)\\ &\leq \frac{(2\mathrm{e})f_{r,J,\Gamma,\kappa,C,S}(t)}{s_\sigma}.\end{aligned}$$

The second is when the last copy of $(r,(K_{3,3})_r)$ rooted at f is completed. There are two cases based on the size of S. When $S=\{f\}$ then

$$\begin{aligned}&(1+o(1))\frac{1}{q(t)}2(2t)^7(2q(t))\exp(-(2t)^2)^6x_{\phi,J,\Gamma,\kappa_{f,k_{max}},C}(t)\frac{n^{f_2(r,J)}}{n^{3/2}}\\ &\leq \frac{q(t)^{e(\Gamma)-e(J)}\beta_{r,J,\Gamma,\kappa,C,S}}{k_{max}!}\frac{n^{f_2(r,J)}}{n^{3/2}}\\ &\leq \frac{f_{r,J,\Gamma,\kappa,C,S}(t)}{s_\sigma}\frac{n^{f_2(r,J)}}{n^{3/2}}.\end{aligned}$$

On the other hand if $|S|>1$ then for a fixed $f\in S$ we have that the expected change is

$$\begin{aligned}&(1+o(1))\frac{1}{q(t)}2(2t)^7(2q(t))\exp(-(2t)^2)^6\frac{f_{r,J,\Gamma,\kappa_{f,k_{max}},C,S\setminus f}}{s_\sigma}\frac{n^{f_2(r,J)}}{n^{3/2}}\\ &\leq \frac{f_{r,J,\Gamma,\kappa,C,S}(t)}{k_{max}!s_\sigma}\frac{n^{f_2(r,J)}}{n^{3/2}}\leq\frac{f_{r,J,\Gamma,\kappa,C,S}(t)}{s_\sigma^2}\frac{n^{f_2(r,J)}}{n^{3/2}}.\end{aligned}$$

Finally the expected change for completing an element $c\in C$

$$(1+o(1))\frac{1}{q(t)}2(2t)(2q(t))\frac{f_{r,J,\Gamma,\kappa,C\setminus c,S}(t)}{s_\sigma}\frac{n^{f_2(r,J)}}{n^{3/2}}(1+o(1))8t\frac{f_{r,J,\Gamma,\kappa,C,S}(t)}{s_\sigma}\frac{n^{f_2(r,J)}}{n^{3/2}}.$$

Since in Lemma 7.11 we have established that all other contributions are small we have that

$$\mathbb{E}(Z^+_{\phi,J,\Gamma,\kappa,C,S}(i))=\left(0\pm\frac{h^+(t)}{s_\sigma}\right)\frac{n^{f_2(r,J)}}{n^{3/2}}$$

where

$$h^+(t)=(e(J)2\mathrm{e}+|S|+|C|8t)f_{r,J,\Gamma,\kappa,C,S}(t)=o(f'_{r,J,\Gamma,\kappa,C,S}(t))$$

completing the proof. ∎

Corollary 7.14
For every $\sigma \in \mathcal{I}$ whenever $\mathcal{G}_i \cap \mathcal{H}_i$ hold then there exists a smooth non-negative function $h_\sigma(t)$ such that

$$\mathbb{E}(X_\sigma^\pm(i)) = \left(x_\sigma^\pm \pm \frac{h_\sigma(t)}{s_\sigma}\right)\frac{S_\sigma}{s},$$

and

$$f_\sigma(t) \geq \int_0^t h_\sigma(\tau)d\tau + \beta_\sigma.$$

7.2.5 Boundedness hypothesis

Note that the boundedness hypothesis holds whenever $\mathcal{H}_i$ and $\mathcal{G}_i$ hold as the event $\mathcal{E}_i \supseteq \mathcal{H}_i$ guarantees it. Next we verify the initial conditions.

7.2.6 Initial conditions

Lemma 7.15
For all $\sigma \in \mathcal{I}$ we have that:

$$X_\sigma(0) = \left(x_\sigma(0) \pm \frac{\beta_\sigma}{3s_\sigma}\right) S_\sigma$$

PROOF Note if $X_\sigma(0) = 0$ then $x_\sigma(0) = 0$ and we are done. In the remaining cases either $\sigma \in \mathcal{I}_q$ in which case we are considering $Q_0(0)$ and we have that

$$Q_0(0) = \binom{n}{2} = \frac{n^2}{2} + O(n) = \left(\frac{1}{2} \pm O\left(\frac{1}{n}\right)\right) n^2.$$

Note that $q_0(0) = 1/2$ and that $n^{-1} = o(1/s_\sigma) = o(\beta_0/s_\sigma)$. Otherwise we have to consider $\sigma \in \mathcal{I}_x$ where σ is a quintuple $(\phi, J, \Gamma, \kappa, C)$ where $e(J) = 0$, κ is the constant 0 function and $C = \emptyset$. In these cases we have that $X_{\phi,J,\Gamma,\kappa,C}(0) = [n]_{v(\Gamma)-2}$ and recall that J is a spanning graph of Γ and thus $f_2(r, J) = v(\Gamma) - 2$. Since

$$|n^{v(\Gamma)-2} - [n]_{v(\Gamma)-2}| = O(n^{-1}) = o\left(\frac{\beta_{r,J,\Gamma,\kappa,C}}{s_\sigma}\right)$$

hence the boundedness condition holds. ∎

7.2.7 Bounded number of variables

Lemma 7.16
For all $j \in \mathcal{V}$ and $\sigma \in \mathcal{I}_j$ we have

$$\max\{|\mathcal{V}|, |\mathcal{I}_j|\} \leq e^{u_\sigma}$$

PROOF Note that $u_\sigma = \log^2 n$ and thus $e^{u_\sigma} = n^{\log n}$. Also note that $|\mathcal{V}| = 4$ and that $|\mathcal{I}_j| \leq O(n^2 \log^{e(\Gamma)} n) \leq n^{\log n}$. ∎

7.2.8 High probability events

For the following proofs we will need the following result of Spencer [41].

Theorem 7.17 ([41])

Let (R, F) be a strictly balanced rooted graph where R forms an independent set and fix $\varepsilon, \delta > 0$. Fix $S \subseteq V(G_{n,p})$ such that $|S| = |R|$ and let $X_{S,F}$ denote the number of copies of (R, F) rooted at S in $G_{n,p}$. Then there exists a K such that if p satisfies $\mathbb{E}(X_{S,F}) > K \ln n$ then

$$\mathbb{P}(|X_{S,F} - \mathbb{E}(X_{S,F})| > \varepsilon\mathbb{E}(X_{S,F})) \leq o(n^{-\delta}).$$

Actually we will only need the following corollary.

