Some problems related to the Karp-Sipser algorithm on random graphs

Eleonora Kreačić

Department of Statistics

University of Oxford

This dissertation is submitted for the degree of

Doctor of Philosophy

St Catherine's College

Trinity, 2017

To my mother, Jelena.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original. This thesis is submitted to the University of Oxford for the degree Doctor of Philosophy, and has not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university.

Eleonora Kreačić

Trinity, 2017

Acknowledgements

Most of all, I would like to thank my supervisors Professor Christina Goldschmidt and Professor James Martin for their guidance during my DPhil. I am deeply grateful for the diligence and patience they invested in my supervision. Their support played a decisive role in the outcome of this journey.

I am grateful to everyone who contributed to my education before Oxford. In the middle of the chaos taking place in Serbia in late 1990s, my teacher Ljupka Todosijević invested enormous amounts of energy in order to equip her class with the values one should strive for in life. In the later grades, my teacher of mathematics Mima Milošević encouraged me to take part in maths competitions and kept me interested in mathematics along the way.

A couple of years later, I started at the Mathematical Grammar School in Belgrade. The school's culture that cherishes curiosity, dialogue, hard work and openness shaped me as a person. Fully aware of the fact that the Mathematical Grammar School's excellence is primarily a result of the enthusiasm and sacrifice of its employees throughout decades, rather than simply a consequence of the systematic approach towards the system of education, I am truly grateful to all my teachers there.

I first got to know Probability Theory in the Faculty of Mathematics of the University of Belgrade, and I would like to thank Professor Pavle Mladenović for introducing me to its basics.

Oxford is a beautiful, welcoming place where one easily feels at home. I am grateful to the members of the Probability group at Department of Statistics who provided a stimulating and supportive working environment. I would like to thank my friends Franz, Dane and Helmut for the lively office dynamic. Our office has been a place of discussions about coffee and politics, where events of probability zero were occasionally realized!

Finally, I would like to thank to my family and friends for their willingness to listen, encourage and cheer. Special thanks go to Tiago, for his selfless support during the various DPhil battles.

Abstract

We study certain questions related to the performance of the Karp-Sipser algorithm on the sparse Erdős-Rényi random graph.

The Karp-Sipser algorithm, introduced by Karp and Sipser [34] is a greedy algorithm which aims to obtain a near-maximum matching on a given graph. The algorithm evolves through a sequence of steps. In each step, it picks an edge according to a certain rule, adds it to the matching and removes it from the remaining graph. The algorithm stops when the remining graph is empty. In [34], the performance of the Karp-Sipser algorithm on the Erdős-Rényi random graphs $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ and $G(n, p = \frac{c}{n})$, $c > 0$ is studied. It is proved there that the algorithm behaves near-optimally, in the sense that the difference between the size of a matching obtained by the algorithm and a maximum matching is at most $o(n)$, with high probability as $n \to \infty$. The main result of [34] is a law of large numbers for the size of a maximum matching in $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ and $G(n, p = \frac{c}{n})$, $c > 0$. Aronson, Frieze and Pittel [2] further refine these results. In particular, they prove that for $c < e$, the Karp-Sipser algorithm obtains a maximum matching, with high probability as $n \to \infty$; for $c > e$, the difference between the size of a matching obtained by the algorithm and the size of a maximum matching of $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ is of order $\Theta_{\log n}(n^{1/5})$, with high probability as $n \to \infty$. They further conjecture a central limit theorem for the size of a maximum matching of $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ and $G(n, p = \frac{c}{n})$ for all $c > 0$. As noted in [2], the central limit theorem for $c < 1$ is a consequence of the result of Pittel [45]. In this thesis, we prove a central limit theorem for the size of a maximum matching of both $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ and $G(n, p = \frac{c}{n})$ for $c > e$. (We do not analyse the case $1 \leq c \leq e$.) Our approach is based on the further analysis of the Karp-Sipser algorithm. We use the results from [2] and refine them.

For $c > e$, the difference between the size of a matching obtained by the algorithm and the size of a maximum matching is of order $\Theta_{\log n}(n^{1/5})$, with high probability as $n \to \infty$, and the study [2] suggests that this difference is accumulated at the very end of the process. The question how the Karp-Sipser algorithm evolves in its final stages for $c > e$, motivated us to consider the following problem in this thesis. We study a model for the destruction of a random network by fire. Let us assume that we have a multigraph with minimum degree at least 2 with real-valued edge-lengths. We first choose a uniform random point from along the length and set it alight. The edges burn at speed 1. If the fire reaches a node of degree 2, it is passed on to the neighbouring edge. On the other hand, a node of degree at least 3 passes the fire either to all its neighbours or none, each with probability 1/2. If the fire extinguishes before the graph is burnt, we again pick a uniform point and set it alight. We study this model in the setting of a random multigraph with N nodes of degree 3 and $\alpha(N)$ nodes of degree 4, where $\alpha(N)/N \to 0$ as $N \to \infty$. We assume the edges to have i.i.d. standard exponential lengths. We are interested in the asymptotic behaviour of the number of fires we must set alight in order to burn the whole graph, and the number of points which are burnt from two different directions. Depending on whether $\alpha(N) >> \sqrt{N}$ or not, we prove that after the suitable rescaling these quantities converge jointly in distribution to either a pair of constants or to (complicated) functionals of Brownian motion. Our analysis supports the conjecture that the difference between the size of a matching obtained by the Karp-Sipser algorithm and the size of a maximum matching of the Erdős-Rényi random graph $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ for $c > e$, rescaled by $n^{1/5}$, converges in distribution.

Keywords: random graph, Erdős-Rényi model, configuration model, exploration process, Markov chain, fluid limit method, diffusion limit, matchings, maximum matching, maximum matching number, Karp-Sipser algorithm, central limit theorem.

Table of contents

Chapter 1

Introduction

Some notation

In this section we introduce some notation we use throughout the thesis.

- o **notation.**

  For non-negative functions f and g, we say that $f(n) = o(g(n))$ as $n \to \infty$, if

  $$\frac{f(n)}{g(n)} \to 0,$$

  as $n \to \infty$.

- O **notation.**

  For non-negative functions f and g, we say that $f(n) = O(g(n))$ as $n \to \infty$, if there exist constants $M < \infty$ and $n_0 \in \mathbb{N}$ such that

  $$\frac{f(n)}{g(n)} < M,$$

  for all $n > n_0$.

- $O_{\log n}$ **notation.**

For non-negative functions f and g, we say that $f(n) = O_{\log n}(g(n))$ as $n \to \infty$, if there exist $a \in \mathbb{R}$ such that

$$f(n) = O(g(n)(\log n)^a).$$

- **Ω notation.**

  For non-negative functions f and g, we say that $f(n) = \Omega(g(n))$ as $n \to \infty$, if there exists constants $0 < M < \infty$ and $n_0 \in \mathbb{N}$ such that

  $$\frac{f(n)}{g(n)} > M,$$

  for all $n > n_0$.

- **$\Omega_{\log n}$ notation.**

  For non-negative functions f and g, we say that $f(n) = \Omega_{\log n}(g(n))$ as $n \to \infty$, if there exist $a \in \mathbb{R}$ such that

  $$f(n) = \Omega(g(n)(\log n)^a).$$

- **Θ notation.**

  For non-negative functions f and g, we say that $f(n) = \Theta(g(n))$ as $n \to \infty$, if both $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$ as $n \to \infty$. That is, there exist constants $0 < M_1 < M_2 < \infty$ and $n_0 \in \mathbb{N}$ such that

  $$M_1 < \frac{f(n)}{g(n)} < M_2,$$

  for all $n > n_0$.

- **$\Theta_{\log}$ notation.**

  For non-negative functions f and g, we say that $f(n) = \Theta_{\log}(g(n))$ as $n \to \infty$, if we have both $f(n) = O_{\log n}(g(n))$ and $f(n) = \Omega_{\log n}(g(n))$ as $n \to \infty$.

- **With high probability.**

Let $(\Omega^n, \mathcal{F}^n, \mathbb{P}^n)$, $n \geq 0$, be a sequence of probability spaces. We say that a sequence of events $\{E^n\}_{n\geq 0}$, $E^n \in \mathcal{F}^n$, happens *with high probability* as $n \to \infty$ if

$$\mathbb{P}^n(E^n) \to 1,$$

as $n \to \infty$. We often slightly abuse the notation by omitting the superscript n in $\mathbb{P}^n$.

- $\Theta(h(n))$ **with high probability.** Let $(\Omega^n, \mathcal{F}^n, \mathbb{P}^n)$, $n \geq 0$, be a sequence of probability spaces, and for $n \geq 0$ let A^n be a non-negative random variable on $(\Omega^n, \mathcal{F}^n, \mathbb{P}^n)$. For a positive function $h(n)$, we say that $A^n = \Theta(h(n))$ with high probability as $n \to \infty$ if there exist constants $0 < M_1 < M_2 < \infty$ such that

$$\mathbb{P}^n\left(M_1 \leq \frac{A^n}{h(n)} \leq M_2\right) \to 1,$$

as $n \to \infty$. We often slightly abuse the notation by omitting the superscript n in $\mathbb{P}^n$.

- $\Theta_{\log n}(h(n))$ **with high probability.** Let $(\Omega^n, \mathcal{F}^n, \mathbb{P}^n)$, $n \geq 0$, be a sequence of probability spaces, and for $n \geq 0$ let A^n be a non-negative random variable on $(\Omega^n, \mathcal{F}^n, \mathbb{P}^n)$. For a positive function $h(n)$, we say that $A^n = \Theta_{\log n}(h(n))$ with high probability as $n \to \infty$ if there exist $a_1, a_2 \in \mathbb{R}$ such that

$$\mathbb{P}^n\left((\log n)^{a_1} \leq \frac{A^n}{h(n)} \leq (\log n)^{a_2}\right) \to 1,$$

as $n \to \infty$. We often slightly abuse the notation by omitting the superscript n in $\mathbb{P}^n$.

- **Bounded with high probability.**

  Let $(\Omega^n, \mathcal{F}^n, \mathbb{P}^n)$, $n \geq 0$, be a sequence of probability spaces, and for $n \geq 0$ let A^n be a non-negative random variable on $(\Omega^n, \mathcal{F}^n, \mathbb{P}^n)$. We say that A^n is *bounded with high probability* as $n \to \infty$ if there is a constant M such that

$$\mathbb{P}^n(A^n < M) \to 1,$$

as $n \to \infty$. We often slightly abuse the notation by omitting the superscript n in $\mathbb{P}^n$.

- **Convergence $\xrightarrow{p}$.**

  Let A^n, $n \geq 0$ and A be random variables defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The convergence $A^n \xrightarrow{p} A$ denotes the usual convergence in probability. That is, $A^n \xrightarrow{p} A$ as $n \to \infty$ if for every $\epsilon > 0$ we have

  $$\mathbb{P}(|A^n - A| > \epsilon) \to 0,$$

  as $n \to \infty$.

  Let $a \in \mathbb{R}$ be a fixed constant, $(\Omega^n, \mathcal{F}^n, \mathbb{P}^n)$, $n \geq 0$, a sequence of probability spaces and A^n a random variable defined on $(\Omega^n, \mathcal{F}^n, \mathbb{P}^n)$. We say that $A^n \xrightarrow{p} a$ as $n \to \infty$ if for each $\epsilon > 0$ we have

  $$\mathbb{P}^n(|A^n - a| > \epsilon) \to 0,$$

  as $n \to \infty$. We often slightly abuse the notation by omitting the superscript n in $\mathbb{P}^n$.

1.1 The Erdős-Rényi random graph

The study of random graphs started flourishing with the groundbreaking research of Paul Erdős and Alfréd Rényi in a series of papers at the beginning of the 1960s [17, 18, 19, 20] where they explore a number of properties of two most famous random graph models. Let us start by introducing these models. The Erdős-Rényi random graph $G(n, M)$ is a graph-valued random variable on the space of graphs with vertices labelled by $\{1, \ldots, n\}$, which is equal to any fixed graph with M edges with probability

$$\binom{\binom{n}{2}}{M}^{-1}.$$

The other model, the Erdős-Rényi random graph $G(n, p)$ is a graph-valued random variable on the space of graphs with vertices labelled by $\{1, \ldots, n\}$, which is equal to any fixed graph with $0 \leq e \leq \binom{n}{2}$ edges with probability

$$p^e(1-p)^{\binom{n}{2}-e}.$$

It is easy to see that the graph $G(n, p)$, conditioned on having exactly M edges, has the same distribution as $G(n, M)$. As the expected number of edges in $G(n, p)$ equals $\binom{n}{2}p$, it is perhaps not surprising that as $n \to \infty$, random graphs $G(n, p)$ and $G(n, M = \binom{n}{2}p)$ exhibit similar behaviour with respect to a wide range of properties, and it is often a routine task to translate a result obtained for one of them to the other.

Before we provide a brief overview of some properties of the Erdős-Rényi random graph, let us note the following convenient feature of the model. One can think of a *random graph process* which begins with a graph with n vertices and no edges, and evolves by adding one edge at a step, up until the natural end of the process when there is an edge between each pair of vertices. At each step, we add a uniform random edge from among those not yet present. It follows from the construction that this process is a Markov chain, and that for $0 \leq M \leq \binom{n}{2}$, the graph we have in the Mth step of the procedure is Erdős-Rényi random graph $G(n, M)$. Thus, we may employ the standard tools of Markov processes theory to study the properties of the Erdős-Rényi model.

In [20], Erdős and Rényi study how the properties of $G(n, M)$ change as the number of edges (as a function of n) varies. Interestingly, it turns out that there are a number of properties which exhibit a sudden change. A well-known such property concerns the size of the largest component. For $c > 0$, in both $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ and $G(n, p = \frac{c}{n})$ the expected degree of a uniform random node tends to c as $n \to \infty$. An interesting feature of both Erdős-Rényi random graph models is the well known *phase transition* for the emergence of the giant component occurring around the *critical value* $c = 1$. For $c < 1$, the graph consists of a forest of trees and a number of components which contain exactly one cycle, with high probability as $n \to \infty$. The size of the largest component is of order $\Theta(\log n)$, with high probability as $n \to \infty$. On the other hand, for $c > 1$, there is a unique *giant component*, i.e. a unique component which contains $\Theta(n)$ nodes, with high probability as $n \to \infty$. What remains outside the giant component is a forest of trees and a number of components which contain exactly one cycle. The second-largest component has a size of order $\Theta(\log n)$, with high probability as $n \to \infty$. For $c = 1$, the largest component contains $\Theta(n^{2/3})$ nodes, and same is true for the kth largest component for each $k \in \mathbb{N}$, with high probability as $n \to \infty$. This abrupt change in the structure near $c = 1$ was noted by Erdős and Rényi themselves in [20], and its discovery marked an important milestone for the area of random graphs.

The study of Bollobás [13] sheds light on how the giant component arises as c gets slightly larger than 1. More precisely, it identifies $c = 1 + 4(\log n)^{1/2} n^{-1/3}$ as a threshold for the property that there is a unique component of size greater than $n^{2/3}$, with high probability as $n \to \infty$. The work by Luczak et al [39] studies the properties of the graph within the critical window, that is for $c = 1 + \lambda n^{-1/3}$, $\lambda \in \mathbb{R}$. In [39] it is proved that the expected number of *complex components*, i.e. components with the number of edges being at least the number of nodes plus 1, is bounded. Moreover, both the largest and the smallest complex component have the size of order $\Theta(n^{2/3})$, with high probability as $n \to \infty$.

The further study of the critical window is done by Aldous [1]. Our study in Chapter 3 is very similar in style to the analysis done in [1], and hence we briefly sketch the result from [1]. The main idea is to access certain *global* characteristic of a graph by the successive use of certain *local* information. This idea is at the core of the *exploration process* (see for example Durrett [15], Chapter 2 for an introduction). During the exploration process, we distinguish between three categories of vertices: explored, discovered (but as-yet unexplored) and undiscovered. The process usually starts with a single node on a stack of discovered nodes and progresses through a large sequence of exploration steps. In each step, we explore one node from the stack of discovered nodes, i.e. we remove it from the stack and place its as-yet undiscovered neighbours on the stack. The study from [1] relies on the following exploration process. Initially, we place a uniform random node on a stack of discovered nodes. In the first step of the exploration, we explore that node, i.e. we remove it from the stack and place its as-yet undiscovered neighbours on the top of the stack. In each subsequent step of the exploration process, we take a node from the *bottom of the stack* and proceed in the similar way – we remove it from the stack and add all its as-yet undiscovered neighbours to the top of the stack. In this way, nodes are explored in the order of their discovery, i.e. we are doing a *breadth-first search* of a component (this differs from the *depth-first search*, where in each step we explore a node from the *top of the stack*, thus giving priority to the newly discovered neighbours). When the size of the stack drops down to 0, we have visited all nodes in the component that contains the initial node. This is due to the fact that there is a path between each two nodes in a component. If the size of the stack is zero and we have not explored the whole graph yet, we start exploring a new component by placing a uniform random as-yet unexplored node to the stack. The analysis from [1] makes use of the exploration process by exploiting the *branching process* heuristics. Namely, in each step of the process the newly discovered neighbours could be seen as children of the node we are currently exploring. Thus, the distribution of the increment of the stack coincides with the number of children of the node currently explored minus one (note that degree of each node added to the stack after the exploration of its neighbour has a size-biased degree distribution). It turns out that as long as we have not explored the whole graph, the size of the stack is a time inhomogeneous Markov chain with the reflection at zero. As the process eventually explores each node in a current component (exactly once), the sizes of components are encoded by the lengths of the intervals between two successive visits to zero of this Markov chain. Let $S^n(i)$ be the size of the stack after the ith step of the exploration process. The functional central limit theorem argument enables one to prove that as $n \to \infty$,

$$n^{-1/3} S^n(\lfloor n^{2/3} s \rfloor),$$

converges in distribution to a certain functional of a Brownian motion, uniformly on finite intervals. Having in mind the time rescaling, this enables us to deduce that the sizes of components, rescaled by $n^{2/3}$, jointly converge to the vector of excursion lengths of the limiting process.

As already mentioned, for $c > 1$ both Erdős-Rényi random graphs $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ and $G(n, p = \frac{c}{n})$ contain a unique giant component, i.e. a unique component with order $\Theta(n)$ vertices, with high probability as $n \to \infty$. We now explore the structure of the graph in this regime in some more detail. For $c > 1$, Pittel [45] proves a central limit theorem for the size of the giant component of both $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ and $G(n, p = \frac{c}{n})$. Moreover, in [45] it is proved that the giant component consists of a *giant* 2*-core*, i.e. the maximal subgraph of the giant component of minimum degree at least 2, and a *mantle of trees* around the core, with high probability as $n \to \infty$. Both the giant 2-core and the mantle of trees contain a number of vertices of order $\Theta(n)$, and [45] proves the law of large numbers for these quantities. Later work by Pittel and Wormald [47] gives a central limit theorem for the joint distribution of the fluctuations.

As long as the total number of edges is $O(n)$, the Erdős-Rényi random graph is said to be in the *sparse regime*. In this thesis, our primary interest lies in the properties of random graphs in this regime and we thus omit to overview the characteristics of Erdős-Rényi models outside the sparse regime. However, we emphasize that these are widely studied and understood and direct an interested reader to consult, for example, Janson et al [30].

1.2 The configuration model

The *configuration model*, introduced by Bender and Canfield [7] (see also Bollobás [12], and Wormald [54, 55, 56]) allows one to construct a graph with given degree sequence. This model will play an important role in the thesis. Given a set of n vertices, let

$(d_1, \cdots, d_n)$ be a vector of non-negative integers such that $\sum_{i=1}^n d_i$ is even and there exists at least one graph such that for each $i \in \{1, \ldots, n\}$, node i has degree d_i. We refer to $(d_1, \cdots, d_n)$ as a *degree sequence.* The construction works as follows. To each vertex i we attach d_i so-called *half-edges.* A *configuration* is a partition of the set of half-edges into $\frac{1}{2}\sum_{i=1}^n d_i$ pairs. We choose a uniform random configuration from the set of all possible configurations. Each pair of half-edges from the chosen configuration is thought of as an edge between the corresponding nodes. In general, in this way we obtain a *multigraph*, i.e. a graph which may have a number of loops and multiple edges, with a given degree sequence. A uniform random choice of a configuration determines a *random multigraph with a given degree sequence.* The probability of obtaining a fixed multigraph $\mathcal{M}$ with n vertices and a degree sequence $(d_1, \cdots, d_n)$ is

$$\frac{1}{(\sum_{i=1}^n d_i - 1)!!} \cdot \frac{\prod_{i=1}^n d_i!}{2^s \prod_{e \in E(\mathcal{M})} m_e!},$$

where $(\sum_{i=1}^n d_i - 1)!! = (\sum_{i=1}^n d_i - 1) \cdot (\sum_{i=1}^n d_i - 3) \ldots 1$, and s denotes the number of loops and m_e the multiplicity of an edge e from the set of edges $E(\mathcal{M})$ of $\mathcal{M}$. Note that a graph gets penalized for each loop and multiple edge, and so the choice of a uniform random configuration *does not* generate a uniform random multigraph with the given degree sequence. However, if we condition the multigraph to be simple, i.e. to have no loops or multiple edges, it is then uniform. Let us note that the configuration model allows us to generate edges one by one by pairing a uniform random pair of as-yet unpaired half-edges at each step. Indeed, as long as each configuration has probability $\frac{1}{(\sum_{i=1}^n d_i - 1)!!}$, the exact order in which we pair the half-edges does not matter. This property makes the configuration model particularly convenient for the exploration-type analysis.

Like the Erdős-Rényi random graph, the structure of the configuration model exhibits a phase transition for the emergence of a giant component. Molloy and Reed [41] prove this under certain conditions on the degree sequence. More precisely, given the sequence $\lambda_0, \lambda_1, \ldots$ of non-negative real numbers which sum to 1, the authors study a graph with approximately $\lambda_i n$ nodes of degree $i = 0, 1, \ldots$. Loosely speaking, their result says that, under certain regularity conditions on the degree sequence, if $\sum_i i(i-2)\lambda_i > 0$, a graph has a giant component (i.e. a component containing order $\Theta(n)$ vertices) with high probability as $n \to \infty$, whilst if $\sum_i i(i-2)\lambda_i < 0$, the size of all components of such a graph is negligible when compared to n, with high probability as $n \to \infty$. The threshold for the occurrence of the giant component becomes natural using the already mentioned branching process heuristics. Namely, the exploration of a node of degree i (which was previously placed to the stack as a consequence of the exploration of its neighbour) increases the size of the stack of discovered nodes by $i-2$. This is due to the fact that such a node has $i-1$ yet-undiscovered neighbours to be added to the stack, and is itself removed from the stack after the current step of the exploration. Recall that we can make use of the branching process heuristics by considering newly discovered neighbours of the node we are exploring as its children. Having in mind that we have approximately $\lambda_i n$ nodes of degree i, the threshold for the appearance of the giant component coincides with the threshold for the criticality of the branching process which is given by $\sum_i i(i-2)\lambda_i > 0$. In [42], the same authors explore the size of the giant component of this model in the supercritical regime, and the structure of the graph remaining after the deletion of the giant component. Joos et al [31] also consider the question whether the configuration model possesses a giant component, and provide other conditions on the degree sequence which guarantee its existence with high probability as $n \to \infty$. Joseph [32] and Riordan [49] (independently) consider the properties of the graph at criticality, and prove that under appropriate conditions on the degree sequence, the ordered component sizes rescaled by $n^{2/3}$ converge in distribution to the sequence of excursion intervals of a certain (reflected) functional of a Brownian motion. These works exploit the exploration process heuristics and are similar in flavour to the result by Aldous [1] about the critical Erdős-Rényi random graph.

1.3 Some techniques for the analysis of large stochastic systems

Some useful tools for studying properties of random graphs belong to the general framework for the analysis of density dependent population processes in large systems (see Ethier and Kurtz [21], Chapter 11). Such systems usually consist of a number n of similar entities which act separately of each other, possibly interacting only locally. As a result of those actions, the system changes its properties over time. Some examples of models falling in this category are: the spread of epidemic in a population consisting of susceptible and infected individuals; the evolution of a population where births and deaths are taking place; reactions in a chemical system consisting of a number of different reactants. It is often of interest to know how such systems evolve over time. The limited interaction between different entities usually guarantees that the dynamics of the process can be described by a Markov chain. Thus, it is natural to ask what can be said about the dynamic of a series of Markov chains parametrized by the size parameter n, as $n \to \infty$.

Kurtz [35] proves that, under certain conditions, such a Markov chain, rescaled by n, remains sharply concentrated around a certain deterministic function. Let us be a bit more precise about this result. Assume that $(A^n(i))_{i=0,1,\dots}$ is a discrete-time Markov chain taking values in $\mathbb{R}$, defined on some probability space. Let $b : \mathbb{R} \to \mathbb{R}$ be a Lipschitz function and $a : \mathbb{R}^+ \to \mathbb{R}$ the solution to a differential equation $a'(s) = b(a(s))$ with the initial condition $a(0) \in \mathbb{R}$. We further assume that: 1) the expected conditional increment of a jump of $\frac{A^n(\cdot)}{n}$ from a state ξ tends to $b(\xi)$ as $n \to \infty$; 2) we have $\frac{A^n(0)}{n} \xrightarrow{p} a(0)$ as $n \to \infty$; 3) with probability 1, there is an upper bound on the increments of $\frac{A^n(\cdot)}{n}$. The Lipschitz property of the function b together with 1), 2) and 3) ensures that the Markov chain evolves smoothly together with the function a, and is not subject to abrupt changes. Theorem 2.11 from [35] gives us that under these conditions, for any $\epsilon > 0$ we have

$$\mathbb{P}\left(\sup_{0\leq s\leq t}\left|\frac{A^n(\lfloor ns\rfloor)}{n} - a(s)\right| > \epsilon\right) \to 0, \tag{1.1}$$

that is, the Markov process after the *law of large numbers rescaling*, remains sharply concentrated around its expected trajectory. This type of reasoning is widely used in the analysis of large stochastic systems and is known as the *fluid limit method* or (in the combinatorics literature) as the *differential equations method.*

There is a rich body of literature that investigates or employs the fluid limit method. Chapter 11 of Ethier and Kurtz [21] provides a general setting for the analysis of sequences of Markov processes in this way. The survey by Darling and Norris [14] about the fluid limit method gives an overview of the state of the art and proves a theorem which guarantees the fluid limit result under fairly general conditions. Their setting, for example, does not require a process to be a Markov chain, but to be "close" to a Markov chain in a sense of condition 1) above. The paper by Wormald [57] also contains a version of a theorem which guarantees the fluid limit result, and provides an overview of some problems involving random processes on random graphs which can be analysed in this way.

Assuming that a Markov chain remains sharply concentrated around its expected trajectory in a sense of (1.1), it is natural to ask how the fluctuations around the expected trajectory behave under the central limit theorem scaling. Kurtz [36] provides conditions which guarantee that the fluctuations converge in distribution to a certain diffusion (in the space of càdlàg functions from $\mathbb{R}_+$ to $\mathbb{R}$, equipped with the Skorokhod topology). Some extensions of this result can be found in Barbour [5] and Kurtz and Protter [37]. Loosely speaking, the main ingredient in the proof of this type of result is the martingale functional central limit theorem. A nice overview is given in Chapter 11 of [21].

In this thesis, we will make an extensive use of the fluid limit method and the framework of the functional central limit theorem for the fluctuations. However, the aim of this section is only to sketch the basic idea, and we postpone greater elaboration to later chapters.

1.4 Random graphs and greedy algorithms

A *greedy* algorithm uses the simple heuristic of making a *locally* optimal choice in each step, in order to achieve a *globally* near-optimal solution to some problem of interest. There is a large body of literature which investigates the performance of greedy algorithms on random graphs. The primary motivation for this approach arises from the fact that random graphs can be used for the average case analysis of graph algorithms. We refer the reader to Frieze and McDiarmid [22] for a survey of the performance of various greedy algorithms on random graphs – algorithms for finding a Hamilton cycle in a graph (i.e. a cycle which visits each node exactly once), the shortest path, a large matching, certain colourings etc.

In this section, we do not aim to provide an extensive review of the literature studying greedy algorithms on random graphs, but rather focus on one particular algorithm which is somewhat related to the questions we study in Chapter 2 and Chapter 3. As the main interest of this thesis lies in some questions related to the matchings on random graphs, in forthcoming sections we further review the performance of some greedy algorithms for finding large matchings in random graphs.

1.4.1 The k-core of a random graph

The k-core of a graph is its maximal subgraph with minimal degree at least k. Note that the k-core of a given graph can be empty. In the Erdős-Rényi random graph, the threshold for the emergence of a *giant* 2-core, i.e. the threshold for the 2-core to contain $\Theta(n)$ vertices with high probability as $n \to \infty$, coincides with the threshold $c = 1$ for the emergence of the giant component. However, the size of the *first non-empty* 2-core which shows up in the random graph process, i.e. the first cycle, is bounded with high probability as $n \to \infty$ (see Janson [26] or Pittel et al [46]). The behaviour of the k-core for $k \geq 3$ at the point of the phase transition is somewhat different. In order to study a k-core for $k \geq 3$, Pittel et al [46] propose an algorithm which finds a k-core in a graph by using simple heuristics. In each step, the algorithm chooses a uniform random node with at least one and at most $k-1$ neighbours, if such a node exists, and deletes all its adjacent edges. The algorithm stops when there are no more edges in the graph (in which case the k-core is empty) or there are some edges in the graph, but all vertices are either isolated or have a degree at least k (in this case we have a non-empty k-core). Let us emphasize that this greedy algorithm always finds the (empty or non-empty) k-core, rather than just providing an approximate solution to the problem studied. It turns out that for a given $k \geq 3$, there is a threshold c_k such that for $c < c_k$ the k-core of the graph is empty with high probability as $n \to \infty$, while for $c > c_k$ there is a k-core which contains $\Theta(n)$ nodes, with high probability as $n \to \infty$. In addition, when the non-empty k-core, $k \geq 3$, appears for the first time, it already contains $\Theta(n)$ vertices, with high probability as $n \to \infty$ (recall that this is not the case with the 2-core).

Janson and Luczak [28] study the k-core, $k \geq 2$, of a configuration model with a given degree sequence. They provide conditions on the asymptotic of a degree sequence which guarantee that the k-core, $k \geq 3$, is empty with high probability as the number of nodes tends to infinity, and other conditions which imply that the number of nodes and edges in the k-core obey a law of large numbers. The situation with the 2-core is similar to that in the Erdős-Rényi model. That is, the probability of obtaining a non-empty 2-core is bounded away from zero even below the threshold for the emergence of the giant 2-core. The further study done in Janson and Luczak [29] investigates the location of the phase transition window, and the distributional properties of the fluctuations of the size of the k-core, $k \geq 2$, around its deterministic approximation at and above the threshold.

1.5 Matchings on random graphs

As already mentioned, the questions studied in this thesis are closely related to the problem of finding a large matching on the sparse Erdős-Rényi random graph. In this section, we briefly recall the prominent role problems related to matchings play in graph theory, and then highlight some important results related to matchings on random graphs.

A *matching* on a given graph is a subset of the set of edges with the property that each node is adjacent to at most one edge from that subset. A *maximal matching* has the property that if any further edge is added to it, one no longer has a matching, i.e. there is a node adjacent to more than one edge from the newly obtained subset of edges. The *size* of a matching is given by the number of its edges. A *maximum matching* is a matching that contains the largest possible number of edges - that is, there is no matching of the graph that contains more edges. In other words, a maximum matching is a maximal matching of the largest possible size. We refer to the size of a maximum matching of a given graph as its *maximum matching number.* A *perfect matching* is a matching which covers all vertices in the graph.

1.5.1 Matchings in combinatorics

Let us begin a brief account on the history of matchings with a quotation from Lovász and Plummer [38], "[matching theory] has often been in attendance when many of the exciting new concepts in combinatorial optimization have been born". Indeed, some of the earliest problems studied in graph theory concern matchings. For a given graph, define a leaf-component to be a connected component of a subgraph obtained by the deletion of a single edge from the graph, which furthermore has the property that any further deletion of one edge does not disconnect it. A famous theorem of Julius Petersen from 1891 says that a connected 3-regular graph with at most two leaf-components contains a perfect matching (see, for example, Mulder [43]). In 1947, Tutte characterized all of the graphs with a perfect matching. Tutte's theorem ([38], Theorem 3.1.1) says that a graph G has a perfect matching if and only if there is no set of vertices S with the property that $G \setminus S$ contains more than $|S|$ components with odd number of vertices. The relationship between the structure of a graph and its matchings is further revealed by the Gallai-Edmonds structure theorem ([38], Theorem 3.2.1) which gives a decomposition of a graph into subgraphs based on the set of vertices which are not incident to at least one maximum matching. In 1965, Edmonds proposed the first polynomial-time algorithm for obtaining a maximum matching ([38], Theorem 9.1.8). Since then, a number of algorithms have been proposed with ever-increasing levels of efficiency (see [38], Table 9.1.1, and [10, 25, 53]).

1.5.2 Matchings in Erdős-Rényi model

Matchings in the Erdős-Rényi random graph have received a considerable amount of attention in the literature. The threshold for the existence of a perfect matching in $G(n, M)$ was proved by Erdős and Rényi themselves [18]. More precisely, let us assume that $c_1, c_2 \dots$ is a sequence of real numbers. Supposing that n is even, the findings from [18] give us that for the Erdős-Rényi random graph with n vertices and $M = \lfloor \frac{n(\log n + c_n)}{2} \rfloor$

edges, as $n \to \infty$ we have

$$\mathbb{P}\Big(G(n, M = \lfloor \tfrac{n(\log n + c_n)}{2} \rfloor) \text{ has a perfect matching } \Big) \to \begin{cases} 0 & \text{if } c_n \to -\infty, \\ e^{-e^{-c}} & \text{if } c_n \to c, \\ 1 & \text{if } c_n \to +\infty. \end{cases}$$

Let us emphasize that for such a graph the symmetric difference between the event that it possesses a perfect matching and the event that it is connected, is an event of probability tending to zero, and thus for $c_n \to c > 0$, the limiting probability $e^{-e^{-c}}$ is equal to the limiting probability that the graph is connected.

Our main interest lies in the properties of sparse Erdős-Rényi models, more precisely $G(n, p = \frac{c}{n})$ and $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ for $c > 0$. In the sparse regime, both models have $\Theta(n)$ isolated vertices, with high probability as $n \to \infty$, and so there is clearly no perfect matching. The interesting question in this case is to determine the size of a maximum matching.

1.5.3 Some greedy algorithms related to matchings

Dyer et al [16] analyse the performance of two greedy algorithms for obtaining a matching on the sparse Erdős-Rényi random graph. The first one, GREEDY uses the simplest possible heuristics: as long as there are some edges in the graph, it matches a uniform random edge and removes the corresponding pair of vertices together with all adjacent edges from the graph. The number of matched edges has order $\Theta(n)$, with high probability as $n \to \infty$; however, the difference between the number of edges in a matching obtained by GREEDY and the maximum matching number is also of order $\Theta(n)$, with high probability as $n \to \infty$. The other algorithm, MODIFIED GREEDY proceeds in the following way. As long as the remaining graph is non-empty, it first chooses a uniform random node and then a uniform random edge adjacent to it, adds that edge to the matching and removes the matched pair of vertices together with any adjacent edges from the graph. The algorithm MODIFIED GREEDY is better than GREEDY in a sense that it matches $\Theta(n)$ more nodes than GREEDY, with high probability as $n \to \infty$. However, MODIFIED GREEDY is still far from optimal as there is a gap of order $\Theta(n)$ between the number of edges matched by it and the maximum matching number, with high probability as $n \to \infty$. Fortunately, it turns out that it is possible to obtain a near-maximum matching on the sparse Erdős-Rényi random graph by using a greedy algorithm introduced by Karp and Sipser [34]. As this algorithm plays an important role in the thesis, in the following section we provide an extensive review of its properties.