Corollary 7.18

Let (R, F) be a strictly balanced rooted graph where R forms an independent set and consider $G_{n,p}$ with $p = n^{-1/2+o(1)}$. Define $a = \max\{0, f_2(R, F)\}$. Then for every fixed $\delta > 0$ we have with probability $1 - o(n^{-\delta})$ that the maximum number of copies of (R, F) rooted at any $S \subseteq V(G_{n,p})$ such that $|S| = |R|$ is $n^{a+o(1)}$.

PROOF Let $X_{S,F}$ denote the number of copies of (R, F) rooted at S in $G_{n,p}$ when $p = n^{-1/2+o(1)}$ and fix $\delta > 0$. Note that $\mathbb{E}(X_{S,F}) = n^{f_2(R,F)+o(1)}$.

First consider the case when $\mathbb{E}(X_{S,F}) = \omega(\log n)$. Note that this holds when $f_2(R, F) > 0$. Applying Theorem 7.17 gives us that

$$\mathbb{P}(X_{S,F} \geq 2\mathbb{E}(X_{S,F})) = o(n^{-\delta-2})$$

since $\omega(\log n) > K \log n$ for every $K > 0$ assuming n is sufficiently large. Note that $2\mathbb{E}(X_{S,F}) = n^{f_2(R,F)+o(1)}$.

Otherwise let $X'_{S,F}$ be the number of copies of (R, F) rooted at S in $G_{n,p'}$ where $p' = n^{-1/d(R,F)} \log^2 n$. We have that $p \leq p'$ as

$$\mathbb{E}(X'_{S,F}) = \Theta(n^{v(F)-|R|} p^{e(F)}) = \Theta(\log^{2e(F)} n).$$

Clearly

$$\mathbb{P}(X_{S,F} \geq 2\mathbb{E}(X'_{S,F})) \leq \mathbb{P}(X'_{S,F} \geq 2\mathbb{E}(X'_{S,F})).$$

Applying Theorem 7.17 implies that $\mathbb{P}(X_{S,F} \geq 2\mathbb{E}(X'_{S,F})) \leq o(n^{-\delta-2})$ and note that $2\mathbb{E}(X'_{S,F}) = n^{o(1)}$. The statement follows by applying the union bound. ∎

Note that this Corollary also applies to graphs as graphs are rooted graphs with empty roots.

Lemma 7.19

The event $\mathcal{B}_i$ satisfies

$$\mathbb{P}[\exists i \leq m : \mathcal{G}_i \cap \neg\mathcal{B}_i] = o(1).$$

PROOF Let (r, F) be the rooted graph composition described in Lemma 7.5 and note that is is strictly balanced with $d(F) = 2$. Instead of counting the total number of these graphs rooted at some open pair in $G_f(K_{3,4}^-)_{n,j}$ we will consider the number of copies of F found in $G_f(K_{3,4}^-)_{n,m}$. Denote by U_F the number of copies of F found in $G_f(K_{3,4}^-)_{n,m}$. Since $\mathcal{G}(m)$ implies that $Q(m) > n^{2-o(1)}$ thus

$$\begin{aligned}\mathbb{P}(\exists i \leq m : \mathcal{G}_i \cap \neg\mathcal{B}_i) &\leq \sum_{i=1}^{m} \mathbb{P}(\mathcal{G}_i \cap \neg\mathcal{B}_i) \\ &\leq \sum_{i=1}^{m} \mathbb{P}(\{Q(i) > n^{2-o(1)}\} \cap \neg\mathcal{B}_i) \\ &\leq m\mathbb{P}(\{Q(m) > n^{2-o(1)}\} \cap \neg\mathcal{B}_m) \\ &\leq m\mathbb{P}(\{Q(m) > n^{2-o(1)}\} \cap \{U_F > n^{o(1)}\}).\end{aligned}$$

Note that $U_F > n^{o(1)}$ is a decreasing property and thus we may apply Theorem 4.6 in order to determine this probability. Select $\lambda = \binom{n}{2}/Q(m) = n^{o(1)}$ and thus applying Theorem 4.6 gives us that it is enough to have a bound on the probability that $G_{n,p}$ has more than $n^{o(1)}$ copies of F, when $p = n^{-1/2+o(1)}$. Corollary 7.18 gives us that this is smaller than $n^{-\delta}$ for any $\delta > 0$. The statement follows by setting $\delta = 2$. ∎

The proofs of the remaining events $\mathcal{C}_i$, $\mathcal{D}_i$ and $\mathcal{E}_i$ will be similar, however in these cases we will have to consider rooted graphs which are not strictly balanced. In order to manage these cases, similarly to Bohman and Keevash [6] we define the *extension series* of a rooted graph. Let (R, F) be a rooted graph. Then the extension series $R = B_0 \subsetneq B_1 \subsetneq B_2 \subsetneq ... \subsetneq B_\upsilon = V(F)$ is created in the following way. If (B_i, F) is not strictly balanced, then let B_{i+1} be a minimal set which minimises $f_2(B_i, F_{B_i}[B_{i+1}])$ otherwise let $B_{i+1} = V(F)$. Bohman and Keevash [6] noted that the rooted graphs $(B_i, F_{B_i}[B_{i+1}])$ are strictly balanced and that for $i \geq 1$ we have that $f_2(B_i, F_{B_i}[B_{i+1}]) \geq 0$. Also we have that

$$f_2(R, F) = \sum_{i=0}^{\upsilon-1} f_2(B_i, F_{B_i}[B_{i+1}]).$$

The following results follow from the proofs of Bohman and Keevash [6] which were necessary for the strictly 2-balanced case. We need to extend their results slightly as we have a 2-balanced graph and thus we will include a simplified version of the proofs.

Lemma 7.20

Let (R, F) be a rooted graph where R forms an independent set and let $B_0, B_1, ..., B_\upsilon$ be its extension series. Assume $\upsilon \geq 1$. Define $a = \max\{f_2(B_0, F_{B_0}), f_2(B_1, F_{B_1})\}$. Then we have with probability $1 - o(n^{-2})$ that for any $S \subseteq V(G_{n,p})$ with $|S| = |R|$ and $p = n^{-1/2+o(1)}$ that the number of copies of (R, F) rooted at S in $G_{n,p}$ is at most $n^{a+o(1)}$.