Before we proceed, let us briefly comment on some related problems. Frieze et al [24] study the performance of the MINGREEDY algorithm for obtaining a matching on a uniform random 3-regular graph with n nodes. As long as the remaining graph is not empty, MINGREEDY chooses a uniform random node among the nodes of *minimum* degree, and then matches a uniform random edge adjacent to it and removes it from the graph. Note that the removal of an edge typically causes nodes to decrease their degrees. Hence, although the initial graph is 3-regular, as the algorithm runs we typically have some nodes of smaller degree. The algorithm turns out to be very successful in obtaining a large matching, as the expected number of nodes left unmatched is upper bounded by a quantity of order $\Theta_{\log n}(n^{1/5})$ [24].

A related problem on a random 3-regular graph is studied by Bal et al [4]. A 2-*matching* of a graph is its subgraph with maximum degree 2 and minimum degree 1. The study done in [4] investigates the performance of a greedy algorithm for obtaining a 2-matching on a random 3-regular graph. Loosely speaking, the algorithm adds edges to a 2-matching one by one so that it gives priority to nodes which are “in danger” of not being included in the final 2-matching due to having low degree in the remaining graph. The algorithm is successful as it fails to include at most $\Theta_{\log n}(n^{1/5})$ nodes to a 2-matching, with high probability as $n \to \infty$.

1.6 The Karp-Sipser algorithm

As already mentioned, the Karp-Sipser algorithm is a greedy algorithm for obtaining a large matching on a given graph. It was introduced by Karp and Sipser [34] for the purpose of obtaining a near-maximum matching in the sparse Erdős-Rényi random graph. The focus of this thesis is on certain questions related to the performance of this algorithm on the sparse Erdős-Rényi random graph.

Before we proceed, let us agree to refer to a vertex with degree 1 as a *pendant* vertex. The algorithm evolves as follows. Whenever there are pendant vertices in the graph, the algorithm chooses a uniform random pendant vertex, adds the adjacent edge to the matching and removes the corresponding pair of vertices together with all edges adjacent to them. We refer to the pendant vertex the algorithm chooses in this type of a step as the *leaf*, and to its only neighbour as the *partner*. If there are no pendant vertices (but the remaining graph is non-empty), the algorithm chooses a uniform random *edge*, adds it to the matching and removes the corresponding pair of nodes together with all edges adjacent to them. The algorithm stops when the remaining graph does not possess any edges.

We refer to steps the algorithm makes up until the first uniform random choice of an edge as *Phase 1*. Let us note that in each step of Phase 1 the algorithm matches a pendant vertex. We refer to the graph remaining at the end of Phase 1 as the *Karp-Sipser core*. The first uniform random choice of an edge we add to a matching marks the beginning of *Phase 2*, and this phase lasts up until the very end of the algorithm. Phase 2 typically combines both kinds of step, i.e. whenever there is a pendant vertex, the algorithm proceeds as before and matches a uniform random pendant vertex, and when this is not the case it matches a uniform random edge.

The algorithm performs optimally during Phase 1, i.e. there exists a maximum matching which contains all edges matched during this phase. In order to see this, let us fix an arbitrary pendant vertex. The neighbour of the fixed pendant vertex is covered by all maximum matchings. Indeed, it is possible to extend any matching which does not cover the neighbour by simply adding the edge adjacent to the fixed pendant vertex to it, and hence any such a matching is not a maximum matching. Now, let us take any maximum matching, remove the edge adjacent to the pendant vertex's neighbour and add the edge adjacent to the fixed pendant vertex instead. The newly obtained matching is clearly a maximum matching and it covers the fixed pendant vertex. This holds for all pendant nodes in the graph, and hence there indeed exists a maximum matching which contains all edges matched during Phase 1. However, the simple heuristic the algorithm uses at the beginning of Phase 2 represents a potential danger from the perspective of obtaining a near-maximum matching. Namely, in this phase the algorithm can match an edge which does not belong to any maximum matching. Despite its naive approach, the algorithm turns out to be very successful on certain classes of sparse random graphs.

The performance of the Karp-Sipser algorithm on the sparse Erdős-Rényi random graph is investigated by Karp and Sipser [34] and their findings are further refined by Aronson et al [2]. These results are crucial for our analysis in Chapter 2 and serve as a motivation for the problem we study in Chapter 3, as we explain in Section 1.8. Thus, we provide a detailed overview of these findings in the following section, and here just highlight the main results. For both sparse Erdős-Rényi random graph models $G(n, p = \frac{c}{n})$ and $G(n, M = \lfloor \frac{cn}{2} \rfloor)$, we have the following. For $c < e$ the Karp-Sipser algorithm finds a maximum matching, with high probability as $n \to \infty$. For $c > e$, the difference between the size of a matching obtained by the algorithm and the maximum matching number is of order $\Theta_{\log n}(n^{1/5})$, with high probability as $n \to \infty$. For $c = e$, the difference is of order $o(n)$, with high probability as $n \to \infty$, although we do not know its true order of size.

Let us recall the notion of the 2-core of a graph, i.e. its maximal subgraph with minimal degree at least 2 and related discussion from Section 1.4. The Karp-Sipser core is a subgraph of the 2-core, and is typically smaller than the 2-core. This is easily seen if we compare the algorithm for finding a 2-core from [46] to the Karp-Sipser algorithm: while the former removes the leaf and the adjacent edge, the latter removes both the leaf and its partner together with any edges adjacent to it. The discussion from Section 1.4 suggests that it would be interesting to explore the phase transition of the Karp-Sipser core in more detail. That is, it would be interesting to know what is the order of size of the number of vertices and edges in the Karp-Sipser core of sparse Erdős-Rényi random graph for $c = e$, as well as what is the size of the critical window. The numerical study done in Bauer and Golinelli [6] suggests that the critical Karp-Sipser core has $\Theta(n^{3/5})$ vertices. However, there are no rigorous results regarding this issue.

We have already emphasized that during Phase 1 the algorithm does not "make any mistakes" in the sense that there exists a maximum matching which contains all edges matched during this phase. Thus, for the Karp-Sipser algorithm to find a near-maximum matching, we need either a sufficiently small Karp-Sipser core (so that even a terrible performance of the algorithm during Phase 2 leads only to a small difference between the size of a matching obtained and the maximum matching number), or a "well behaved" Karp-Sipser core, i.e. such a core that even the naive heuristics causes only a negligible difference between the size of the obtained matching and the maximum matching number. A class of such "well behaved" cores is studied by Bohman and Frieze [11]. More precisely, in [11] the class of random graphs with a given degree sequence and minimal degree at least 2 is studied. The authors provide asymptotic conditions on the degree sequence which guarantee that the Karp-Sipser algorithm matches all but $o(n)$ vertices, with high probability as the number of vertices n tends to infinity. Consequently, the Karp-Sipser algorithm finds a near-maximum matching in any random graph whose Karp-Sipser core has the distribution of the model studied in [11]. The sparse Erdős-Rényi model belongs to this class. The study from [2] provides more precise bounds on the discrepancy between the size of the obtained matching and the maximum matching number for the Erdős-Rényi model, while the findings from [11] identify a wider class of graphs where the algorithm behaves nearly optimally.

The approach the Karp-Sipser algorithm takes by giving priority to pendant vertices, served as a motivation for a construction of similar algorithms for related problems. This is for example the case of the greedy algorithm for obtaining a large 2-matching on a random 3-regular graph studied in [4], which we already mentioned in Section 1.5.3. Recall that this algorithm also gives priority to vertices with low degree. Interestingly, the order of size of the number of nodes which remain uncovered by the 2-matching is $\Theta_{\log n}(n^{1/5})$ with high probability as $n \to \infty$, and this coincides with the order of size of the number of nodes the Karp-Sipser algorithm fails to match during Phase 2 for the sparse Erdős-Rényi model for $c > e$.

1.6.1 The Karp-Sipser core

In this section, we present the proof from Bauer and Golinelli [6] that the Karp-Sipser core of a given graph does not depend on the order in which we match pendant vertices. For this purpose, we study the *leaf removal algorithm.* The leaf removal algorithm on a fixed graph behaves in the same way as the Karp-Sipser algorithm, with the amendment that once there are no pendant vertices, the algorithm finishes and returns the remaining graph as its output. Obviously, for a fixed graph, the subgraph induced by non-isolated vertices of the output graph of the leaf removal algorithm coincides with the Karp-Sipser core. Here we prove that, for a given graph, the subgraph induced by non-isolated vertices of the output graph of the leaf removal algorithm is deterministic, i.e. it does not depend on the order in which we remove pendant vertices.

Definition 1. *For a given graph G^0, a history of the leaf removal algorithm on G^0 is a sequence*

$$G^0, (u^0, w^0), G^1, (u^1, w^1), \dots, G^{k-1}, (u^{k-1}, w^{k-1}), G^k$$

where $k \geq 0$ and for each $0 \leq i \leq k-1$, G^i is the remaining graph after i steps of the leaf removal algorithm and (u^i, w^i) is the ordered pair where w^i is the leaf chosen for the removal in the $(i+1)$st step, and u^i is its partner. The output graph G^k is the first graph in the sequence with no pendant vertices.

Equipped with this notion, we are ready to prove the following lemma.

Lemma 2 (Bauer and Golinelli [6], Appendix)**.** *For a fixed graph, both the Karp-Sipser core and the number of the isolated vertices of the output of the leaf removal algorithm, do not depend on a history.*

Proof. We prove the claim by induction in the number of vertices of the initial graph. Let us note that, if the initial graph has either 0 or 1 vertices, there are no pendant vertices to choose. Thus, we have a unique history of the leaf removal algorithm and so the claim holds trivially. Let us suppose that the claim holds for any graph with at most $n-1$ vertices, and consider an arbitrary graph G with $n \geq 2$ vertices. We distinguish between the following cases.

1. If a graph G does not have any pendant vertices, there is a unique history and so the claim holds.

2. If G has exactly one pendant vertex, all histories start with the removal of the same pair of vertices, thus leading to the same remaining graph with $n-2$ vertices. The induction hypothesis guarantees that the final output does not depend on the history.

3. If G has at least two pendant vertices, we will compare two histories

$$\mathcal{H} = G, (u^0, w^0), G^1, \dots \quad \text{and}$$
$$\tilde{\mathcal{H}} = G, (\tilde{u}^0, \tilde{w}^0), \tilde{G}^1, \dots$$

arising as a result of two pendant nodes w^0 and $\tilde{w}^0$ being chosen by the algorithm.

   (a) If $\{u^0, w^0\} = \{\tilde{u}^0, \tilde{w}^0\}$, then $G^1 = \tilde{G}^1$ and the induction hypothesis guarantees that the final output does not depend on the history.

   (b) If $u^0 = \tilde{u}^0$ but $w^0 \neq \tilde{w}^0$, then the two pendant vertices w^0 and $\tilde{w}^0$ have a common neighbour. In this case, G^1 has $\tilde{w}^0$ as an isolated vertex, $\tilde{G}^1$ has w^0 as an isolated vertex, but $G^1 \setminus \{\tilde{w}^0\}$ and $\tilde{G}^1 \setminus \{w^0\}$ coincide. The inductive hypothesis yields the proof.

   (c) Assume that u^0, w^0, $\tilde{u}^0$, $\tilde{w}^0$ are four different vertices. Let us note that $\tilde{w}^0$ is pendant in G^1 and w^0 is pendant in $\tilde{G}^1$, and that the graph obtained by a removal of $(\tilde{u}^0, \tilde{w}^0)$ from G^1 coincides with the graph obtained by a removal of (u^0, w^0) from $\tilde{G}^1$. Denote that graph by G'', and let $\mathcal{H}''$ be a history on G''. A history on G given by $\mathcal{H}_1 = G, (u^0, w^0), G^1, (\tilde{u}^0, \tilde{w}^0), \mathcal{H}''$ and the history $\mathcal{H}$ have to result in the same core and the same number of isolated vertices, as the inductive hypothesis applies to the graph G^1. The same argument gives us that $\tilde{\mathcal{H}}_1 = G, (\tilde{u}^0, \tilde{w}^0), \tilde{G}^1, (u^0, w^0), \mathcal{H}''$ and $\tilde{\mathcal{H}}$ result in the same core and the same number of isolated vertices. Thus, it is enough to note that $\mathcal{H}_1$ and $\tilde{\mathcal{H}}_1$ yield the same core and the same number of isolated vertices.

We explored all possible cases, and so the result follows by induction. □

1.7 The Karp-Sipser algorithm on the sparse Erdős-Rényi random graph

As already mentioned, the Karp-Sipser algorithm was first introduced in [34] in order to analyse the sparse Erdős-Rényi random graph $G(n, p = \frac{c}{n})$, $c > 0$. The paper [34] provides a law of large numbers for the size of a maximum matching of $G(n, p = \frac{c}{n})$, $c > 0$. This result translates straightforwardly to the Erdős-Rényi random graph $G(n, M = \lfloor \frac{cn}{2} \rfloor)$, $c > 0$. The work of Aronson, Frieze and Pittel [2] refines these findings and provides precise bounds on the discrepancy between the size of a matching obtained by the algorithm and the size of a maximum matching. The analysis from [2] is an essential ingredient of our analysis in Chapter 2. Thus, in this section we highlight the major aspects of the study done in [2].

1.7.1 The model

The model studied in [2] is based on the idea used in the configuration model we have introduced in Section 1.2. Let us note that we do not know the degree sequence of the Erdős-Rényi random graph. However, it is possible to adjust the idea to the case of $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ in the following way. Firstly, distribute independently each of $2M$ half-edges among the n vertices with the same probability (i.e. distribute $2M$ balls into n bins). This step generates a degree sequence (provided that there exists at least one graph with that degree sequence). Secondly, given the degree sequence, choose a uniform random pairing of half-edges. Thirdly, declare each pair of half-edges to be an edge in the graph. In general, in this way we get a random multigraph with n nodes and M edges. Conditional on being simple, the graph constructed in this way has the same distribution as the Erdős-Rényi random graph $G(n, M)$. As the probability of obtaining a simple graph is bounded away from 0 as $n \to \infty$ ([2], Lemma 1 or see Remark 26 in Chapter 2), any event that happens with high probability in this model happens with high probability in the Erdős-Rényi random graph $G(n, M)$ too.

The performance of the Karp-Sipser algorithm on a multigraph is somewhat delicate. When there are no pendant vertices, the algorithm chooses a uniform random edge in order to add it to the matching, but in a case of a multigraph this could be a loop. If this happens, we say that the algorithm made an error and remove the adjacent node together with all edges adjacent to it from the graph. Fortunately, in the model we study, the total number of loops converges in distribution as $n \to \infty$ (see Lemma 24 in Chapter 2), and so the errors made in this way do not affect the size of a matching we obtain on the order we are interested in.

1.7.2 The Markov chain

The analysis from [34] and [2] relies on the study of the asymptotic properties of a Markov chain which carries enough information in order to recover the size of the matching we obtain at the end of the algorithm. For $i = 0, 1, \dots,$ let $Y_1^n(i)$ and $Y_2^n(i)$ denote the number of nodes of degree 1 and the number of nodes of degree at least 2, respectively, and let $Y_3^n(i)$ denote the number of edges in the graph, after i steps of the Karp-Sipser algorithm. The results from [34] and [2] are based on the study of the random process

$$Y^n(i) = (Y_1^n(i), Y_2^n(i), Y_3^n(i)), \quad 0 \le i \le \mathcal{T}_3^n,$$

where

$$\mathcal{T}_3^n = \inf\{i \ge 0 : Y_3^n(i) = 0\},$$

denotes the natural end of the algorithm when there are no more edges to match. The state space of Y^n is the set of triples of the form

$$S = \{(y_1, y_2, y_3) : y_1, y_2, y_3 \in \mathbb{N}_0, y_1 + 2y_2 \le 2y_3, y_1 + y_2 \le n, y_3 \le M\}.$$

Note that we have $y_1 + 2y_2 \le 2y_3$ as each pendant node has an adjacent edge and each node of degree at least two has at least two adjacent edges. Let $(\mathcal{F}_i^n)_{i \ge 0}$ denote the natural filtration of the process Y^n.

The Markovian nature of Y^n is somewhat delicate. When the algorithm matches an edge, all neighbours of the matched pair of vertices decrease their degrees. Thus, we need information about the local structure of the graph around the matched edge in order to update $Y_1^n(\cdot)$, $Y_2^n(\cdot)$ and $Y_3^n(\cdot)$ correctly, and it is not immediately clear that this depends on the past only through the previous state of the Markov chain. The rest of this section considers this issue.

For $y \in S$, let $\mathfrak{C}_y$ be the set of the configurations building multigraphs with $n - y_1 - y_2$ nodes of degree 0, y_1 nodes of degree 1, y_2 nodes of degree at least 2 and y_3 edges. For a fixed configuration $C_y \in \mathfrak{C}_y$, let $\mathfrak{T}(C_y)$ denote the collection of elements of the form $\{e, \tilde{C}\}$ where e represents an edge (given by the two half-edges) eligible for the matching (that is, an edge adjacent to a pendant node if such a node exists, or any existing edge otherwise), and $\tilde{C}$ the resulting configuration after the edge removal. Let us note that two different elements of $\mathfrak{T}(C_y)$ can have the same resulting configuration (for example, matching any of the two pendant nodes with a common neighbour results in a same configuration). We will refer to $\mathfrak{T}(C_y)$ as a set of possible transitions from a fixed configuration C_y. Also, for a fixed transition $T \in \mathfrak{T}(C_y)$ of C_y given by $T = (e, \tilde{C})$, we will write $\mathbf{c}(T)$ for the resulting configuration $\tilde{C}$.

Lemma 3 ([2], Claim within Lemma 2 there)**.** *Let $x, y \in S$. Each configuration $C_y \in \mathfrak{C}_y$ can be obtained by the same number of transitions from a configuration from $\mathfrak{C}_x$, that is*

$$\sum_{C_x \in \mathfrak{C}_x} \sum_{T \in \mathfrak{T}(C_x)} \mathbb{1}_{\{C_y \equiv \mathbf{c}(T)\}}$$

is equal for all $C_y \in \mathfrak{C}_y$.