PROOF Note that Corollary 7.18 is a special case when $B_1 = V(F)$. Note that according to Corollary 7.18 we have with probability $1 - o(n^{-2})$ that for any $0 \le i \le v - 1$ and any $S \subseteq V(G_{n,p})$ with $|S| = |B_i|$ that the number of copies of $(B_i, F_{B_i}[B_{i+1}])$ rooted at S in $G_{n,p}$ is at most $n^{\max\{f_2(B_i, F_{B_i}[B_{i+1}]),0\}+o(1)}$. Note that $f_2(B_i, F_{B_i}[B_{i+1}]) \ge 0$ for $i \ge 1$. Also since $\sum_{i=1}^{v-1} f_2(B_i, F_{B_i}[B_{i+1}]) = f_2(B_1, F_{B_1})$ thus the total number of copies of (R, F) rooted at any set S with $|S| = |R|$ in $V(G_{n,p})$ it at most

$$n^{\max\{f_2(B_0, F_{B_0}[B_1]),0\}+o(1)} n^{f_2(B_1, F_{B_1})+o(1)} = n^{a+o(1)}$$

as $f_2(B_0, F_{B_0}[B_1]) + f_2(B_1, F_{B_1}) = f_2(B_0, F_{B_0})$. ■

Lemma 7.21
Let (R, F) be a rooted graph such that R forms an independent set and for every $R \subseteq B \subsetneq V(F)$ we have that $f_2(B, F) \le a$ for some $a \ge 0$. Then we have with probability $1 - o(n^{-2})$ that for any $S \subseteq V(G_{n,p})$ with $|S| = |R|$ and $p = n^{-1/2+o(1)}$ that the number of copies of (R, F) rooted at S in $G_{n,p}$ is at most $n^{a+o(1)}$.

PROOF Let $B_0, B_1, ..., B_v$ be the extension series of (R, F). We have to consider 3 cases when $B_0 = V(F)$ when $B_1 = V(F)$ and when neither of these hold. The case $B_0 = V(F)$ is trivial as there can be only $O(1)$ copies of F rooted at S which is less than $n^{o(1)}$. In case $B_1 = V(F)$ then we can apply Lemma 7.20 and since $f_2(B_1, F_{B_1}) = 0$ there are at most $n^{\max\{f_2(B_0, F_{B_0}),0\}+o(1)} \le n^{a+o(1)}$ copies. The final case also follows from applying Lemma 7.20 as both $f_2(B_0, F_{B_0}) \le a$ and $f_2(B_1, F_{B_1}) \le a$. ■

Lemma 7.22
Let (R, F) be a rooted graph such that R forms an independent set and for every $R \subseteq B \subsetneq V(F)$ we have that $f_2(B, F) \le a$ for some $a \ge 0$. Let $\mathcal{U}_i$ denote the event that for every $j \le i$ we have that for every $S \subseteq V(G_f(K_{3,4}^-)_{n,i})$ the number of copies of (R, F) rooted at S in $G_f(K_{3,4}^-)_{n,i}$ is at most $n^{a+o(1)}$. Then we have that

$$\mathbb{P}(\exists i \le m : \mathcal{G}_i \cap \neg \mathcal{U}_i) = o(1).$$

PROOF Note that $\mathcal{G}_m$ implies that $Q(m) > n^{2-o(1)}$ thus

$$\begin{aligned}
\mathbb{P}(\exists i \le m : \mathcal{G}_i \cap \neg \mathcal{U}_i) &\le \sum_{i=1}^{m} \mathbb{P}(\mathcal{G}_i \cap \neg \mathcal{U}_i) \\
&\le \sum_{i=1}^{m} \mathbb{P}(\{Q(i) > n^{2-o(1)}\} \cap \neg \mathcal{U}_i) \\
&\le m\mathbb{P}(\{Q(m) > n^{2-o(1)}\} \cap \neg \mathcal{U}_m).
\end{aligned}$$

Note that $\mathcal{U}_m$ is a decreasing property and thus we may apply Theorem 4.6 in order to determine this probability. Let $p = n^{-1/2+o(1)}$ and let $\mathcal{U}_i$ denote the event that for every $S \subseteq V(G_{n,p})$ the number of copies of (R, F) rooted at S in $G_{n,p}$ is at most $n^{a+o(1)}$.

Select $\lambda = \binom{n}{2}/Q(m) = n^{o(1)}$ and thus applying Theorem 4.6 gives us that

$$\mathbb{P}(\{Q(m) > n^{2-o(1)}\} \cap \neg\mathcal{U}_m) \leq \mathbb{P}(\neg\mathcal{U}) + o(n^{-2}).$$

Lemma 7.21 gives us that $\mathbb{P}(\neg\mathcal{U}) = o(n^{-2})$ and the statement follows. ■

Lemma 7.23

Let (r_1, F_1) and (r_2, F_2) be strictly balanced non-trivial rooted graphs, such that r_1 and r_2 are non-edges and $d(r_1, F_1) = d(r_2, F_2) = 2$. Let $(g_1, g_2, F) \in (r_1, F_1) \oplus (r_2, F_2)$ and let $S_{g_1,g_2,F} \subseteq B \subsetneq V(F)$. Then we have that

$$f_2(B, F_B) < 0.$$

PROOF Define B_1 as the maximal subset of $V(F_1)$ such that $g_1(B_1) \subseteq B$ and similarly let B_2 be the maximal subset of $V(F_2)$ such that $g_2(B_2) \subseteq B$. Also define $T_1 \subseteq V(F_1)$, $T_2 \subseteq V(F_2)$ as the maximal subsets such that $g_1(T_1) \subseteq V(T_{g_1,g_2,F})$ and $g_2(T_2) \subseteq V(T_{g_1,g_2,F})$. Let $e(B_1 \backslash T_1, T_1)$ denote the number of edges between $B_1 \backslash T_1$ and T_1 and $e(B_2 \backslash T_2, T_2)$ is defined analogously

Note that

$$f_2(B, F_B) \leq f_2(B_1, (F_1)_{B_1}) + f_2(B_2 \cup T_2, (F_2)_{B_2 \cup T_2}) - e(B_2 \backslash T_2, T_2)/2$$
$$f_2(B, F_B) \leq f_2(B_1 \cup T_1, (F_1)_{B_1 \cup T_1}) + f_2(B_2, (F_2)_{B_2}) - e(B_1 \backslash T_1, T_1)/2$$

as the first inequality might ignore some edges in $E(F_2[B_2 \cup T_2])$ and the second some edges in $E(F_1[B_1 \cup T_1])$.

Note that since (R_1, F_1) and (R_2, F_2) are strictly balanced rooted graphs we have that for i=1,2 and every $R_i \subseteq S \subseteq V(F_i)$ that $f_2(S, (F_i)_S) \leq 0$ where equality holds only if $S = R_i$ or $S = V(F_i)$. It is enough to consider the case when each of $f_2(B_1, (F_1)_{B_1})$,$f_2(B_2 \cup T_2, (F_2)_{B_2 \cup T_2})$, $f_2(B_1 \cup T_1, (F_1)_{B_1 \cup T_1})$, $f_2(B_2, (F_2)_{B_2})$, $e(B_1 \backslash T_1, T_1)$, and $e(B_2 \backslash T_2, T_2)$ is zero.

Note that we have that $B \subsetneq V(F)$ thus either $B_1 \subsetneq V(F_1)$ or $B_2 \subsetneq V(F_2)$ and without loss of generality we may assume that $B_1 \subsetneq V(F_1)$. Since $f_2(B_1, (F_1)_{B_1}) = 0$ we have that $B_1 = r_1$. Since $f_2(B_2, (F_2)_{B_2}) = 0$ we have that either $B_2 = r_2$ or $B_2 = V(F_2)$. However the second contradicts the assumption that $e(T_{g_1,g_2,F}) > 0$ as in this case no edge in $g_1(E(F_1))$ is spanned by B and every edge in $g_2(E(F_2))$ is spanned by B.

Thus we have that $B_1 = r_1$ and $B_2 = r_2$. Note that neither B_1 or B_2 spans any edges and thus we have that $T_1 \not\subseteq B_1$ and $T_2 \not\subseteq B_2$. We also have that $B_1 \cup T_1 = V(F_1)$ and $B_2 \cup T_2 = V(F_2)$. Since $e(B_1 \backslash T_1, T_1) = 0$ and because (r_1, F_1) is non-trivial we have that $T_1 = V(F_1)$ and similarly $T_2 = V(F_2)$. Therefore the two copies overlap on every vertex and $g_1(r_1) = g_2(r_2) = B$. However this means that the graph we are considering is a copy of (r_1, F_1) with an additional edge added, and the statement holds. ∎

Lemma 7.24

The event $\mathcal{C}_i$ satisfies

$$\mathbb{P}[\exists i \leq m : \mathcal{G}_i \cap \neg\mathcal{C}_i] = o(1).$$

PROOF Recall that $\mathcal{C}_i$ was the event that for every $j \leq i$ we have that

- for any pair of vertices $u, v \in V(G_f(K_{3,4}^-)_{n,j})$ with $u \neq v$ and for any rooted graph $(r, F) \in \{(r_1, (K_{3,4}^-)_{r_1}), (r, (K_3)_r), (r, (K_{3,3})_r)\}$ we have that the number of edges which when inserted into $G_f(K_{3,4}^-)_{n,j}$ would create multiple copies of (r, F) rooted at u, v is at most $f(t)\sqrt{n}/n^{1/11}$

- for any four vertices $u_1, u_2, v_1, v_2 \in V(G_f(K_{3,4}^-)_{n,j})$, with $u_1 \neq v_1$, $u_2 \neq v_2$ and for $(r_1, F_1), (r_2, F_2) \in \{(r_1, (K_{3,4}^-)_{r_1}), (r, (K_3)_r), (r, (K_{3,3})_r)\}$ the number of edges which when inserted into $G_f(K_{3,4}^-)_{n,j}$ would create both a copy of (r_1, F_1) at $\{u_1, v_1\}$ and (r_2, F_2) at u_2, v_2 is at most $f(t)\sqrt{n}/n^{1/11}$.

Note that inserting this last edge would not only create a copy of (r_1, F_1) rooted at f_1 and a copy of (r_2, F_2) at f_2, but also it would create a copy of a rooted graph $(R_{g_1,g_2,F}, F)$ for some $(g_1, g_2, F) \in (r_1, F_1) \oplus (r_2, F_2)$ rooted at $f_1 \cup f_2$. A similar argument holds for creating multiple copies of (r_1, F_1) at f_1.

We will show that the maximal number of copies of $(R_{g_1,g_2,F}, F_e)$ for any $e \in E(F)$ rooted at any set $S \subseteq V(G_f(K_{3,4}^-))$ such that $|S| = |R_{g_1,g_2,F}|$ is $n^{o(1)}$.

Note that $f_2(B, (F_e)_B) \leq f_2(B, F_B) + 1/2$. We also have by Lemma 7.23 that $f_2(B, F_B) < 0$ for $R_{g_1,g_2,F} \subseteq B \subsetneq V(F)$ and note that $f_2(B, F_B) < 0$ is equivalent to $f_2(B, F_B) \leq -1/2$ and the result follows by applying Lemma 7.22. ∎

Lemma 7.25

Let $(r_1, F_1), (r_2, F_2)$ be strictly balanced non-trivial rooted graphs such that r_1 and r_2 are non-edges and let $(g_1, g_2, F) \in F_1 \otimes F_2$ such that the following hold

- $g_1(E(F_1)) \cap g_2(E(F_2)) \neq \emptyset$

- $g_2(r_2) \subseteq g_1(V(F_1))$

- $g_2(E(F_2)) \not\subseteq g_1(E(F_1))$

Then for any $g_1(r_1) \subseteq B \subsetneq V(F)$ *we have that* $f_2(B, F_B) < 0$.

PROOF Define B_1 as the maximal subset of $V(F_1)$ such that $g_1(B_1) \subseteq B$ and similarly let B_2 be the maximal subset of $V(F_2)$ such that $g_2(B_2) \subseteq B$. Also define $T_2 \subseteq V(F_2)$ as the maximal subset such that $g_2(T_2) \subseteq g_1(V(F_1)) \cap g_2(V(F_2))$. Let $e(B_2 \backslash T_2, T_2)$ denote the number of edges between $B_2 \backslash T_2$ and T_2. Note that

$$f_2(B, F_B) \leq f_2(B_1, (F_1)_{B_1}) + f_2(B_2 \cup T_2, (F_2)_{B_2 \cup T_2}) - e(B_2 \backslash T_2, T_2)/2.$$