Proof. Firstly, let us note that if either $y_2 > x_2$ or $y_3 \geq x_3$ then the removal of an edge from a configuration from $\mathfrak{C}_x$ cannot result in a configuration from $\mathfrak{C}_y$, and so in this case the claim trivially holds. Let us thus assume that $y_2 \leq x_2$ and $y_3 < x_3$. Let us firstly assume that $x_1 > 0$ and so we necessarily removed a pendant vertex from a configuration from $\mathfrak{C}_x$ which preceded the given configuration $C_y \in \mathfrak{C}_y$ (recall that Karp-Sipser always matches a pendant vertex, if such a vertex exists). In order to recover the configuration from $\mathfrak{C}_x$ which preceded $C_y \in \mathfrak{C}_y$, we have to make the following sequence of steps.

1. Choose an ordered pair (u, w) of vertices from the set of isolated vertices of $C_y \in \mathfrak{C}_y$ and add an edge between them. This is the edge which was matched in the previous step of the algorithm. Let w represent the leaf and u its partner.

2. Choose $l = (n - y_1 - y_2) - (n - x_1 - x_2) - 2$ isolated vertices (other than u and w) from the set of isolated vertices of $C_y \in \mathfrak{C}_y$. These are vertices which had some adjacent edges before the last step of the algorithm and became isolated after that step.

3. Choose a number m'' which satisfies $l \leq m'' \leq x_3 - y_3 - 1$. Let m'' represent the number of edges joining u to vertices other than w which became isolated in the previous step of the algorithm (recall that we allow multiple edges and self-loops).

4. Choose a number $\alpha < l$ such that $y_1 + \alpha + 1 + \mathbb{1}_{\{x_3=y_3+1\}} - x_1 \geq 0$ and $x_1 - \alpha - 1 - \mathbb{1}_{\{x_3=y_3+1\}} \geq 0$. We let α represent the number of those isolated nodes (other than u and w) in C_y which had degree 1 before the last step of the algorithm. Let l' be given by $l' = y_1 + \alpha + 1 + \mathbb{1}_{\{x_3=y_3+1\}} - x_1$. Thus, l' represents the number of vertices which are pendant in C_y but had a degree greater than 1 before the last step of the algorithm.

5. Choose a surjection $\phi : [m''] \to [l]$ such that $\alpha = |\{v | \exists! e \text{ s.t. } \phi(e) = v\}|$. This step enables us to connect l isolated vertices to the matched vertex u, taking care that exactly α of them obtain a degree equal to 1.

6. Choose l' vertices of degree 1 in C_y.

7. Choose a number m''' such that $l' \leq m''' \leq x_3 - y_3 - m'' - 1$ (unless $l' = 0$ in which case $m''' = 0$). We let m''' represent the number of edges joining u to nodes which stayed pendant in C_y but had a degree greater than 1 before the last step of the algorithm.

8. Choose a surjection $\psi : [m'''] \to [l']$. This step enables us to connect the matched vertex u to those pendant nodes in C_y which had a degree greater than 1 before the last step of the algorithm.

9. With each of the remaining $x_3 - y_3 - m'' - m''' - 1$ edges, we join u either with itself (thus forming a self-loop) or with a vertex of degree at least 2 in C_y.

Following this procedure step by step, we recover a configuration from $\mathfrak{C}_x$ which could have preceded C_y before the step of the Karp-Sipser algorithm. Let us note that in each step of the procedure the number of possible choices depends exclusively on x and y. Let us recall that the previous procedure applies in the case when $x_1 > 0$, that is, when we had a pendant node in the graph. The case $x_1 = 0$ requires minor changes. Loosely speaking, when recovering edges from the previous configuration, we allow both

u and w to form edges. More precisely, in step (1) we choose a (unordered) pair of isolated vertices (u, w), allowing $u \equiv w$. These vertices form an edge picked by the algorithm (the situation when the two vertices coincide corresponds to the case when the algorithm made an error by choosing a loop). In step (3), we have to choose m'' such that $2l \leq m'' \leq x_3 - y_3 - 1$ as no node was previously pendant, and so each node (other than u and w) which became isolated had at least two edges connecting it to a node from $\{u, w\}$. Given that u and w do not coincide, after choosing m'', we choose m_1'' and m_2'' so that $m'' = m_1'' + m_2''$ and m_1'' and m_2'' represent the number of edges joining w and u, respectively, to nodes which became isolated in the previous step (other than w and u). In step (4) we necessarily have $\alpha = 0$ as $x_1 = 0$. In step (7), after choosing m''' we proceed as before and choose m_1''' and m_2''', so that $m''' = m_1''' + m_2'''$, and these represent the number of edges connecting w and u, respectively, to the pendant nodes which previously had degree at least 2. We modify step (8) accordingly. In step (9), we let each of the remaining edges either connect v and w, become a loop on v or w, or connect any of v and w with a node of degree at least 2.

Thus, for each element from $\mathfrak{C}_y$ the number of possible transitions from an element from $\mathfrak{C}_x$ is the same. □

The following lemma tells us that, starting from a uniform random configuration $C_{Y^n(0)}$ for some $Y^n(0) \in S$, in each step of the algorithm we have a uniform random configuration corresponding to the current state of the process Y^n.

Lemma 4 ([2], Lemma 2)**.** *For a given $Y^n(0) \in S$, let $C^n(0)$ be a uniform random element from $\mathfrak{C}_{Y^n(0)}$. Let $C^n(0), C^n(1), \ldots, C^n(i)$ be the sequence of intermediate stages of the graph configuration as we run the Karp-Sipser algorithm for i steps, and $Y^n(0), Y^n(1), \ldots, Y^n(i)$ be the corresponding sequence of states of the random process Y^n up until step i. For any $C \in \mathfrak{C}_{Y^n(i)}$, we have*

$$P(C^n(i) = C | \mathcal{F}_i^n) = \frac{1}{|\mathfrak{C}_{Y^n(i)}|}.$$

Proof. We prove the claim by induction. The claim holds for $i = 0$. Let us assume that it holds for a fixed i, and prove that it holds for $i + 1$ too. For $Y^n(i + 1) = y \in S$, let C

be a fixed element from $\mathfrak{C}_y$. As $Y^n(i+1) = y$ is equivalent to $C^n(i+1) \in \mathfrak{C}_y$, we have

$$\begin{aligned} P\big(C^n(i+1) = C|\mathcal{F}^n_{i+1}\big) &= P\big(C^n(i+1) = C|C^n(i+1) \in \mathfrak{C}_y, \mathcal{F}^n_i\big) \\ &= \frac{P(C^n(i+1) = C|\mathcal{F}^n_i)}{P(Y^n(i+1) = y|\mathcal{F}^n_i)}, \end{aligned}$$

and this is further equal to

$$\begin{aligned} &P\big(C^n(i+1) = C|\mathcal{F}^n_{i+1}\big) \\ &= \frac{1}{P(Y^n(i+1) = y|\mathcal{F}^n_i)} \sum_{C' \in \mathfrak{C}_{Y^n(i)}} P\big(C^n(i+1) = C, C^n(i) = C'|\mathcal{F}^n_i\big) \\ &= \frac{1}{P(Y^n(i+1) = y|\mathcal{F}^n_i)} \sum_{C' \in \mathfrak{C}_{Y^n(i)}} P\big(C^n(i+1) = C|C^n(i) = C', \mathcal{F}^n_i\big) \cdot P\big(C^n(i) = C'|\mathcal{F}^n_i\big) \\ &= \frac{1}{P(Y^n(i+1) = y|\mathcal{F}^n_i)} \cdot \frac{1}{|\mathfrak{C}_{Y^n(i)}|} \sum_{C' \in \mathfrak{C}_{Y^n(i)}} P\big(C^n(i+1) = C|C^n(i) = C', \mathcal{F}^n_i\big), \end{aligned}$$

where the last equality holds by the induction hypothesis.

Starting from a configuration C', there are $|\mathfrak{T}(C')|$ possible transitions the algorithm can make, some of which might result in a given configuration C. Thus, we have

$$P\big(C^n(i+1) = C|C^n(i) = C', \mathcal{F}^n_i\big) = \frac{1}{|\mathfrak{T}(C')|} \cdot \sum_{T \in \mathfrak{T}(C')} \mathbb{1}_{\{C \equiv \mathbf{c}(T)\}},$$

and so the previous becomes

$$\begin{aligned} &P\big(C^n(i+1) = C|\mathcal{F}^n_{i+1}\big) \\ &= \frac{1}{P(Y^n(i+1) = y|\mathcal{F}^n_i)} \cdot \frac{1}{|\mathfrak{C}_{Y^n(i)}|} \cdot \sum_{C' \in \mathfrak{C}_{Y^n(i)}} \frac{1}{|\mathfrak{T}(C')|} \sum_{T \in \mathfrak{T}(C')} \mathbb{1}_{\{C \equiv \mathbf{c}(T)\}}. \end{aligned} \tag{1.2}$$

For any $C' \in \mathfrak{C}_{Y^n(i)}$, the number of possible transitions is given by the number of candidates for the matched edge, and so we have $|\mathfrak{T}(C')| = Y^n_1(i)$ for $Y^n_1(i) > 0$ and $|\mathfrak{T}(C')| = Y^n_3(i)$ otherwise. Thus, $|\mathfrak{T}(C')|$ is equal for all $C' \in \mathfrak{C}_{Y^n(i)}$. Also, Lemma 3 gives us that each element $C \in \mathfrak{C}_y$ can be obtained by the same number of transitions from an element from $\mathfrak{C}_{Y^n(i)}$, and so

$$\sum_{C' \in \mathfrak{C}_{Y^n(i)}} \sum_{T \in \mathfrak{T}(C')} \mathbb{1}_{\{C \equiv \mathbf{c}(T)\}}$$

is equal for all $C \in \mathfrak{C}_y$. Thus, we get that the right hand side of (1.2) is equal for all $C \in \mathfrak{C}_y$ and so the distribution is uniform on $\mathfrak{C}_y$. This completes the proof. □

Finally, we are ready to prove the following lemma.

Lemma 5. *The process Y^n is a Markov chain.*

Proof. This is a direct consequence of Lemma 4. From (1.2), for $i \geq 0$ and $y \in S$ we have

$$\mathbb{P}(Y^n(i+1) = y | \mathcal{F}_i^n) = \frac{1}{|\mathfrak{C}_{Y^n(i)}|} \sum_{C \in \mathfrak{C}_y} \sum_{C' \in \mathfrak{C}_{Y^n(i)}} \frac{1}{|\mathfrak{T}(C')|} \sum_{T \in \mathfrak{T}(C')} \mathbb{1}_{\{C \equiv \mathbf{c}(T)\}},$$

and this depends on y and $Y^n(i)$ exclusively. □

1.7.3 The degree distribution

Let us recall that our Markov chain Y^n keeps track of the number of nodes of degree 1 and at least 2, as well as the number of edges in the graph, as we run the algorithm. Lemma 4 tells us that, starting from a random multigraph with n nodes and $M = \lfloor \frac{cn}{2} \rfloor$ edges, the remaining graph after i steps of the algorithm is a random multigraph with $Y_1^n(i)$ nodes of degree 1, $Y_2^n(i)$ nodes of degree at least 2, and $Y_3^n(i)$ edges. In each step of the algorithm, we match a pair of vertices and remove them together with all adjacent edges, and this drives the Markov chain to its new state. In order to be able to determine the jump probabilities of Y^n, we need to be able to control the degree distribution. The essential ingredients of the analysis done in [2], which we also exploit in Chapter 2, are sharp approximations on the degree distribution. We now review this analysis in some detail, as it will be useful to us later.

Throughout this section, let

$$f(u) = e^u - 1 - u, \quad u > 0.$$

We will refer to a distribution of a Poisson random variable with the parameter $\lambda > 0$ conditioned on being greater than or equal to 2, as the *truncated Poisson distribution*. That is, the distribution of a truncated Poisson random variable Q with the parameter

$\lambda > 0$ is determined by

$$\mathbb{P}(Q = k) = \frac{\lambda^k}{k! f(\lambda)} \quad k = 2, 3, \ldots.$$

Lemma 6 (Lemma 4 in [2]). *Conditioned on $\mathcal{F}_i^n$, let $(D_1, \ldots, D_{Y_2^n(i)})$ be the degree sequence of the nodes of degree at least 2 and $(Q_1, \ldots, Q_{Y_2^n(i)})$ be a vector of $Y_2^n(i)$ independent truncated Poisson random variables with the parameter $\lambda > 0$. Then, $(D_1, \ldots, D_{Y_2^n(i)})$ has the same distribution as $(Q_1, \ldots, Q_{Y_2^n(i)})$ conditioned on*

$$\sum_{j=1}^{Y_2^n(i)} Q_j = 2Y_3^n(i) - Y_1^n(i).$$

Proof. Conditioned on $\mathcal{F}_i^n$, let $\mathcal{Q}$ be the set of all $(k_1, k_2, \ldots, k_{Y_2^n(i)})$ where for each $1 \le j \le Y_2^n(i)$ we have $k_j \ge 2$ and $k_1 + k_2 + \cdots + k_{Y_2^n(i)} = 2Y_3^n(i) - Y_1^n(i)$. Provided that this set is non-empty, for a fixed $(k_1, k_2, \ldots, k_{Y_2^n(i)}) \in \mathcal{Q}$, we have

$$\begin{aligned}
&\mathbb{P}\Big(Q_1 = k_1, \ldots, Q_{Y_2^n(i)} = k_{Y_2^n(i)} | Q_1 + \cdots + Q_{Y_2^n(i)} = 2Y_3^n(i) - Y_1^n(i), \mathcal{F}_i^n\Big) \\
&= \frac{\prod_{j=1}^{Y_2^n(i)} \frac{\lambda^{k_j}}{k_j! f(\lambda)}}{\sum_{(l_1, \ldots, l_{Y_2^n(i)}) \in \mathcal{Q}} \prod_{j=1}^{Y_2^n(i)} \frac{\lambda^{l_j}}{l_j! f(\lambda)}} \\
&= \frac{\prod_{j=1}^{Y_2^n(i)} \frac{1}{k_j!}}{\sum_{(l_1, \ldots, l_{Y_2^n(i)}) \in \mathcal{Q}} \prod_{j=1}^{Y_2^n(i)} \frac{1}{l_j!}}.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\mathbb{P}\Big(D_1 = k_1, \ldots, D_{Y_2^n(i)} = k_{Y_2^n(i)} | \mathcal{F}_i^n\Big) &= \frac{\frac{(2Y_3^n - Y_1^n(i))!}{k_1! k_2! \ldots k_{Y_2^n(i)}!}}{\sum_{(l_1, \ldots, l_{Y_2^n(i)}) \in \mathcal{Q}} \frac{(2Y_3^n - Y_1^n(i))!}{l_1! l_2! \ldots l_{Y_2^n(i)}!}} \\
&= \frac{\prod_{j=1}^{Y_2^n(i)} \frac{1}{k_j!}}{\sum_{(l_1, \ldots, l_{Y_2^n(i)}) \in \mathcal{Q}} \prod_{j=1}^{Y_2^n(i)} \frac{1}{l_j!}},
\end{aligned}$$

equals the same value, which completes the proof. $\square$

The parameter λ from Lemma 6 is an arbitrary positive number. However, it will turn out that a specific choice of the parameter significantly eases the approximation of the degree distribution. In order to prepare the ground for the approximation, let us introduce the following. For $x = (x_1, x_2, x_3) \in (\mathbb{R}_+ \cup \{0\})^3$ such that $x_2 > 0$, let z_x be the unique solution to

$$\frac{z_x(e^{z_x} - 1)}{f(z_x)} = \frac{2x_3 - x_1}{x_2}. \tag{1.3}$$

Let us introduce the auxiliary process Z^n by $Z^n(i) = z_{Y^n(i)}$ or equivalently

$$\frac{Z^n(i)(e^{Z^n(i)} - 1)}{f(Z^n(i))} = \frac{2Y_3^n(i) - Y_1^n(i)}{Y_2^n(i)}, \quad 0 \le i < \mathcal{T}_2^n, \tag{1.4}$$

where

$$\mathcal{T}_2^n = \min\{i \ge 0 : Y_2^n(i) = 0\}.$$

Note that at time $\mathcal{T}_2^n$ the remaining graph is either empty or it consists of a number of vertex-disjoint edges. It will turn out that $\mathcal{T}_2^n$ is close to $\mathcal{T}_3^n$, the time when the remaining graph is empty, with high probability as $n \to \infty$.

As we will shortly see, the auxiliary process Z^n turns out to be a particularly convenient choice of the parameter of the truncated Poisson distribution in Lemma 6, as it enables sharp approximations of the degree distribution. This is due to the fact that (1.4) ensures that the expected value of the sum of Y_2^n independent truncated Poisson random variables with the parameter Z^n equals the total number of half-edges adjacent to nodes of degree at least 2 in the given step. This choice of the parameter in Lemma 6 means that we are conditioning there on an event which is expected to happen. This will turn out to be a convenient feature, as we will see in a moment. Thus, from now on, conditioned on $\mathcal{F}_i^n$, we will be interested in the properties of the vector of $Y_2^n(i)$ independent truncated Poisson random variables with the parameter $Z^n(i)$.

For an arbitrary function $h(n)$ such that $h(n) \to \infty$ as $n \to \infty$, let

$$\Lambda_h^n = \inf\{i \ge 0 : Y_2^n(i) \cdot Z^n(i) < h(n)\}.$$

Remark 7. *Conditioned on $\mathcal{F}_i^n$, for a vector $(Q_1, \ldots, Q_{Y_2^n(i)})$ of $Y_2^n(i)$ independent truncated Poisson random variables with the parameter $Z^n(i)$, we have*

$$\mathbb{E}\left[\sum_{j=1}^{Y_2^n(i)} Q_j \cdot \mathbb{1}_{\{Q_j \geq 3\}} \Big| \mathcal{F}_i^n\right] = Y_2^n(i) \cdot Z^n(i),$$

(note that we do not impose any condition on the value of their sum). Thus, the stopping time Λ_h^n marks the time when the expectation on the left hand side drops below the threshold $h(n)$. Once we know that this vector of independent truncated Poisson random variables is a good approximation of the degree distribution (which we show shortly), Λ_h^n can be thought of as a time when the expected number of half-edges adjacent to nodes of degree at least 3 drops down below $h(n)$. For the appropriate choice of $h(n)$, it turns out that at time Λ_h^n the number of nodes in the remaining graph is sufficiently small, and so it is enough to follow the algorithm up until Λ_h^n in order to estimate the size of a matching it obtains sufficiently precisely.