Note that when $r_2 \subseteq B_2$ then this is equivalent to Lemma 7.23. So we will assume that $B_2 \cap r_2 \subsetneq r_2$. Again it is enough to concentrate on the case when each of $f_2(B_1, (F_1)_{B_1})$, $f_2(B_2 \cup T_2, (F_2)_{B_2 \cup T_2})$, $e(B_2 \backslash T_2, T_2)$ is zero. We have that $f_2(B_1, (F_1)_{B_1}) = 0$ only if either $r_1 = B_1$ or $V(F_1) = B_1$. However this second case is ruled out as $g_2(r_2) \subseteq g_1(V(F_1))$ and $g_2(r_2) \not\subseteq B$. Clearly $r_2 \subseteq T_2$ and since it spans no edges we have that $r_2 \subsetneq T_2$. Therefore since $f_2(B_2 \cup T_2, (F_2)_{B_2 \cup T_2}) = 0$ we have that $B_2 \cup T_2 = V(F_2)$. Note that if (r_2, F_2) is strictly balanced it is connected. Therefore since $e(B_2 \backslash T_2, T_2) = 0$ and $V(F_2) = B_2 \cup T_2$ we have that $B_2 \backslash T_2 = \emptyset$. Thus we have that $V(F_2) = T_2$ thus $g_2(V(F_2)) \subseteq g_1(V_{F_1})$. However this implies that $(g_1(r_1), F)$ is just a copy of (r_1, F_1) with some additional edges added. The statement follows. ■

Lemma 7.26

The event $\mathcal{D}_i$ *satisfies*

$$\mathbb{P}[\exists i \leq m : \mathcal{G}_i \cap \neg \mathcal{D}_i] = o(1).$$

PROOF Recall that the event $\mathcal{D}_i$ holds if for every $j \leq i$ we have the following. Select $(r, \Gamma) \in \{(r, (K_3)_r), (r, (K_{3,3})_r), (r_1, (K_{3,4}^-)_{r_1})\}$ and let (r, J) be a proper spanning subgraph of (r, Γ). Also let $(r_2, F_2) \in \{(r, (K_3)_r), (r, (K_{3,3})_r), (r_1, (K_{3,4}^-)_{r_1})\}$. Then for any $(g_1, g_2, F) \in (J \otimes F_2)$ such that $g_2(r_2) \subseteq g_1(E(J))$ and $g_1(E(\Gamma) \backslash E(J)) \cap g_2(E(F_2)) = \emptyset$ we have that for any pair of distinct vertices $u, v \in V(G_f(K_{3,4}^-)_{n,j})$ the number of edges which when inserted into the graph would create a copy of $(g_1(r), F)$ rooted at $\{u, v\}$ in $G_f(K_{3,4}^-)_{n,j}$ is $O(n^{f_2(r,J)} n^{\varepsilon})$.

Fix $(r_2, F_2) \in \{(r, (K_3)_r), (r, (K_{3,3})_r), (r_1, (K_{3,4}^-)_{r_1})\}$ and we will show that there are at most $n^{f_2(r,J)+o(1)}$ such edges. We are looking for the number of copies of $(g_1(r), F_e)$ rooted at $\{u, v\}$ in $G_f(K_{3,4}^-)_{n,j}$ for some fixed $e \in E(T_{g_1, g_2, F})$. Lemma 7.22 implies the result if we have that $f_2(B, F_B) \leq f_2(r, J)$ for every $g_1(r) \subseteq B \subsetneq V(F)$.

Note that for the same g_1, g_2 we have that $(g_1, g_2, F') \in \Gamma \otimes F_2$ such that all the conditions of Lemma 7.25 apply as in case the last condition did not hold there would be an edge in $E(\Gamma)\backslash E(J)$, a case which is not considered. Note that any $g_1(r) \subseteq B \subsetneq V(F_e)$ that $f_2(B,(F_e)_B) \leq f_2(B,(F')_B) + (e(\Gamma) - e(J) + 1)/2 = f_2(B,(F')_B) + f_2(r, J) + 1/2$. Lemma 7.25 gives us that $f_2(B,(F')_B) \leq -1/2$. ■

Lemma 7.27

Let $(r_1, F_1), (r_2, F_2)$ be strictly balanced non-trivial rooted graphs such that r_1 and r_2 are non-edges and let $(g_1, g_2, F) \in F_1 \otimes F_2$ such that the following hold

- $g_2(r_2) \subseteq g_1(V(F_1))$
- $g_2(E(F_2)) \nsubseteq g_1(E(F_1))$

Then for any $g_1(r_1) \subsetneq B \subsetneq V(F)$ we have that $f_2(B, F_B) \leq 0$ and equality holds only if $B = V(F_1)$ and $g_1(V(F_1)) \cap g_2(V(F_2)\backslash r_2) = \emptyset$.

PROOF Define B_1 as the maximal subset of $V(F_1)$ such that $g_1(B_1) \subseteq B$ and similarly let B_2 be the maximal subset of $V(F_2)$ such that $g_2(B_2) \subseteq B$. Lemma 7.25 gives us that the statement holds when $g_1(E(V(F_1))) \cap g_2(E(V(F_2))) \neq \emptyset$. Otherwise let $T = g_1(V_1) \cap g_2(V_2\backslash r_2)$. We have that

$$f_2(B, F_B) \leq f_2(B_1,(F_1)_{B_1}) + f_2(r_2 \cup B_2,(F_2)_{B_2}) - |T\backslash B|.$$

Next we examine the cases when all of these terms are zero. As always we have that $f_2(B_1,(F_1)_{B_1}) = 0$ if $B_1 = r_1$ and when $B_1 = V(F_1)$. In case $B_1 = r_1$ then $r_2 \cap B_2 = \emptyset$ and $B_2 \neq \emptyset$, thus $f_2(r_2 \cup B_2,(F_2)_{B_2}) < 0$. On the other hand if $B_1 = V(F_1)$ then $B_2 \neq V(F_2)$ so we only have to consider the case when $B_2 = r_2$, in which case $T\backslash B = \emptyset$ completing the proof. ■