Let also

$$\Gamma^n = \inf\{i \geq 0 : Z^n(i) > 3c\}.$$

Remark 8. *Conditioned on $\mathcal{F}_i^n$, let Q be a truncated Poisson random variable with the parameter $Z^n(i)$. For $i \leq \Gamma^n$, we have*

$$\begin{aligned}
\frac{\mathrm{Var}(Q|\mathcal{F}_i^n)}{Z^n(i)} &= \frac{1}{Z^n(i)} \cdot \frac{Z^n(i)(e^{Z^n(i)}-1)^2 - (Z^n(i))^3 e^{Z^n(i)}}{(f(Z^n(i)))^2} \\
&= 1 + \frac{2Z^n(i)\Big(\frac{(Z^n(i))^3}{3!} + \frac{(Z^n(i))^4}{4!} + \ldots\Big) - (Z^n(i))^2\Big(\frac{(Z^n(i))^2}{2!} + \frac{(Z^n(i))^3}{3!} + \ldots\Big)}{(f(Z^n(i)))^2} \\
&\leq 1 + \frac{4}{(Z^n(i))^4}\left(2Z^n(i)\left(\frac{(Z^n(i))^3}{3!} + \frac{(Z^n(i))^4}{4!} + \ldots\right)\right) \\
&\leq 1 + \frac{8}{3!} + 8e^{Z^n(i)} \\
&\leq 1 + \frac{4}{3} + 8e^{3c},
\end{aligned}$$

where the first inequality follows as $f(u) > \frac{u^2}{2}$, and the last one is the consequence of $Z^n(i) \leq 3c$ for $i < \Gamma^n$.

The following lemma from [2] will be a crucial tool for the approximation of the degree distribution.

Lemma 9 ([2], (5) and (6)). *For a given function $h(n)$ such that $h(n) \to \infty$ as $n \to \infty$, let $g(n)$ be a function such that $g(n) \le (h(n))^{1/4}$. Conditioned on $\mathcal{F}_i^n$, let $(Q_1, \dots, Q_{Y_2^n(i)})$ be a vector of $Y_2^n(i)$ independent truncated Poisson random variables with the parameter $Z^n(i)$. For $0 \le i < \Gamma^n \wedge \Lambda_h^n$ and $k \le g(n)$, we have*

$$\mathbb{P}\left(\sum_{j=1}^{Y_2^n(i)} Q_j = 2Y_3^n(i) - Y_1^n(i) \middle| \mathcal{F}_i^n\right) = \frac{1}{\sqrt{2\pi \mathrm{Var}(Q|\mathcal{F}_i^n) Y_2^n(i)}}\Big(1 + L_{h,g}^n(i)\Big),$$
$$\mathbb{P}\left(\sum_{j=2}^{Y_2^n(i)} Q_j = 2Y_3^n(i) - Y_1^n(i) - k \middle| \mathcal{F}_i^n\right) = \frac{1}{\sqrt{2\pi \mathrm{Var}(Q|\mathcal{F}_i^n) Y_2^n(i)}}\Big(1 + P_{h,g,k}^n(i)\Big),$$

where Q denotes a truncated Poisson random variable with the parameter $Z^n(i)$, and

$$\sup_{0 \le i < \Gamma^n \wedge \Lambda_h^n} Z^n(i) \cdot Y_2^n(i) \cdot |L_{h,g}^n(i)|,$$

as well as

$$\sup_{0 \le i < \Gamma^n \wedge \Lambda_h^n} \sup_{0 \le k \le g(n)} \frac{Z^n(i) \cdot Y_2^n(i)}{(g(n))^2} \cdot |P_{h,g,k}^n(i)|,$$

are bounded with high probability as $n \to \infty$.

We do not present the proof of Lemma 9, but only briefly comment on this result. Let us recall the definition of the process Z^n given by (1.4). Conditioned on $\mathcal{F}_i^n$, the left hand side of (1.4) is the expected value of a truncated Poisson random variable with the parameter $Z^n(i)$. Hence, we have

$$\mathbb{E}\left[\sum_{j=1}^{Y_2^n(i)} Q_j \middle| \mathcal{F}_i^n\right] = 2Y_3^n(i) - Y_1^n(i).$$

Lemma 9 deals with the deviations of $\sum_{j=1}^{Y_2^n(i)} Q_j$ from its expected value, and it thus represents a local limit theorem result. Let us note that Remark 8 guarantees that $\mathrm{Var}(Q|\mathcal{F}_i^n)$ and $Z^n(i)$ have the same order of size for $i < \Gamma^n$. This property is used in the proof of Lemma 9 from [2], as it enables one to express the asymptotics of the errors in terms of Y_2^n and Z^n (instead of Y_2^n and $\mathrm{Var}(Q|\mathcal{F}_i^n)$). The conditions of Lemma 9 ensure that the error terms $L^n_{h,g}$ and $P^n_{h,g,k}$ are tending to zero, with high probability as $n \to \infty$. Indeed, for $i < \Lambda^n_h$ we have $Z^n(i) \cdot Y^n_2(i) \geq h(n)$, and this together with $g(n) \leq (h(n))^{1/4}$ implies that

$$\sup_{0 \leq i < \Gamma^n \wedge \Lambda^n_h} \sup_{0 \leq k \leq g(n)} \frac{(g(n))^2}{Z^n(i) \cdot Y^n_2(i)} \leq \frac{1}{\sqrt{h(n)}},$$

and the quantity on the right hand side tends to zero as $n \to \infty$. Consequently, the asymptotic for $P^n_{h,g,k}$ given by Lemma 9 guarantees that $P^n_{h,g,k}$ tend to zero as $n \to \infty$. Similar reasoning gives us that the $L^n_{g,h}$ tend to zero.

Finally, we can provide the approximation of the degree distribution.

Lemma 10. *Let $D^n(i)$ be a degree of an arbitrary node of degree at least 2 in a graph after i steps of the algorithm. Let $h(n)$ be a function such that $h(n) \to \infty$ as $n \to \infty$, and $g(n)$ be such that $g(n) \leq (h(n))^{1/4}$. For $0 \leq i < \Gamma^n \wedge \Lambda^n_h$ and $2 \leq k \leq g(n)$, we have*

$$\mathbb{P}(D^n(i) = k | \mathcal{F}^n_i) = \frac{(Z^n(i))^k}{k! f(Z^n(i))} + R^n_{h,g,k}(i),$$

where as $n \to \infty$,

$$\sup_{0 \leq i < \Gamma^n \wedge \Lambda^n_h} \frac{h(n)}{(g(n))^3} \sup_{0 \leq k \leq g(n)} |R^n_{h,g,k}(i)| \xrightarrow{p} 0.$$

Also, letting $D^n_1(i)$ and $D^n_2(i)$ be the degrees of an arbitrary pair of distinct nodes of degree at least 2, for $0 \leq i < \Gamma^n \wedge \Lambda^n_h$ and $2 \leq k_1, k_2 \leq g(n)$ we have

$$\mathbb{P}(D^n_1(i) = k_1, D^n_2(i) = k_2 | \mathcal{F}^n_i) = \frac{(Z^n(i))^{k_1}}{k_1! f(Z^n(i))} \cdot \frac{(Z^n(i))^{k_2}}{k_2! f(Z^n(i))} + T^n_{h,g,k_1,k_2}(i),$$

where

$$\sup_{0 \leq i < \Gamma^n \wedge \Lambda^n_h} \frac{h(n)}{(g(n))^3} \sup_{0 \leq k_1, k_2 \leq g(n)} |T^n_{h,g,k_1,k_2}(i)| \xrightarrow{p} 0,$$

as $n \to \infty$.

Proof. Conditioned on $\mathcal{F}^n_i$, let $Q_1, \ldots, Q_{Y^n_2(i)}$ be $X^n_2(i)$ independent truncated Poisson random variables with the parameter $Z^n(i)$. Having in mind Lemma 6, for $2 \leq k \leq g(n)$

we have

$$\begin{aligned}
\mathbb{P}\Big(D^n(i)=k\Big|\mathcal{F}_i^n\Big) &= \mathbb{P}\left(Q_1=k\left|\sum_{j=1}^{Y_2^n(i)} Q_j = 2Y_3^n(i)-Y_1^n(i),\mathcal{F}_i^n\right.\right)\\
&= \frac{\mathbb{P}\Big(Q_1=k,\sum_{j=2}^{Y_2^n(i)} Q_j = 2Y_3^n(i)-Y_1^n(i)-k\Big|\mathcal{F}_i^n\Big)}{\mathbb{P}\Big(\sum_{j=1}^{Y_2^n(i)} Q_j = 2Y_3^n(i)-Y_1^n(i)\Big|\mathcal{F}_i^n\Big)}\\
&= \frac{(Z^n(i))^k}{k!f(Z^n(i))}\cdot\frac{\mathbb{P}\Big(\sum_{j=2}^{Y_2^n(i)} Q_j = 2Y_3^n(i)-Y_1^n(i)-k\Big|\mathcal{F}_i^n\Big)}{\mathbb{P}\Big(\sum_{j=1}^{Y_2^n(i)} Q_j = 2Y_3^n(i)-Y_1^n(i)\Big|\mathcal{F}_i^n\Big)}.
\end{aligned}$$

By Lemma 9, this is further equal to

$$\begin{aligned}
\mathbb{P}\Big(D^n(i)=k\Big|\mathcal{F}_i^n\Big) &= \frac{(Z^n(i))^k}{k!f(Z^n(i))}\left(1-\frac{L_{h,g}^n(i)-P_{h,g,k}^n(i)}{1+L_{h,g}^n(i)}\right)\\
&=: \frac{(Z^n(i))^k}{k!f(Z^n(i))}+R_{h,g,k}^n(i).
\end{aligned}$$

Let us note that

$$|R_{h,g,k}^n(i)| \le \left|\frac{L_{h,g}^n(i)-P_{h,g,k}^n(i)}{1+L_{h,g}^n(i)}\right| \le |L_{h,g}^n(i)|+|P_{h,g,k}^n(i)|,$$

and so the asymptotic behaviour of $R_{h,g,k}^n$ is a consequence of Lemma 9.

The proof for the asymptotic of T_{h,g,k_1,k_2}^n goes analogously with the obvious modifications. □

For a fixed function $h(n)$ such that $h(n)\to\infty$ as $n\to\infty$, Lemma 10 enables us to approximate the degree distribution up until time $\Gamma^n\wedge\Lambda_h^n$. Fortunately, this will turn out to be long enough for us to analyse the performance of the algorithm with sufficient precision. We will explain this in more detail in Chapter 2.

1.7.4 Phase 1

Let us recall that Phase 1 of the Karp-Sipser algorithm ends when for the first time there are no pendant vertices to match. In this section, we highlight the results from [34] and [2] about the properties of the multigraph during and at the end of Phase 1. In particular, this enables us to understand the phase transition occurring around $c=e$ in Erdős-Rényi random graph $G(n,M=\lfloor\frac{cn}{2}\rfloor)$.

The main object of study in [34] and [2] is the rescaled process

$$\frac{Y^n(\lfloor nt \rfloor)}{n}, \quad t \geq 0.$$

The expected infinitesimal changes of this process suggest that its expected trajectory tends to the solution of a certain differential equation as $n \to \infty$. In [34] and [2] it is proved that the rescaled process remains in an arbitrarily small tube around the limiting expected trajectory, with high probability as $n \to \infty$ (we will be more explicit about this in a moment). The proof is based on the fluid limit method we briefly introduced in Section 1.3.

Let us firstly define the deterministic function which represents the limiting expected trajectory of the rescaled process $\frac{Y^n(\lfloor n \cdot \rfloor)}{n}$ during Phase 1. As we will not present the proof of the main results from [34] and [2] here, the relationship between the process and the deterministic function will remain somewhat mysterious for the moment. However, in Chapter 2 we analyse Phase 1 closely for $c > e$, and the relationship between the two should become clearer.

Recall that for $u \geq 0$ we defined $f(u) = e^u - u - 1$. Let also

$$p(u) = e^{-u} f(u).$$

Next, for $u > 0$ define $\beta(u)$ (implicitly) by the unique solution to

$$\beta(u) e^{c\beta(u)} = e^u.$$

Finally, let the function $\vartheta(t)$, $\vartheta : \mathbb{R}_+ \to \mathbb{R}_+$ be given (implicitly) by

$$t = \frac{1}{c}\Big(c(1 - \beta(\vartheta(t))) - \frac{1}{2}\log^2 \beta(\vartheta(t))\Big),$$

and let $\chi(t) = (\chi_1(t), \chi_2(t), \chi_3(t))^T$, $\chi : \mathbb{R}_+ \to \mathbb{R}^3$ be given by

$$\begin{aligned}
\chi_1(t) &= \frac{1}{c}\Big(\vartheta^2(t) - \vartheta(t) \cdot c \cdot \beta(\vartheta(t))(1 - e^{-\vartheta(t)})\Big),\\
\chi_2(t) &= p(\vartheta(t))\beta(\vartheta(t)),\\
\chi_3(t) &= \frac{1}{2c}\vartheta^2(t).
\end{aligned}$$

We now reveal the nature of the parameter $\vartheta(t)$, $t \geq 0$. It turns out that we have the following relationship

$$\frac{\vartheta(t)\Big(e^{\vartheta(t)} - 1\Big)}{f(\vartheta(t))} = \frac{2\chi_3(t) - \chi_1(t)}{\chi_2(t)},$$

for $t \geq 0$. Having in mind that the function $\chi(t)$ should be the limiting expected trajectory of the rescaled process $\frac{Y^n(\lfloor n\cdot\rfloor)}{n}$, as well as the definition of the process Z^n given by (1.4), the function $\vartheta(t)$ represents the idealized trajectory of the process $Z^n(\lfloor n\cdot\rfloor)$ (note that the right-hand side of (1.4) remains unaffected by the space scaling imposed on Y^n).

Let

$$t^* = \inf\{t \geq 0 : \chi_1(t) \leq 0\},$$

be the time when $\chi_1(t)$ hits zero, and let

$$\mathcal{T}_1^n = \inf\{i \geq 0 : Y_1^n(i) = 0\},$$

denote the end of Phase 1.

The following theorem establishes the closeness of the suitably rescaled Markov chain and the function χ during Phase 1.

Theorem 11 ([2], Lemma 11, Lemma 12, Section 5.1. and Section 5.2.)**.** *For $t < t^*$ and every $\epsilon > 0$ we have*

$$\mathbb{P}\left(\sup_{s\leq t}\left|\frac{Y^n(\lfloor ns\rfloor)}{n} - \chi(s)\right| > \epsilon\right) \to 0,$$

as $n \to \infty$, where $|\cdot|$ denotes the supremum norm. Also,

$$\mathbb{P}\left(\left|\frac{Y^n(\mathcal{T}_1^n)}{n} - \chi(t^*)\right| > \epsilon\right) \to 0,$$

as $n \to \infty$.

The phase transition occurring around $c = e$ is then readable from the properties of $\chi(t)$. Namely, one can check that for $c \le e$, $\chi_1(t) = 0$ implies $\chi_2(t) = 0$ and $\chi_3(t) = 0$ and so the size of Karp-Sipser core is negligible compared to n. On the other hand, for $c > e$, when $\chi_1(t)$ hits zero both $\chi_2(t)$ and $\chi_3(t)$ are strictly greater than zero, and so we have a Karp-Sipser core of size of order n. Exactly this property, the existence of a giant Karp-Sipser core for $c > e$, enables our refinement of the results from [2] in Chapter 2. Loosely speaking, the main virtue of this property is that the rescaled Markov chain $\frac{Y^n(\lfloor n\cdot \rfloor)}{n}$ evolves smoothly up until the very end of Phase 1, and thus we can easily control the deviations from the expected trajectory.

1.7.5 Phase 2

Let us recall that Phase 2 begins when for the first time there are no pendant vertices to match, and the algorithm matches a uniform random edge. The remaining graph at this point is the Karp-Sipser core of the initial graph. From that time on, the algorithm matches a uniform random pendant vertex if one exists, and otherwise picks a uniform random edge (and matches it if it connects two distinct vertices). Phase 2 ends when the remaining graph is empty. The analysis of Phase 2 is similar in style to the analysis of Phase 1 – one looks at the Markov chain under the fluid limit scaling and proves that it remains sharply concentrated around a certain deterministic function. Our analysis in Chapter 2 does not involve a study of Phase 2, and so we here only highlight the main results from [34] and [2] related to this phase.

For $c > 0$, let K_c^n denote the number of vertices in the Karp-Sipser core. Let also A_c^n be the number of vertices *from the Karp-Sipser core* which remain unmatched.

Theorem 12 ([2], Theorem 2)**.** *For $c < e$, the Karp-Sipser core is a collection of vertex disjoint cycles, with high probability as $n \to \infty$. Thus, the Karp-Sipser algorithm finds a maximum matching with high probability as $n \to \infty$.*

The following theorem is in fact a consequence of Theorem 11, but we state it here in order to have the complete picture as c varies.

Theorem 13. *For $c = e$ and any $\epsilon > 0$, as $n \to \infty$ we have*

$$\mathbb{P}\Big(\frac{K_e^n}{n} > \epsilon\Big) \to 0.$$

Theorem 14 ([2], Theorem 3)**.** *For $c > e$, there are constants $C_1, C_2 > 0$ such that*

$$C_1 n^{1/5}/(\log n)^{75/2} \leq \mathbb{E}[A_c^n] \leq C_2 n^{1/5}(\log n)^{12}.$$

Let us recall that the Karp-Sipser algorithm behaves optimally during Phase 1, in the sense that there is a maximum matching which contains all edges matched during this phase. Thus, results from [34] and [2] prove that for all $c > 0$ the difference between the size of a matching obtained by the algorithm and the size of a maximum matching is not visible on order $\Theta(n)$, with high probability as $n \to \infty$. By Theorem 12, for $c < e$ the Karp-Sipser core is a collection of vertex disjoint cycles with high probability as $n \to \infty$, and hence the algorithm behaves optimally in Phase 2 too. For $c = e$, we have a Karp-Sipser core whose size is negligible compared to n, although we do not know its true order of size. On the other hand, as the consequence of Theorem 11, for $c > e$ we have a Karp-Sipser core of size of order $\Theta(n)$, with high probability as $n \to \infty$. Given the naive heuristics the algorithm uses in Phase 2, this looks like a potentially dangerous situation from the perspective of obtaining a near maximum matching. Recall that during Phase 2 the Karp-Sipser algorithm can match edges that do not belong to any maximum matching. However, Theorem 14 tells us that the difference between the size of a matching obtained by the algorithm and a maximum matching is $\Theta_{\log n}(n^{1/5})$, with high probability as $n \to \infty$. Loosely speaking, this is the consequence of the fact that for the majority of Phase 2, a removal of a matched edge does not create "too many" nodes of degree 1, and so it does not happen too often that a vertex of degree 1 gets isolated due to its only neighbour being matched to someone else (note that each initially non-isolated vertex which the algorithm fails to match, has a unique neighbour prior to the step in which the algorithm isolates it.)