Lemma 7.28

Let (r, Γ) be a strictly balanced rooted graph with density 2 such that r is a non-edge. Let $(g_1, g_2, g_3, F) \in \Gamma \otimes (K_{3,3})_r \otimes (K_3)_{r,e}$ such that $g_2(r) \subseteq g_1(V(J))$, $g_2(E(F_2)) \nsubseteq g_1(V(F_1))$, and $g_3((K_3)_{r,e}) \cap g_2(V((K_{3,3})_r)) = g_3(r)$. Then we have that for any $g_1(r) \cup g_3(e) \subseteq B \subsetneq V(F)$ we have that

$$f_2(B, F_B) < 0.$$

PROOF Let $V' = g_1(V(\Gamma)) \cup g_2(V((K_{3,3})_r))$. Note that

$$f_2(B, F_B) \leq f_2(B \cap V', F_B[V']).$$

Lemma 7.27 gives us that $f_2(B \cap V', F_B[V']) = 0$ only if $B \cap V' = g_1(V(\Gamma))$ and $g_1(V(\Gamma)) \cap g_2(V((K_{3,3})_r)) = g_2(r)$. Let $V'' = g_2(V((K_{3,3})_r)) \cup g_3(V((K_3)_{r,e}))$. For the remainder of the proof assume that $B \cap V' = g_1(V(\Gamma))$ and $g_1(V(\Gamma)) \cap g_2(V((K_{3,3})_r)) = g_2(r)$. Then we have that

$$f_2(B, F_B) = f_2(B \cap V'', F_B[V'']).$$

Note that $g_2(r) \subsetneq B \cap V''$. Therefore Lemma 7.27 implies that $f_2(B \cap V'', F_B[V'']) < 0$ as $B \cap V'' \neq g_2(V((K_{3,3})_r))$ completing the proof. ∎

Lemma 7.29

The event $\mathcal{E}_i$ satisfies

$$\mathbb{P}[\exists i \leq m : \mathcal{G}_i \cap \neg \mathcal{E}_i] = o(1).$$

PROOF Recall that $\mathcal{E}_i$ is the event that the boundedness hypothesis holds i.e.

$$X_\sigma^\pm \leq \frac{\beta_\sigma^2}{s_\sigma^2 \lambda_\sigma \tau_\sigma} \frac{S_\sigma}{u_\sigma} = \frac{\beta_\sigma n^\varepsilon S_\sigma}{n^{1/11} \log^3 n}.$$

We will show that these bounds hold on a case by case basis. In every case we will need an upper bound on the number of some rooted graphs (R, F) rooted at any $S \subseteq V(G_f(H)_{n,j})$ such that $|S| = |R|$. Similarly to the previous cases our aim is to determine an upper bound for $f_2(B, F_B)$ and then apply Lemma 7.22 to determine the upper bound.

We start by analysing the case when our random variable is $Q_k(i)$ and $Y_o(i)$. Assume that the edge inserted in step i is f. Then the change in these variables is determined by the number of copies of $(r, (K_{3,3})_{r,e})$ rooted at f, the number of copies of $(r_1, (K_{3,4}^-)_{r_1,e})$ rooted at f and the number of copies of $(r, (K_3)_{r,e})$ rooted at some non-edge of some copy of $(r, (K_{3,3})_r)$ rooted at f. Note that for any strictly balanced rooted graph (r, Γ) such that $d(r, \Gamma) = 2$ we have that for every $r \subseteq B \subsetneq V(\Gamma)$ that $f_2(B, \Gamma_B) \leq 0$ and thus after removing an edge from it we have that $f_2(B, (\Gamma_e)_B) \leq 1/2$. Since there are at most $n^{o(1)}$ copies of $(r, (K_{3,3})_r)$ rooted at any pair of vertices Lemma 7.22 implies that the maximal change is at most $n^{1/2+o(1)}$, which satisfies the conditions.

Now for $X_{\phi,J,\Gamma,\kappa,C}(i)$ and $Z_{\phi,J,\Gamma,\kappa,C,S}(i)$. In these cases we aim to show that the maximal change can be bounded by $S_\sigma n^{o(1)}/\sqrt{n}$ which implies the result. Let (r, J) be a subgraph of a strictly balanced rooted graph (r, Γ). The first case considered is when inserting an edge creates the last edge needed in (r, J). We bound the number of such copies created by inserting an edge, by giving bounds on the number of rooted graphs $(r \cup e, J_e)$ rooted at any set $S \subseteq V(G_{n,p})$ with $|S| = |r \cup e|$ and

$p = n^{-1/2+o(1)}$. We have that for any $r \cup e \subseteq B \subsetneq V(J)$ that $f_2(B, J_e) \leq f_2(B, \Gamma_B) + (e(\Gamma)-e(J))/2 = f_2(B, \Gamma_B)+f_2(r, J)$ as $e \subseteq B$. Also since (r, Γ) is a strictly balanced rooted graph we have that $f_2(B, \Gamma_B) \leq -1/2$ for any $r_1 \cup e \subseteq B \subsetneq V(\Gamma)$ and the statement follows from Lemma 7.22.

The second case we consider is when a mutual neighbour of a non-edge of a copy of $(r, (K_{3,3})_r)$ rooted at a specific non-edge of a copy of (r, J) rooted with respect to ϕ is completed. In order to estimate the maximal change caused by completing such a construction we aim to bound the number of copies of the rooted graphs $(g_1(r) \cup g_3(e), F)$ rooted at any set $S \subseteq G_{n,p}$ such that $|S| = |g_1(r) \cup g_3(e)|$ where $(g_1, g_2, g_3, F) \in J \otimes (K_{3,3})_r \otimes (K_3)_{r,e}$ such that $g_2(r) \subseteq g_1(V(J))$, $g_2(E((K_{3,3})_r)) \not\subseteq g_1(E(J))$ and $g_3(V((K_3)_{r,e})) \cap g_2(V((K_{3,3})_r)) = g_3(r)$. This follows in the usual way from Lemma 7.28 and Lemma 7.22.

In the remaining cases let $(r, \Gamma) \in \{(r, (K_{3,3})_r), (r, (K_3)_r), (r_1, (K^-_{3,4})_{r_1})\}$ and consider $(g_1, g_2, F) \in J \otimes \Gamma_e$ such that $g_2(r) \subseteq g_1(V(J))$ and $g_2(E(\Gamma_e)) \not\subseteq g_1(E(J))$. Note that the maximal number of copies of $(g_1(r) \cup g_2(e), F)$ rooted at any set of vertices $S \subseteq V(G_{n,p})$ with $|S| = \{g_1(r) \cup g_2(e)\}$ is an upper bound on the maximal change in these cases. Note that $B \neq g_1(V(J))$, as $g_2(e) \subseteq B$ and $g_2(e) \not\subseteq V(J)$. Therefore the result follows from Lemma 7.27 and Lemma 7.22. ■

Corollary 7.30

The event $\mathcal{H}_i$ satisfies

$$\mathbb{P}[\exists i \leq m : \mathcal{G}_i \cap \neg\mathcal{H}_i] = o(1).$$

PROOF Follows trivially from the fact that $\mathcal{H}_i = \mathcal{B}_i \cap \mathcal{C}_i \cap \mathcal{D}_i \cap \mathcal{E}_i$ and that

$$\mathbb{P}[\exists i \leq m : \mathcal{G}_i \cap \neg\mathcal{H}_i] \leq \sum_{\mathcal{S} \in \{\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D},\mathcal{E}\}} \mathbb{P}[\exists i \leq m : \mathcal{G}_i \cap \neg\mathcal{S}_i] = o(1).$$

■

7.2.9 Additional technical assumptions

Lemma 7.31

The additional technical assumptions hold.