1.7.6 The weak law of large numbers for the maximum matching number

The main result of [34] and [2] is a weak law of large numbers for the maximum matching number of the sparse Erdős-Rényi random graph.

Let us recall that Theorems 12, 13 and 14 give us that during Phase 2 the algorithm fails to match a number of vertices which is negligible compared to n, with high probability as $n \to \infty$. Thus, for all $c > 0$, the difference between the number of nodes the algorithm matches during Phase 2, and the number of nodes in the Karp-Sipser core, is invisible on order $\Theta(n)$. Also, in each step of Phase 1, the algorithm matches one edge. Recalling that $\mathcal{T}_1^n$ denotes the number of steps the algorithm makes during Phase 1, we get that the difference between the size of a matching obtained by the algorithm and $\mathcal{T}_1^n + \frac{1}{2}Y_2^n(\mathcal{T}_1^n)$ is not visible on order $\Theta(n)$, with high probability as $n \to \infty$. Having in mind the discussion from the previous section, this gives us that the difference between the maximum matching number of the random multigraph and $\mathcal{T}_1^n + \frac{1}{2}Y_2^n(\mathcal{T}_1^n)$, is not visible on order $\Theta(n)$ with high probability as $n \to \infty$. As the random multigraph construction produces a simple graph with the probability bounded away from zero as $n \to \infty$ (see Remark 26 from Chapter 2), this enables one to obtain a result about the maximum matching number of Erdős-Rényi random graph $G(n, M = \lfloor \frac{cn}{2} \rfloor)$.

Theorem 15 ([2], Theorem 4)**.** *For $c > 0$, let μ_c^n denote the maximum matching number of the Erdős-Rényi random graph $G(n, M = \lfloor \frac{cn}{2} \rfloor)$. Let γ_1 and γ_2 be the smallest and the largest root, respectively, of the equation $x = c\exp\{-ce^{-x}\}$. For $\epsilon > 0$, we have*

$$\mathbb{P}\left(\left|\frac{\mu_c^n}{n} - \left(1 - \frac{\gamma_2 + \gamma_1 + \gamma_2\gamma_1}{2c}\right)\right| > \epsilon\right) \to 0,$$

as $n \to \infty$.

Remark 16. *Theorem 15 implies that for the maximum matching number $\tilde{\mu}_c^n$ of the Erdős-Rényi random graph $G(n, p = \frac{c}{n})$, $c > 0$ and $\epsilon > 0$, as $n \to \infty$ we have*

$$\mathbb{P}\left(\left|\frac{\tilde{\mu}_c^n}{n} - \left(1 - \frac{\gamma_2 + \gamma_1 + \gamma_2\gamma_1}{2c}\right)\right| > \epsilon\right) \to 0.$$

As we have not presented the detailed analysis of Phase 1 from [2], the relationship between the deterministic approximation χ of the rescaled process $\frac{Y^n(\lfloor n\cdot \rfloor)}{n}$ in Phase 1, and the roots of the equation $x = c\exp\{-ce^{-x}\}$ is not immediately visible. However, it turns out that we have the following relationship (see [2], (48))

$$\vartheta(t^*) = \gamma_2 - \gamma_1,$$

and in particular $\chi_3(t^*) = \frac{1}{2c}(\gamma_2 - \gamma_1)^2$ (recall that t^* denotes the time when χ_1 first hits zero), and thus having in mind Theorem 11 we see that the distance between the two roots of the equation determines the number of edges in the Karp-Sipser core. Let us also emphasize that for $c \leq e$, we have $\gamma_1 = \gamma_2$, while for $c > e$, we have $\gamma_1 < 1 < \gamma_2$, and so the phase transition occurring around $c = e$ could be seen as the consequence of the properties of the equation $x = c\exp\{-ce^{-x}\}$.

1.8 Outline of the thesis

The thesis consists of two separate pieces of research. The problems studied in both Chapter 2 and Chapter 3 are motivated by questions arising from [2].

In Chapter 2, we prove the central limit theorem for the maximum matching number of Erdős-Rényi random graphs $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ and $G(n, p = \frac{c}{n})$ for $c > e$.

The Chapter 3 studies the fire propagation on the random multigraph. This is the joint work with my supervisor Professor Christina Goldschmidt.

1.8.1 Chapter 2

As explained in Section 1.7.6, [34] and [2] provide us with a weak law of large numbers for the maximum matching number of both sparse Erdős-Rényi random graph models. Results of Pittel [45] imply that for $c < 1$ we have a central limit theorem for the maximum matching number. Loosely speaking, this is a consequence of the fact that for $c < 1$ the graph mainly consists of a forest of trees, with high probability as $n \to \infty$, and a result of Pittel [45] about a central limit theorem for the number of nodes adjacent to trees of different shapes. In [2], the authors conjecture a central limit theorem for the maximum matching number for all $c > 0$. Our main result in Chapter 2 is the following central limit theorem for the maximum matching number of both $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ and $G(n, p = \frac{c}{n})$ for $c > e$.

Theorem 17. *Let $c > e$. Let μ_c^n and $\tilde{\mu}_c^n$ denote the size of a maximum matching of the Erdős-Rényi random graph $G(n, M = \lfloor cn/2 \rfloor)$ and $G(n, p = \frac{c}{n})$, respectively. Let γ_1 and γ_2 be the smallest and the largest root, respectively, of the equation $x = c\exp\{-ce^{-x}\}$.*

i) As $n \to \infty$, the random variable

$$\frac{\mu_c^n - n\Big(1 - \frac{\gamma_2+\gamma_1+\gamma_2\gamma_1}{2c}\Big)}{\sqrt{n}}$$

converges in distribution to a normal distribution with mean 0, and the variance $\sigma(c) > 0$ given by (2.54) in Chapter 2.

ii) As $n \to \infty$, the random variable

$$\frac{\tilde{\mu}_c^n - n\Big(1 - \frac{\gamma_2+\gamma_1+\gamma_2\gamma_1}{2c}\Big)}{\sqrt{n}}$$

converges in distribution to a normal distribution with mean 0, and the variance $\tilde{\sigma}(c) > 0$ given by (2.55) in Chapter 2.

The case $c > e$ is particularly convenient for the analysis, as the Karp-Sipser core has the size of order $\Theta(n)$ with high probability as $n \to \infty$. This guarantees that the Markov chain $\frac{Y^n(\lfloor n \cdot \rfloor)}{n}$ evolves smoothly during Phase 1, and we exploit this property in our analysis. As a consequence of Theorem 14, the difference between the maximum matching number and $\mathcal{T}_1^n + \frac{1}{2}Y_2^n(\mathcal{T}_1^n)$ is not visible on order $\Omega(\sqrt{n})$, and so it suffices to follow the Markov chain up until the end of Phase 1 in order to obtain the central limit theorem for the maximum matching number. In order to fill the gap and prove the central limit theorem for $1 \leq c \leq e$, one would have to overcome the following obstacles. For $1 \leq c < e$, we would have to control the evolution of the Markov chain close to the end of Phase 1, when there is no upper bound on the derivative of the deterministic approximation and the process thus may exhibit abrupt changes. Moreover, for $c = e$, the true order of size of the number of nodes and edges in the Karp-Sipser core is not known, hence there is a possibility that in Phase 2 the algorithm fails to match a number of vertices which is visible on order $\Omega(\sqrt{n})$. We do not pursue the case $1 \leq c \leq e$ any further here.

1.8.2 Chapter 3

Theorem 14 says that for $c > e$, the expected number of nodes the algorithm fails to match during Phase 2 is of order $\Theta_{\log n}(n^{1/5})$. As conjectured in [2], it is plausible that the number of nodes from the Karp-Sipser core which remain unmatched, rescaled by $n^{1/5}$, converges in distribution. The analysis done in [2] suggests that at some time close to the end of the algorithm the remaining graph has $\Theta_{\log n}(n^{3/5})$ nodes of degree 2, $\Theta_{\log n}(n^{2/5})$ nodes of degree 3 and $\Theta_{\log n}(n^{1/5})$ nodes of degree at least 4 (see [2], Section 6.1, (172)), and that from this time on the number of nodes the algorithm fails to match is of order $\Theta_{\log n}(n^{1/5})$, with high probability as $n \to \infty$. Thus, it is worth exploring which structural properties of such a graph cause $\Theta_{\log n}(n^{1/5})$ nodes to remain unmatched. There is one more question related to the performance of the Karp-Sipser algorithm on the sparse Erdős-Rényi random graph which directs us to the study of the same model. As already explained, for $c = e$ we do not know the true order of size of the number of nodes and edges in the Karp-Sipser core. The numerical study done by Bauer and Golinelli in [6] provides us with the following conjecture: for $c = e$, the Karp-Sipser core has $\Theta(n^{3/5})$ nodes of degree 2, $\Theta(n^{2/5})$ nodes of degree 3 and $\Theta(n^{1/5})$ nodes of degree at least 4 (see [6], Section 5.3, (7) and (8) and Table 1). Thus, understanding the behaviour of the Karp-Sipser algorithm on such a graph could provide an insight into its performance on the Karp-Sipser core for the critical $c = e$. The problem we study in Chapter 3 is motivated by the question of how the Karp-Sipser algorithm progresses on such a graph, and whether the (suitably rescaled) number of nodes which remain unmatched converges in distribution.

In Chapter 3 we study a fire propagation on a multigraph generated according to the configuration model with N nodes of degree 3 and $\alpha(N)$ nodes of degree 4 where $\alpha(N) = o(N)$ as $N \to \infty$. We assign the edges of this multigraph i.i.d. standard exponential lengths. The fire propagation starts by a uniform random choice of a point from the total length, which we set alight. The edges burn at speed 1. When the fire reaches a node of degree at least 3, it is passed along all adjacent edges with probability $\frac{1}{2}$, and is suppressed otherwise. Note that the burning procedure may cause nodes to decrease their degrees and so we can create nodes of degree 2 on the way. If the fire reaches a node of degree 2, it always gets passed along to the remaining adjacent edge. If the fire goes out before the whole graph is burnt, we again pick a uniform random point and set it alight. Our primary interest is to determine the limiting distribution of the suitably rescaled number of fires we set alight in order to burn the whole graph, and the number of points which get caught by the fire from two different directions.

More precisely, let F^N be the number of fires we set alight in order to burn the whole multigraph, and C^N be the number of points which get caught by fire from two different directions. It turns out that both quantities scale as $\sqrt{N}$ as long as $\alpha(N) = O(\sqrt{N})$ and otherwise as $\alpha(N)$. In order to state the main result of Chapter 3, we introduce the following stochastic process.

Let $(B_t)_{t\geq 0}$ be a standard Brownian motion and for $a \geq 0$ let $(X_t^a, L_t^a)_{0\leq t\leq 1}$ be the unique solution to the stochastic differential equation with reflection determined by

$$dX_t^a = -\frac{2}{3}\frac{X_t^a}{(1-t)}dt + \frac{2a}{3}(1-t)^{1/3}dt + dB_t + dL_t^a, \quad 0 \leq t < 1,$$

where $(L_t^a)_{0\leq t<1}$ is the local time process of X^a at level 0. As we will see in Chapter 3, we can write this solution explicitly as a functional of the Brownian motion.

The main result of Chapter 3 is the following theorem.

Theorem 18. *(i) Suppose that $\alpha(N)/\sqrt{N} \to a$ as $N \to \infty$, where $a \geq 0$. Then, as $N \to \infty$,*

$$\frac{1}{\sqrt{N}}(F^N, C^N) \xrightarrow{d} \left(\frac{1}{2}L_1^a, \int_0^1 \frac{X_s^a}{3(1-s)}ds\right),$$

where the limiting random variables are almost surely finite.

(ii) Suppose that $\alpha(N) >> \sqrt{N}$. Then, as $N \to \infty$,

$$\frac{1}{\alpha(N)}F^N \xrightarrow{p} 0, \quad \frac{1}{\alpha(N)}C^N \xrightarrow{p} \frac{1}{4}.$$

Let us briefly explain the analogy between the performance of the Karp-Sipser algorithm on the graph with $\Theta(n^{3/5})$ nodes of degree 2, $\Theta(n^{2/5})$ nodes of degree 3 and $\Theta(n^{1/5})$ nodes of degree at least 4 and our fire propagation (in Chapter 3 this is done in more detail). On such a graph, we expect Karp-Sipser to start by matching an edge which lives within a long path consisting of degree-2 vertices. This creates two pendant vertices, each at the end of a path consisting of degree-2 vertices and a vertex of a higher degree at the other end. If the length of the path is odd, the Karp-Sipser algorithm will match the degree-2 vertices along the path, but will not remove the higher-degree vertex at the end of the path. On the other hand, if the length of the path is even, the higher degree vertex gets matched and thus removed, and this causes its degree-2 neighbours to become pendant. In this case, the Karp-Sipser algorithm eats away further into the graph. As we typically have a large number of degree-2 vertices along each path, we expect a path to be even or odd, both with probability approximately $\frac{1}{2}$. Let us note that whenever a path gets eaten away by the algorithm from two different ends, there is a chance that a vertex in the path remains unmatched (this again depends on the parity).

It remains to note that the behaviour of the Karp-Sipser algorithm along a path of degree-2 vertices with a pendant vertex in the end, does not involve any randomness. Hence, we can see the evolution of the Karp-Sipser algorithm on such a path as a fire propagation along the length proportional to the number of degree-2 vertices. Then, the number of the nodes Karp-Sipser fails to match corresponds (approximately) to half the number of points which get caught by fire from two different directions in the fire propagation model. This provides a heuristic explanation for the similarity between the Karp-Sipser algorithm on a graph with $\Theta(n^{3/5})$ nodes of degree 2, $\Theta(n^{2/5})$ nodes of degree 3 and $\Theta(n^{1/5})$ nodes of degree at least 4 and the fire propagation on the random multigraph with $N = n^{2/5}$ nodes of degree 3 and $\alpha(N) = \Theta(\sqrt{N})$ nodes of degree 4. Our Theorem 18 implies the convergence in distribution of the number of points which get burnt from two different directions scaled by $\sqrt{N}$. Note that this is in line with the conjecture that the number of nodes which remain unmatched, when the Karp-Sipser is applied on the corresponding graph, scaled by $n^{1/5}$ converges in distribution. We do not claim that our fire propagation is a good approximation of the behaviour of the algorithm on the appropriate graph; we rather use this analogy as a motivation for the study of the problem we consider.

Chapter 2

The central limit theorem for the maximum matching number of the sparse Erdős-Rényi random graph for $c > e$

Abstract

We obtain a central limit theorem for the size of a maximum matching of the Erdős-Rényi random graph models $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ and $G(n, p = \frac{c}{n})$ for $c > e$. Our approach is based on the analysis of the Karp-Sipser algorithm that uses simple heuristics to obtain a matching. We use results from [2] and refine them.

Keywords: random graph, Erdős-Rényi model, matching, maximum matching, maximum matching number, Karp-Sipser algorithm, central limit theorem.

2.1 Introduction

In Chapter 1, we introduced the Karp-Sipser algorithm for obtaining a matching in a graph, and commented on its performance on some classes of sparse random graphs. In particular, we provided a review of the results from Karp and Sipser [34] and Aronson et al [2] which consider the Karp-Sipser algorithm on Erdős-Rényi random graph models $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ and $G(n, p = \frac{c}{n})$. In this chapter, we rely on the results from [2] and further refine them. Before we proceed, we briefly recall main aspects of the study done in [2].

2.1.1 The model

We consider a random multigraph with n nodes and $M = \lfloor \frac{cn}{2} \rfloor$ edges for $c > e$. We generate such a multigraph by adjusting the well known configuration model construction [7, 12, 54, 55, 56]. Firstly, distribute independently $2M$ half-edges among the n vertices with the same probability (i.e. distribute $2M$ balls into n bins). Secondly, choose a uniform random partition of the set of half-edges into M pairs. Thirdly, declare each pair of half-edges in the partition to be an edge in the multigraph. Conditioned on being simple, such a graph has the same distribution as the well known Erdős-Rényi random graph $G(n, M = \lfloor \frac{cn}{2} \rfloor)$. Moreover, the probability of obtaining a simple graph is bounded away from 0 as $n \to \infty$ (see Remark 26 below). Thus, any event that happens with high probability in the random multigraph happens with high probability in the Erdős-Rényi graph too. Carrying over results about convergence in distribution is a more challenging task. Fortunately, it will turn out that our analysis for the random multigraph can be easily adjusted to the Erdős-Rényi random graph. Loosely speaking, this is due to the fact that the total number of half-edges taking part in loops and multiple edges converges in distribution, and so their effect is not visible on the order of $\sqrt{n}$.

2.1.2 Karp-Sipser algorithm

Let us recall that for a given graph, a *matching* is any set of pairwise non-adjacent edges. The size of a matching is given by the number of its edges. A *maximum matching* is a matching with the largest possible size, i.e. such that there is no matching containing more edges. The *maximum matching number* of a graph is the size of its maximum matching. As already explained in Chapter 1, the Karp-Sipser algorithm, introduced in [34], is a greedy algorithm for finding a near-maximum matching in a fixed graph. Let us recall how the algorithm works. Call a vertex of degree 1 a *pendant* vertex. If there is at least one pendant vertex in the graph, choose one uniformly at random and include the edge adjacent to it in the matching. Remove this edge, the pair of vertices that forms it and any other edges adjacent to them. If there are no pendant vertices in the graph, choose one of the existing edges uniformly at random, and include it in the matching. Remove the chosen edge together with the two vertices that form it, as well as any other edges adjacent to those vertices. Repeat the procedure on the remaining graph, until there are no more edges.