PROOF Note that $u_\sigma = \omega(1)$. Also note that by definition $\lambda_\sigma/\beta_\sigma = n^\varepsilon$ therefore

$$s = n^{3/2} \geq= 15 \log^2 n \log n (n^{1/22-\varepsilon} n^\varepsilon)^2 = 15 u_\sigma \tau_\sigma (s_\sigma \lambda_\sigma/\beta_\sigma)^2.$$

also

$$s = n^{3/2} \geq 9n^{1/22-\varepsilon} n^\varepsilon = 9 s_\sigma \lambda_\sigma/\beta_\sigma.$$

We also have that

$$\frac{s\beta_\sigma}{18 s_\sigma \lambda_\sigma} = \frac{n^{3/2}}{18 n^{1/22}} < m = \mu n^{3/2} \sqrt{\log \log n} \leq \frac{n^{3/2} \log n}{1944} = \frac{s\tau_\sigma}{1944}$$

Also

$$\sup_{0 \leq t \leq m/s} q_k^{\pm}(t) = O(k_{max}(t) q_k(t)) = O(k_{max} t \beta_k) = o(\beta_k n^{\varepsilon})$$

$$\sup_{0 \leq t \leq m/s} x_\sigma^{\pm}(t) = O(k_{max} t \beta_\sigma) = \beta_\sigma n^{\varepsilon}.$$

We also have that

$$\int_0^{m/s} |x''_\sigma(t)| dt \leq \lambda_\sigma.$$

Clearly we may set $h_\sigma(t) = f'_\sigma(t)/2$ and thus

$$h_\sigma(0) = O(k_{max})\beta_\sigma \leq n^{1/22} \beta_\sigma = s_\sigma \lambda_\sigma.$$

Finally:

$$\int_0^{m/s} |h'_\sigma(t)| dt \leq h_\sigma(m/s) \leq f'_\sigma(m/s) \leq n^{1/22} \beta_\sigma = s_\sigma \lambda_\sigma.$$

■

Bibliography

[1] ALDOUS, D. A random tree model associated with random graphs. *Random Structures Algorithms 1*, 4 (1990), 383–402. Available from: http://dx.doi.org/10.1002/rsa.3240010402. 21

[2] ALLEN, P. Personal communication. 25

[3] ALON, N., RÓNYAI, L., AND SZABÓ, T. Norm-graphs: variations and applications. *J. Combin. Theory Ser. B 76*, 2 (1999), 280–290. Available from: http://dx.doi.org/10.1006/jctb.1999.1906. 38

[4] BOHMAN, T. The triangle-free process. *Adv. Math. 221*, 5 (2009), 1653–1677. Available from: http://dx.doi.org/10.1016/j.aim.2009.02.018. 8, 71, 72, 75

[5] BOHMAN, T., FRIEZE, A., AND LUBETZKY, E. Random triangle removal. *Preprint* (2012). arXiv:1203.4223. 22

[6] BOHMAN, T., AND KEEVASH, P. The early evolution of the H-free process. *Invent. Math. 181*, 2 (2010), 291–336. Available from: http://dx.doi.org/10.1007/s00222-010-0247-x. 9, 12, 26, 30, 32, 37, 38, 71, 72, 86, 112

[7] BOLLOBÁS, B. Threshold functions for small subgraphs. *Math. Proc. Cambridge Philos. Soc. 90*, 2 (1981), 197–206. Available from: http://dx.doi.org/10.1017/S0305004100058655. 25, 69

[8] BOLLOBÁS, B. *Random graphs*, second ed., vol. 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2001. 19

[9] BOLLOBÁS, B., AND RIORDAN, O. Constrained graph processes. *Electron. J. Combin. 7* (2000), Research Paper 18, 20 pp. (electronic). Available from: http://www.combinatorics.org/Volume_7/Abstracts/v7i1r18.html. 8, 23

[10] BOLLOBÁS, B., AND RIORDAN, O. Random graphs and branching processes. In *Handbook of large-scale random networks*, vol. 18 of *Bolyai Soc. Math. Stud.* Springer, Berlin, 2009, pp. 15–115. 7, 21

[11] BROWN, W. G. On graphs that do not contain a Thomsen graph. *Canad. Math. Bull. 9* (1966), 281–285. 38

[12] ERDŐS, P. Graph theory and probability. II. *Canad. J. Math. 13* (1961), 346–352. 64

[13] ERDŐS, P., AND RÉNYI, A. On random graphs. I. *Publ. Math. Debrecen 6* (1959), 290–297. 18

[14] ERDŐS, P., AND RÉNYI, A. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl. 5* (1960), 17–61. 7, 70

[15] ERDŐS, P., RÉNYI, A., AND SÓS, V. T. On a problem of graph theory. *Studia Sci. Math. Hungar. 1* (1966), 215–235. 38

[16] ERDÖS, P., AND STONE, A. H. On the structure of linear graphs. *Bull. Amer. Math. Soc. 52* (1946), 1087–1091. 38

[17] ERDŐS, P., SUEN, S., AND WINKLER, P. On the size of a random maximal graph. In *Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, "Random Graphs '93" (Poznań, 1993)* (1995), vol. 6, pp. 309–318. 7, 20, 21, 23, 44

[18] ERDŐS, P., AND TETALI, P. Representations of integers as the sum of k terms. *Random Structures Algorithms 1*, 3 (1990), 245–261. Available from: http://dx.doi.org/10.1002/rsa.3240010302. 17

[19] FÜREDI, Z. An upper bound on Zarankiewicz' problem. *Combin. Probab. Comput. 5*, 1 (1996), 29–33. Available from: http://dx.doi.org/10.1017/S0963548300001814. 38

[20] GERKE, S., AND MAKAI, T. No dense subgraphs appear in the triangle-free graph process. *Electron. J. Combin. 18*, 1 (2011), Research Paper 168, 7. 71

[21] GERKE, S., SCHLATTER, D., STEGER, A., AND TARAZ, A. The random planar graph process. *Random Structures Algorithms 32*, 2 (2008), 236–261. 21