As already explained in Chapter 1, whenever there exists a pendant vertex in a graph, that vertex and its neighbour are included in *some* maximum matching. So the Karp-Sipser algorithm *never* makes a "mistake" in including such an edge in its matching. On the other hand, in the other type of move (picking a uniform edge and including it in the matching) it is possible that it includes an edge which would *not* be in any maximum matching. We will refer to the steps the algorithm makes until the moment when the number of nodes of degree 1 drops to 0 for the first time as *Phase 1*. Also, we will refer to the graph remaining in the end of Phase 1 as *the Karp-Sipser core*. Since the algorithm performs optimally as long as we have some pendant vertices, it is clear that the distance of the final matching from a maximum matching is determined by how well the algorithm performs when run on the Karp-Sipser core.

2.1.3 The heuristics for the CLT for the maximum matching number

The performance of the Karp-Sipser algorithm on the sparse Erdős-Rényi random graph $G(n, M = \lfloor \frac{cn}{2} \rfloor)$, $c > 0$ is studied in [2] and [34]. The main result of [34] is a law of large numbers for the size of a matching obtained via Karp-Sipser algorithm. In [2] the results from [34] are further refined. Let us comment on how successful is the algorithm in obtaining a near-maximum matching as c varies. Theorem 2 of [2] states that for $c < e$, the Karp-Sipser core consists of a collection of vertex disjoint cycles with high probability as $n \to \infty$, and so in this case the Karp-Sipser algorithm finds a maximum matching with high probability as $n \to \infty$. Theorem 3 of [2] says that for $c > e$, the expected number of the nodes *of the Karp-Sipser core* that remain unmatched at the end of the algorithm is bounded below by a quantity of order $n^{1/5}/(\log n)^{75/2}$ and bounded above by a quantity of order $n^{1/5}(\log n)^{12}$. Hence, for $c > e$, the expected difference between a maximum matching and that obtained by the Karp-Sipser algorithm is bounded above by a quantity of order $n^{1/5}(\log n)^{12}$, as $n \to \infty$. For $c = e$, the size of the Karp-Sipser core is negligible compared to n, with high probability as $n \to \infty$ (see [34], Theorem 3), however the true order of its size is not known.

For $c < 1$, the graph $G(n, M = \lfloor cn/2 \rfloor)$ consists of a forest of trees plus possibly a few unicyclic components of a total size bounded in probability, with high probability as $n \to \infty$ (see Erdős and Rényi [20]). Pittel [45] proves that the number of trees of a given size fluctuates around its expected value on order $\sqrt{n}$. Even more, it is proved that the vector consisting of such fluctuations for all possible sizes of trees, rescaled by $\sqrt{n}$, converges to a Gaussian limit. As noted in [2], a central limit theorem for the maximum matching number of $G(n, M = \lfloor cn/2 \rfloor)$ is a consequence of this result.

In [2], it is conjectured that a central limit theorem also holds in the case $c \geq 1$; this is left as an open problem. Let us provide a heuristic argument why is such a conjecture reasonable. For $c > 1$, the graph consists of a unique giant component which contains $\Theta(n)$ vertices, a forest of trees which also contains $\Theta(n)$ vertices, and a number of unicyclic components of total size bounded in probability, as $n \to \infty$ (see, for example Janson et al [30], Chapter 5). Having in mind that the supercritical Erdős-Rényi random graph with the giant component deleted has (almost) the same distribution as some subcritical Erdős-Rényi graph ([30], Chapter 5, Section 5.6), we expect to have a central limit theorem for the number of edges matched in the forest of trees. We now turn to the giant component. Theorem 3 of [45] states that the giant component consists of a giant 2-core (this is the intersection of the 2-core and the giant component; recall that the Karp-Sipser core is a subgraph of the 2-core) and a *mantle* of trees, both of them containing $\Theta(n)$ vertices, with high probability as $n \to \infty$. Moreover, Theorem 6 of Pittel and Wormald [47] gives us that the joint distribution of the fluctuations of the number of the vertices of the giant 2-core and the number of the vertices of the mantle around their expected values, rescaled by $\sqrt{n}$, tends to Gaussian as $n \to \infty$. Thus, the results from [45] suggest that it is reasonable to expect a central limit theorem for the matching number of the mantle "chopped off" the giant 2-core. Finally, it remains to see what happens on the giant 2-core. Theorem 2 of Frieze and Pittel [23] gives us that the giant 2-core, conditioned on the number of its vertices and edges, possesses an almost perfect matching (in the sense that at most one vertex stays unmatched), with high probability as $n \to \infty$. This, together with the central limit theorem result for the size of the 2-core, suggests a central limit theorem for the number of matched edges in the giant 2-core. Finally, as the maximum matching number of each of the three subgraphs considered separately is expected to obey a central limit theorem, it seems reasonable to conjecture a central limit theorem for the maximum matching number of a graph as a whole.

2.1.4 Our analysis

The main result we prove in this chapter is a central limit theorem for the size of the maximum matching for both Erdős-Rényi random graphs $G(n, M = \lfloor cn/2 \rfloor)$ and $G(n, p = \frac{c}{n})$ for $c > e$.

Theorem 19. *Let $c > e$. Let μ_c^n and $\tilde{\mu}_c^n$ denote the size of a maximum matching of the Erdős-Rényi random graph $G(n, M = \lfloor cn/2 \rfloor)$ and $G(n, p = \frac{c}{n})$, respectively. Let γ_1 and γ_2 be the smallest and the largest root, respectively, of the equation $x = c\exp\{-ce^{-x}\}$.*

i) As $n \to \infty$, the random variable

$$\frac{\mu_c^n - n\Big(1 - \frac{\gamma_2 + \gamma_1 + \gamma_2\gamma_1}{2c}\Big)}{\sqrt{n}}$$

converges in distribution to a normal distribution with mean 0, and variance $\sigma(c) > 0$ given by (2.54) below.

ii) As $n \to \infty$, the random variable

$$\frac{\tilde{\mu}_c^n - n\Big(1 - \frac{\gamma_2 + \gamma_1 + \gamma_2\gamma_1}{2c}\Big)}{\sqrt{n}}$$

converges in distribution to a normal distribution with mean 0*, and variance* $\tilde{\sigma}(c) > 0$ *given by (2.55) below.*

Our analysis relies on results obtained in [2] and extends them. As already explained, as long as one is interested in the size of the maximum matching on order $\Omega(\sqrt{n})$, it is enough to analyse the algorithm up until the end of Phase 1, as "almost all" vertices from the Karp-Sipser core get matched. Recall that, for $c > e$, the Karp-Sipser core contains $\Theta(n)$ nodes, with high probability as $n \to \infty$. This property guarantees that the degree distribution of the graph evolves smoothly as we run the algorithm, and this enables a straightforward analysis. Our main tool is the framework on convergence of density dependent Markov processes from Ethier and Kurtz [21], Chapter 11. We emphasize that Theorem 19 deals with the case $c > e$ only. As we have explained earlier, the central limit theorem for a matching number for $c < 1$ is the consequence of the results from [45]. Note that this leaves a gap for $1 \leq c \leq e$.

2.2 The Markov chain

Let us start with a random multigraph on n vertices and $M = \lfloor \frac{cn}{2} \rfloor$ edges. In order to be able to use results from [21] it is convenient to study the evolution of the Karp-Sipser algorithm in continous time. Thus, let the time between two steps of the Karp-Sipser algorithm have exponential distribution with rate 1. For $t \geq 0$, let $X_1^n(t)$ and $X_2^n(t)$ be the number of nodes of degree 1 and at least 2 respectively at time t. Let $X_3^n(t)$ be the number of edges in the graph at time t, and $X_4^n(t)$ the number of steps the algorithm made up until time t. We are interested in the evolution of the process

$$X^n(t) = (X_1^n(t), X_2^n(t), X_3^n(t), X_4^n(t))^T$$

during Phase 1 of the Karp-Sipser algorithm, that is up until time

$$\tau^n = \inf\{t \geq 0 : X_1^n(t) = 0\}, \tag{2.1}$$

when the number of pendant nodes drops down to zero for the first time. The analysis in [2] relies on the study of this process (without the fourth component) in a discrete time. The process X^n is a Markov chain with state space

$$\{(y_1, y_2, y_3, y_4) : y_1, y_2, y_3, y_4 \in \mathbb{N} \cup \{0\}, y_1 + 2y_2 \leq 2y_3, y_1 + y_2 \leq n, y_3 \leq M\}.$$

Note that $y_1 + 2y_2 \leq 2y_3$ holds since each pendant node has an adjacent edge and each node of degree at least two has at least two adjacent edges. The Markov property is the consequence of Lemma 5 from Chapter 1. In the following, we let $(\hat{\mathfrak{F}}^n_t)_{t\geq 0}$ denote the natural filtration of the Markov chain X^n.

2.2.1 The fluid limit approximation in Phase 1

Here we highlight the fluid limit result from [2], which says that the suitably rescaled Markov chain remains sharply concentrated around a certain deterministic function all the way up until the end of Phase 1 (recall Chapter 1, Section 1.3 for the brief introduction to the fluid limit method). In this section we let the relationship between the process and the deterministic approximation stay somewhat mysterious and we reveal its nature later. We pursue this order in presentation as the fluid limit result enables us to estimate the errors in the approximation of rates of transitions with the precision required for our analysis.

For $u \geq 0$, let

$$f(u) = e^u - u - 1,$$
$$p(u) = e^{-u} f(u),$$

and let $\beta(u)$ be the unique solution to

$$\beta(u) e^{c\beta(u)} = e^u. \tag{2.2}$$

Let the function $\vartheta(t)$, $\vartheta : \mathbb{R}_+ \to \mathbb{R}$ be given (implicitly) by

$$t = \frac{1}{c}\Big(c(1 - \beta(\vartheta(t))) - \frac{1}{2}\log^2 \beta(\vartheta(t))\Big),$$

and let $\chi(t) = (\chi_1(t), \chi_2(t), \chi_3(t), \chi_4(t))^T$, $\chi : \mathbb{R}_+ \to \mathbb{R}^4$ be given by

$$\begin{aligned}
\chi_1(t) &= \frac{1}{c}\Big(\vartheta^2(t) - \vartheta(t) \cdot c \cdot \beta(\vartheta(t))(1 - e^{-\vartheta(t)})\Big), \\
\chi_2(t) &= p(\vartheta(t))\beta(\vartheta(t)), \\
\chi_3(t) &= \frac{1}{2c}\vartheta^2(t), \\
\chi_4(t) &= t.
\end{aligned} \tag{2.3}$$

Let

$$t^* = \inf\{t \geq 0 : \chi_1(t) = 0\}, \tag{2.4}$$

be the time when $\chi_1(t)$ hits zero. The following theorem establishes the closeness of the suitably rescaled Markov chain and the function $\chi(t)$ during Phase 1.

Theorem 20 ([2], Lemma 11, Lemma 12, Section 5.1. and Section 5.2.)**.** *For $t < t^*$ and every $\epsilon > 0$ we have*

$$\mathbb{P}\left(\sup_{u \leq t}\left|\frac{X^n(\lfloor nu \rfloor)}{n} - \chi(u)\right| > \epsilon\right) \to 0, \tag{2.5}$$

as $n \to \infty$, where $|\cdot|$ denotes the supremum norm. Also,

$$\mathbb{P}\left(\left|\frac{X^n(\tau^n)}{n} - \chi(t^*)\right| > \epsilon\right) \to 0, \tag{2.6}$$

as $n \to \infty$.

Let us emphasize that $\chi(0) = (ce^{-c}, 1 - e^{-c} - ce^{-c}, c/2, 0)$, and so $\chi_1(0)$, $\chi_2(0)$ and $\chi_3(0)$ represent distributional limits of the number of nodes of degree 1, the number of nodes of degree least 2, and the total number of edges in the initial graph, after the weak law of large numbers rescaling.

The phase transition occuring around $c = e$ is readable from the properties of $\chi(t)$. Namely, one can check that for $c \leq e$, $\chi_1(t) = 0$ implies $\chi_2(t) = 0$ and $\chi_3(t) = 0$ and so Theorem 20 gives us that the size of Karp-Sipser core is negligible compared to n. On the other hand, for $c > e$, when $\chi_1(t)$ hits zero, both $\chi_2(t)$ and $\chi_3(t)$ are strictly greater than zero, and so we have a giant Karp-Sipser core. We will exploit this property in our analysis. Let

$$\zeta^n = \inf\left\{t \geq 0 : X_2^n(t) \leq \frac{\chi_2(t^*)}{2}n \text{ or } X_3^n(t) \leq \frac{\chi_3(t^*)}{2}n\right\}. \tag{2.7}$$

Remark 21. *As a consequence of Theorem 20, for $c > e$ as $n \to \infty$ we have*

$$\mathbb{P}(\tau^n \leq \zeta^n) \to 1.$$

Let us note that Remark 21 guarantees that for $c > e$ the Karp-Sipser core contains $\Theta(n)$ nodes and $\Theta(n)$ edges with high probability as $n \to \infty$.

2.2.2 The degree distribution

In Chapter 1, we presented findings from [2] which enable one to control the degree sequence of the remaining graph as the algorithm evolves. Here we briefly recall these results and state them in a form convenient for the refinement of the analysis of Phase 1 for $c > e$. Recall that our Markov chain keeps track of the number of nodes of degree at least 2, but we do not know their exact degrees. However, in order to follow the evolution of the graph, it will be necessary to control the degree sequence to some extent.

Before we proceed, for $x = (x_1, x_2, x_3, x_4)^T \in (\mathbb{R}_+ \cup \{0\})^4$ such that $x_2 > 0$, let z_x be the unique solution to

$$\frac{z_x(e^{z_x} - 1)}{f(z_x)} = \frac{2x_3 - x_1}{x_2}, \tag{2.8}$$

where as before, $f(u) = e^u - u - 1$, $u > 0$.

In the sequel, we will refer to a Poisson random variable with the parameter $\lambda > 0$ conditioned on being greater than or equal to 2 as a *truncated* Poisson random variable with the parameter λ. Conditional on $\hat{\mathcal{F}}_t^n$, the vector of degrees of the $X_2^n(t)$ nodes of degree at least 2 has the same distribution as a vector of $X_2^n(t)$ i.i.d. truncated Poisson random variables with the parameter $\lambda > 0$, conditioned on their sum being equal to $2X_3^n(t) - X_1^n(t)$ ([2], Lemma 4 or see Chapter 1, Lemma 6). Let us emphasize that this holds for any $\lambda > 0$. However, a particularly convenient choice of the parameter will guarantee the sharp approximation of the degree distribution. In order to obtain such an approximation, let us define an auxiliary process Z^n by $Z^n(t) = z_{X^n(t)}$, that is $Z^n(t)$ satisfies

$$\frac{Z^n(t)(e^{Z^n(t)}-1)}{f(Z^n(t))} = \frac{2X_3^n(t) - X_1^n(t)}{X_2^n(t)}, \quad 0 \le t \le \zeta^n. \tag{2.9}$$

If we condition on $\hat{\mathfrak{F}}_t^n$, the left hand side of (2.9) represents the expectation of a Poisson random variable with mean $Z^n(t)$ conditioned on taking value at least 2. The right hand side of (2.9) is the average degree of a node of degree at least 2 in our multigraph. As we will see shortly, the process Z^n plays a key role in approximating the degree distribution.

Before we proceed, let us note that the auxiliary function $\vartheta(t)$ which shows up in the definition (2.3) of the fluid limit approximation $\chi(t)$ represents an idealized trajectory of the process $Z^n(t)$. Indeed, one can check that

$$\frac{\vartheta(t)\cdot(e^{\vartheta(t)}-1)}{f(\vartheta(t))} = \frac{2\chi_3(t) - \chi_1(t)}{\chi_2(t)}, \tag{2.10}$$

or equivalently $\vartheta(t) = z_{\chi(t)}$. Now we are ready to provide the approximation of the degree distribution as long as the remaining graph is large enough.

Lemma 22 ([2], Lemma 5)**.** *Let $D^n(t)$ be the degree of an arbitrary node of degree at least 2 at time t. For $t \le \tau^n \wedge \zeta^n$ and $2 \le k \le \log n$, we have*

$$\mathbb{P}\Big(D^n(t) = k | \hat{\mathfrak{F}}_t^n\Big) = \frac{(Z^n(t))^k}{k! f(Z^n(t))} + R_k^n(t),$$

where

$$\frac{n}{\log^3(n)} \sup_{0\le t\le \tau^n\wedge\zeta^n} \sup_{2\le k\le \log n} |R_k^n(t)| \xrightarrow{p} 0,$$

as $n \to \infty$. Furthermore, letting $D_1^n(t)$ and $D_2^n(t)$ be the degrees of an arbitrary pair of nodes of degree at least 2, for $t \le \tau^n \wedge \zeta^n$ and $2 \le k_1, k_2 \le \log n$, we have

$$\mathbb{P}\Big(D_1^n(t) = k_1, D_2^n(t) = k_2 | \hat{\mathfrak{F}}_t^n\Big) = \frac{(Z^n(t))^{k_1}}{k_1! f((Z^n(t)))} \frac{(Z^n(t))^{k_2}}{k_2! f(Z^n(t))} + T_{k_1,k_2}^n(t),$$

where

$$\frac{n}{\log^3(n)} \sup_{0\le t\le \tau^n\wedge\zeta^n} \sup_{2\le k_1,k_2\le \log n} \Big|T_{k_1,k_2}^n(t)\Big| \xrightarrow{p} 0,$$

as $n \to \infty$*.*

Sketch of the proof. For $t \le \tau^n \wedge \zeta^n$ we have $Z^n(t)X_2^n(t) \ge \frac{\vartheta(t^*)}{2} \cdot \frac{\chi_2(t^*)}{2} n$, and so this is the special case of Lemma 10 from Chapter 1 for the choice of $h(n) = \frac{\vartheta(t^*)}{2} \cdot \frac{\chi_2(t^*)}{2} n$ and $g(n) = \log n$ there.