[22] GRABLE, D. A. On random greedy triangle packing. *Electron. J. Combin. 4*, 1 (1997), Research Paper 11, 19 pp. (electronic). Available from: http://www.combinatorics.org/Volume_4/Abstracts/v4i1r11.html. 22

[23] HARRIS, T. E. A lower bound for the critical probability in a certain percolation process. *Proc. Cambridge Philos. Soc. 56* (1960), 13–20. 15

[24] JANSON, S. New versions of Suen's correlation inequality. In *Proceedings of the Eighth International Conference "Random Structures and Algorithms" (Poznan, 1997)* (1998), vol. 13, pp. 467–483. Available from: http://dx.doi.org/10.1002/(SICI)1098-2418(199810/12)13:3/4<467::AID-RSA15>3.3.CO;2-E. 16

[25] JANSON, S., ŁUCZAK, T., AND RUCIŃSKI, A. An exponential bound for the probability of nonexistence of a specified subgraph in a random graph. In *Random graphs '87 (Poznań, 1987)*. Wiley, Chichester, 1990, pp. 73–87. 16

[26] JANSON, S., ŁUCZAK, T., AND RUCIŃSKI, A. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000. 68

[27] KIM, J. H. The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$. *Random Structures Algorithms 7*, 3 (1995), 173–207. Available from: http://dx.doi.org/10.1002/rsa.3240070302. 8

[28] KÖVARI, T., SÓS, V. T., AND TURÁN, P. On a problem of K. Zarankiewicz. *Colloquium Math. 3* (1954), 50–57. 38

[29] ŁUCZAK, T. On the equivalence of two basic models of random graphs. In *Random graphs '87 (Poznań, 1987)*. Wiley, Chichester, 1990, pp. 151–157. 19

[30] OSTHUS, D. Personal communication. 25

[31] OSTHUS, D., AND TARAZ, A. Random maximal H-free graphs. *Random Structures Algorithms 18*, 1 (2001), 61–82. Available from: http://dx.doi.org/10.1002/1098-2418(200101)18:1<61::AID-RSA5>3.3.CO;2-K. 8, 9, 12, 23, 24, 64, 77

[32] PICOLLELLI, M. The diamond-free process. *Preprint* (2010). arXiv:1010.5207. 9, 33, 36, 38

[33] PICOLLELLI, M. The final size of the C_4-free process. *Combinatorics, Probability and Computing 1*, 20 (2011), 939–955. 9

[34] PICOLLELLI, M. The final size of the C_ℓ-free process. *Preprint* (2011). 9, 38

[35] PÓSA, L. Hamiltonian circuits in random graphs. *Discrete Math. 14*, 4 (1976), 359–364. 25

[36] RÖDL, V., AND THOMA, L. Asymptotic packing and the random greedy algorithm. *Random Structures Algorithms 8*, 3 (1996), 161–177. Available from: http://dx.doi.org/10.1002/(SICI)1098-2418(199605)8:3<161::AID-RSA1>3.0.CO;2-W. 22

[37] RUCIŃSKI, A., AND VINCE, A. Balanced graphs and the problem of subgraphs of random graphs. In *Proceedings of the sixteenth Southeastern international conference on combinatorics, graph theory and computing (Boca Raton, Fla., 1985)* (1985), vol. 49, pp. 181–190. 69

[38] RUCIŃSKI, A., AND WORMALD, N. C. Random graph processes with degree restrictions. *Combinatorics, Probability and Computing 1*, 2 (1992), 169–180. Available from: http://dx.doi.org/10.1017/S0963548300000183. 8, 28, 30

[39] RUCIŃSKI, A., AND WORMALD, N. C. Connectedness of graphs generated by a random d-process. *J. Aust. Math. Soc. 72*, 1 (2002), 67–85. Available from: http://dx.doi.org/10.1017/S1446788700003591. 30

[40] SHEARER, J. B. A note on the independence number of triangle-free graphs. *Discrete Math. 46*, 1 (1983), 83–87. Available from: http://dx.doi.org/10.1016/0012-365X(83)90273-X. 66

[41] SPENCER, J. Counting extensions. *J. Combin. Theory Ser. A 55*, 2 (1990), 247–255. Available from: http://dx.doi.org/10.1016/0097-3165(90)90070-D. 111

[42] SPENCER, J. Threshold functions for extension statements. *J. Combin. Theory Ser. A 53*, 2 (1990), 286–305. Available from: http://dx.doi.org/10.1016/0097-3165(90)90061-Z. 39

[43] SPENCER, J. Asymptotic packing via a branching process. *Random Structures Algorithms 7*, 2 (1995), 167–172. Available from: http://dx.doi.org/10.1002/rsa.3240070206. 22

[44] SUDAKOV, B., AND VU, V. H. Local resilience of graphs. *Random Structures Algorithms 33*, 4 (2008), 409–433. Available from: http://dx.doi.org/10.1002/rsa.20235. 25

[45] TELCS, A., WORMALD, N., AND ZHOU, S. Hamiltonicity of random graphs produced by 2-processes. *Random Structures Algorithms 31*, 4 (2007), 450–481. Available from: http://dx.doi.org/10.1002/rsa.20133. 30

[46] WARNKE, L. When does the K_4-free process stop? *Preprint* (2010). arXiv:1007.3037. 9, 26, 38

[47] WARNKE, L. The C_ℓ-free process. *Preprint* (2011). arXiv:1101.0693. 9, 24, 38

[48] WARNKE, L. Dense subgraphs in the H-free process. *Discrete Mathematics 311*, 2324 (2011), 2703 – 2707. Available from: http://www.sciencedirect.com/science/article/pii/S0012365X11003621. 75, 76

[49] WOLFOVITZ, G. Lower bounds for the size of random maximal H-free graphs. *Electron. J. Combin. 16*, 1 (2009), Research Paper 4, 26. Available from: http://www.combinatorics.org/Volume_16/Abstracts/v16i1r4.html. 24

[50] WOLFOVITZ, G. Triangle-free subgraphs in the triangle-free process. *Preprint* (2009). arXiv:0903.1756. 9, 71, 72

[51] WOLFOVITZ, G. The K_4-free process. *Preprint* (2010). arXiv:1008.4044. 9, 38

[52] WORMALD, N. The differential equation method for random graph processes and greedy algorithms. *M. Karonski, H.J. Prmel (Eds.), Lectures on Approximation and Randomized Algorithms, PWN, Warsaw* (1999), 73–155. Available from: http://www.math.uwaterloo.ca/~nwormald/papers/de.pdf. 26, 28, 29, 30