In order to emphasize the role of the auxiliary process Z^n, we briefly recall the main idea of the proof. Conditional on $\hat{\mathfrak{F}}_t^n$, let $Q_1, \ldots, Q_{X_2^n(t)}$ be $X_2^n(t)$ independent truncated Poisson random variables with the parameter $Z^n(t)$. We have

$$\begin{aligned}
\mathbb{P}\Big(D^n(t) = k | \hat{\mathfrak{F}}_t^n\Big) &= \mathbb{P}\left(Q_1 = k \middle| \sum_{j=1}^{X_2^n(t)} Q_j = 2X_3^n(t) - X_1^n(t), \hat{\mathfrak{F}}_t^n\right) \\
&= \frac{\mathbb{P}\Big(Q_1 = k, \sum_{j=2}^{X_2^n(t)} Q_j = 2X_3^n(t) - X_1^n(t) - k \Big| \hat{\mathfrak{F}}_t^n\Big)}{\mathbb{P}\Big(\sum_{j=1}^{X_2^n(t)} Q_j = 2X_3^n(t) - X_1^n(t) \Big| \hat{\mathfrak{F}}_t^n\Big)} \\
&= \frac{(Z^n(t))^k}{k! f(Z^n(t))} \cdot \frac{\mathbb{P}\Big(\sum_{j=2}^{X_2^n(t)} Q_j = 2X_3^n(t) - X_1^n(t) - k \Big| \hat{\mathfrak{F}}_t^n\Big)}{\mathbb{P}\Big(\sum_{j=1}^{X_2^n(t)} Q_j = 2X_3^n(t) - X_1^n(t) \Big| \hat{\mathfrak{F}}_t^n\Big)}.
\end{aligned}$$

At this point, the specific choice of the $Z^n(t)$ as the parameter of the truncated Poisson distribution eases the approximation of the degree distribution. Having in mind (2.9), we can use the local limit theorem in order to approximate the probabilities on the right hand side of the last equality (see (5) and (6) in [2], or Lemma 9 in Chapter 1), and these sharp estimates lead to the asymptotics for the errors. □

2.2.3 Some useful properties of the initial multigraph

Here we collect certain properties of the initial multigraph which will be useful in the forthcoming analysis.

Lemma 23. *Let* $\mathcal{A}^n$ *denote the event that the maximum degree of the initial multigraph is at most* $\log n$*. As* $n \to \infty$*, we have*

$$\mathbb{P}(\mathcal{A}^n) \to 1.$$

Proof. This is a consequence of the well known result that the maximum load in a balls-in-bins problem is of order $\Theta\Big(\frac{\log n}{\log \log n}\Big)$ with high probability as $n \to \infty$ (see Mitzenmacher and Upfal [40], Chapter 5). □

Next, we will find it useful to know the distribution of the total number of loops and multiple edges in a multigraph. For a vertex $v \in \{1, \dots, n\}$, let ς_v^n be the number of loops adjacent to it. Also, for a pair $\{v, u\}$ of distinct vertices, let ς_{vu}^n be the number of edges between v and u. Let

$$\Lambda^n = \sum_{v=1}^{n} \varsigma_v^n + \frac{1}{2} \sum_{v=1}^{n} \sum_{u \neq v} \binom{\varsigma_{vu}^n}{2}$$

denote the total number of loops and pairs of multiple edges in the multigraph. We prove the following lemma.

Lemma 24. *Let us consider the random multigraph with n nodes and $M = \lfloor \frac{cn}{2} \rfloor$ edges, for $c > 0$. As $n \to \infty$, the random variable Λ^n converges in distribution to a Poisson with parameter $\frac{c}{2} + \frac{c^2}{4}$. Moreover, all moments of Λ^n converge to appropriate moments of the limiting distribution.*

The following proposition will be a useful tool in the proof. Let us emphasize that it treats a wider class of models than the one we consider, as it allows any degree distribution with certain properties, and the degree distribution of our multigraph is multinomial with $M = 2\lfloor \frac{cn}{2} \rfloor$ trials and n categories with equal chances for success.

Proposition 25 (based on Proposition 7.13 from van der Hofstad [52], Chapter 7)**.** *Consider a random multigraph with n nodes constructed by a configuration model in the following way. Firstly, we choose a degree sequence $\delta^n = (\delta_1, \dots, \delta_n)$ from some given distribution, and then pick a uniform random configuration with that degree sequence. For the degree Δ^n of a uniform random node, and some random variable Δ, let us assume that as $n \to \infty$ we have*

$$\mathbb{P}(\Delta^n = k | \delta^n) \xrightarrow{p} \mathbb{P}(\Delta = k), \tag{2.11}$$

for every $k \in \mathbb{Z}_+$. Also, assume that as $n \to \infty$ we have

$$\mathbb{E}[\Delta^n | \delta^n] \xrightarrow{p} \mathbb{E}[\Delta], \tag{2.12}$$

$$\mathbb{E}\Big[(\Delta^n)^2 | \delta^n\Big] \xrightarrow{p} \mathbb{E}\Big[\Delta^2\Big]. \tag{2.13}$$

Then, the total number of loops and pairs of multiple edges Λ^n converges in distribution to a Poisson random variable with the parameter $\frac{\nu}{2} + \frac{\nu^2}{4}$, where

$$\nu = \frac{\mathbb{E}[\Delta(\Delta - 1)]}{\mathbb{E}[\Delta]}.$$

Moreover, all moments of Λ^n converge to the appropriate moments of the limiting distribution.

Proof. This proposition is a direct consequence of Proposition 7.13 from van der Hofstad [52], Chapter 7. The Proposition 7.13 in [52] provides a convergence in distribution in the case of a *fixed* degree sequence which satisfies certain conditions. However, Remark 7.9 there provides us with conditions on the degree distribution which enable carrying over the results to the case of a random degree sequence (our conditions (2.11), (2.12) and (2.13) are conditions (7.2.13) and (7.2.15) there). Moreover, although the convergence of moments is not explicitly stated in the statement of Proposition 7.13 from [52], it is proved there. Namely, let

$$\Lambda_1^n = \sum_{i=1}^{n} \varsigma_i^n \quad \text{and} \quad \Lambda_2^n = \frac{1}{2} \sum_{i=1}^{n} \sum_{j \neq i} \binom{\varsigma_{ij}^n}{2},$$

be the number of self-loops and pairs of multiple edges, respectively. By (7.4.3) in the proof of Proposition 7.13 in [52], for any $s, r \geq 0$ as $n \to \infty$ we have

$$\mathbb{E}[(\Lambda_1^n)_s (\Lambda_2^n)_r] \to \left(\frac{\nu}{2}\right)^s \left(\frac{\nu^2}{4}\right)^r,$$

where $(u)_r = u(u-1)\cdots(u-r+1)$ denotes the rth factorial moment. This implies the convergence of moments of Λ^n to the moments of the limiting distribution. □

Proof of Lemma 24. As already explained, the degree sequence of our multigraph has multinomial distribution with $M = 2\lfloor \frac{cn}{2} \rfloor$ trials and n categories with equal chances for success. Let V_k^n denote the number of nodes of degree k. Let D be a Poisson random variable with the parameter c. As the consequence of Theorem 4.1 from Janson [27], for any $k \geq 0$ as $n \to \infty$ we have

$$\frac{V_k^n}{n} \xrightarrow{p} \mathbb{P}(D = k),$$

as well as

$$\sum_{k \geq 0} \frac{kV_k^n}{n} \xrightarrow{p} \mathbb{E}[D], \quad \sum_{k \geq 0} \frac{k^2 V_k^n}{n} \xrightarrow{p} \mathbb{E}\Big[D^2\Big].$$

It remains to note that

$$\frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]} = c,$$

and Proposition 25 completes the proof. $\square$

Remark 26. *As a consequence of Lemma 24, the probability that a random multigraph with n vertices and $M = \lfloor \frac{cn}{2} \rfloor$ edges is simple tends to $e^{-c/2-c^2/4}$ as $n \to \infty$.*

The following lemma tells us that in the neighbourhood of a fixed vertex, we typically do not see loops or multiple edges.

Lemma 27. *For an arbitrary node in the initial multigraph, let Ξ^n be the sum of the number of multiple edges adjacent to that node and the number of loops adjacent to that node or any of its neighbours. For n large enough, we have*

$$\mathbb{P}(\Xi^n > 0, \mathcal{A}^n) \leq \frac{(c + c^2)\log^2 n}{n}.$$

Proof. Firstly, let us give an upper bound on the expectation of $\Xi^n \cdot \mathbb{1}_{\mathcal{A}^n}$. By Lemma 24, for n large enough we have $\mathbb{E}[\Lambda^n] \leq c + c^2$ and so expected number of pairs of multiple edges and loops adjacent to a fixed node is at most $\frac{c+c^2}{n}$. On the event $\mathcal{A}^n$, a fixed node has at most $\log n$ neighbours and so we have $\mathbb{E}[\Xi^n \cdot \mathbb{1}_{\mathcal{A}^n}] \leq \frac{(c+c^2)\log n}{n}$. Markov's inequality gives us $\mathbb{P}\Big(\Xi^n > \frac{1}{\log n}, \mathcal{A}^n\Big) \leq \frac{(c+c^2)\log^2 n}{n}$, and so having in mind that Ξ^n takes only integer values and that n is larger than 3, the proof is completed. $\square$

2.2.4 The rates of the Markov chain in Phase 1

The analysis from [2] relies on a counting argument in order to determine the expected size of a jump of the Markov chain. In order to refine the results from [2], we need a precise estimate of the rates of the Markov chain. We achieve this by using a consequence of Theorem 20 highlighted in Remark 21, which states that up until the end of Phase 1 a positive fraction of the initial number of nodes has degree at least 2. This property guarantees that all the way up until the end of Phase 1 we can use Lemma 22 to give a precise estimate on the degree distribution of the remaining graph. Let us recall that in Phase 1 the algorithm always chooses a pendant *vertex* uniformly at random and adds the corresponding edge to the matching. We refer to that chosen vertex as the leaf. Also, we refer to its unique neighbour as the partner. Recall that all edges adjacent to a matched vertex are removed from the graph. Thus, the partner rips off all the edges adjacent to it, causing its neighbours to change their degrees. This guides the Markov chain to its new state. Let $(Y_1^n(i), Y_2^n(i), Y_3^n(i))_{i\geq 0}$ denote the jump chain of $(X_1^n(t), X_2^n(t), X_3^n(t))_{t\geq 0}$. For the steps the algorithm makes in Phase 1, that is for $0 \leq i \leq X_4^n(\tau^n)$, let $I^n(i)$ be the indicator that the partner has degree at least 2. Before we proceed, let

$$\mathcal{K} = \{(k_1, k_2, k_3) \in \mathbb{Z}^3 : k_1 \geq 1, k_2, k_3 \geq 0, k_1 + k_2 + k_3 \geq 2\}. \tag{2.14}$$

Let us note that for $I^n(i) = 1$ the set of possible triples of numbers of the partner's neighbours of degree 1, 2 and at least 3, respectively, belongs to $\mathcal{K}$. That is the case as there is an edge between the leaf and the partner, and for $I^n(i) = 1$ the partner has at least one more adjacent edge. Having in mind that the maximum degree of the multigraph is at most $\log n$ with high probability, let us for each n introduce $\mathcal{K}^n$ by

$$\mathcal{K}^n = \{(k_1, k_2, k_3) \in \mathbb{Z}^3 : k_1 \geq 1, k_2, k_3 \geq 0, \log n \geq k_1 + k_2 + k_3 \geq 2\}. \tag{2.15}$$

It will be convenient to denote by $(\mathcal{F}_i^n)_{i\geq 0}$ the filtration of the jump chain of our process.

The purpose of this section is to prove the following theorem.

Theorem 28. *We work on the event $\mathcal{A}^n$, i.e. the initial graph has maximum degree at most* $\log n$. *Suppose that for some $0 \le i \le X_4^n(\tau^n \wedge \zeta^n)$, we have $(Y_1^n(i), Y_2^n(i), Y_3^n(i)) = (y_1, y_2, y_3)$ where $y_1, y_2, y_3 > 0$, and let z denote the state of the jump chain of Z^n after i jumps. Then, for each triple $(k_1, k_2, k_3) \in \mathcal{K}^n$, we have*

$$\mathbb{P}((Y_1^n(i+1), Y_2^n(i+1), Y_3^n(i+1)) = (y_1 - k_1 + k_2, y_2 - 1 - k_2, y_3 - k_1 - k_2 - k_3)|\mathcal{F}_i^n)$$
$$= \frac{y_2}{2y_3}\frac{z}{f(z)}\left(\frac{y_1 z}{2y_3}\right)^{k_1-1}\left(\frac{y_2 z^3}{2y_3 f(z)}\right)^{k_2}\left(\frac{y_2 z^2}{2y_3}\right)^{k_3}\frac{1}{(k_1-1)!k_2!k_3!} + E^n_{k_1,k_2,k_3}(i)$$

where

$$\frac{n}{\log^3 n}\sup_{0\le i\le X_4^n(\tau^n\wedge\zeta^n)}\ \sup_{(k_1,k_2,k_3)\in\mathcal{K}^n}|E^n_{k_1,k_2,k_3}(i)| \xrightarrow{p} 0,$$

as $n \to \infty$, and

$$\mathbb{P}((Y_1^n(i+1), Y_2^n(i+1), Y_3^n(i+1)) = (y_1 - 2, y_2, y_3 - 1)|\mathcal{F}_i^n) = \frac{y_1}{2y_3} + B^n(i),$$

where $\sup_{0\le i\le X_4^n(\tau^n\wedge\zeta^n)} n \cdot |B^n(i)|$ *is bounded with high probability as $n \to \infty$.*

We start with the following lemma.

Lemma 29. *We work on the event $\mathcal{A}^n$, i.e. the initial graph has maximum degree at most* $\log n$. *Suppose that for some $0 \le i \le X_4^n(\tau^n)$, we have $(Y_1^n(i), Y_2^n(i), Y_3^n(i)) = (y_1, y_2, y_3)$ where $y_1, y_2, y_3 > 0$. For the leaf chosen to be matched in the $(i+1)$st step, let $\alpha(i)$, $\kappa(i)$ and $\omega(i)$ be the number of half-edges adjacent to its partner which lead to a node of degree 1, 2 and at least 3, respectively.*

i) If $I^n(i) = 1$, we have

$$\begin{aligned} Y_1^n(i+1) &= y_1 - \alpha(i) + \kappa(i) + A_1^n(i), \\ Y_2^n(i+1) &= y_2 - 1 - \kappa(i) + A_2^n(i), \\ Y_3^n(i+1) &= y_3 - \alpha(i) - \kappa(i) - \omega(i) + A_3^n(i), \end{aligned} \tag{2.16}$$

where

$$\mathbb{P}(\max\{A_1^n(i), A_2^n(i), A_3^n(i)\} > 0 | \mathcal{F}_i^n) \leq \tfrac{(c+c^2)(\log n)^2}{n}. \tag{2.17}$$

ii) If $I^n(i) = 0$, we have

$$\begin{aligned} Y_1^n(i+1) &= y_1 - 2, \\ Y_2^n(i+1) &= y_2, \\ Y_3^n(i+1) &= y_3 - 1. \end{aligned} \tag{2.18}$$

Proof. We firstly consider the case i), when the partner has degree at least 2. The change in the number of nodes of degree 1 arises as follows. Each neighbour of the partner which had degree 1 prior to the $(i+1)$st step of the algorithm is now isolated. Also, removal of an edge adjacent to the partner and a node of degree 2 causes the degree of that node to drop down to 1. The quantity $A_1^n(i)$ stands for the correction due to the possibility of loops and multiple edges. For the number of nodes of degree at least 2, we have the following situation. Firstly, the partner gets matched and thus removed from the graph. Also, each neighbour of the partner previously having degree 2 now becomes pendant. Again, the correction $A_2^n(i)$ stands for the error caused by loops and multiple edges. Finally, the number of edges we remove from the graph is the sum of the total number of edges adjacent to the partner and the number of loops belonging to vertices whose only neighbour is the partner. It remains to note that Lemma 27 justifies the asymptotics for the corrections. The result in case ii) follows trivially as in this case we remove an edge adjacent to two nodes of degree 1, leaving the rest of the graph intact. □

In order to determine the rates of our Markov chain, it is crucial to understand the degree distribution of the partner and its neighbours other than the leaf. The following lemma is the first step in that direction.

Lemma 30. *Let us fix a half-edge in the graph we have after i steps of the Karp-Sipser algorithm. Let $\Psi^n(i)$ be the degree of the node adjacent to the half-edge paired with it. For $i \leq X_4^n(\tau^n \wedge \zeta^n)$, let $(Y_1^n(i), Y_2^n(i), Y_3^n(i)) = (y_1, y_2, y_3)$ and let z denote the state* *of the jump chain of Z^n after i jumps. We have*

$$\mathbb{P}(\Psi^n(i) = 1|\mathcal{F}_i^n) = \frac{y_1}{2y_3} + S^n(i),$$

and

$$\mathbb{P}(\Psi^n(i) = k|\mathcal{F}_i^n) = \frac{y_2}{2y_3} \cdot \frac{z^k}{(k-1)!f(z)} + J_k^n(i),$$

for $2 \le k \le \log n$, where $\sup_{0 \le i \le X_4^n(\tau^n \wedge \zeta^n)} n \cdot |S^n(i)|$ is bounded with high probability and, as $n \to \infty$,

$$\sup_{0 \le i \le X_4^n(\tau^n \wedge \zeta^n)} \sup_{2 \le k \le \log n} \frac{n}{\log^3 n} \cdot |J_k^n(i)| \xrightarrow{p} 0.$$

Proof. Depending on whether the node adjacent to the fixed half edge has a degree 1 or more, there are either $y_1 - 1$ or y_1 available half-edges adjacent to nodes of degree 1 out of the total of $2y_3 - 1$ available half-edges. Bearing in mind that $\frac{y_1-1}{2y_3-1} = \frac{y_1}{2y_3} - \frac{2y_3-y_1}{2y_3(2y_3-1)}$ and $\frac{y_1}{2y_3-1} = \frac{y_1}{2y_3} + \frac{y_1}{2y_3(2y_3-1)}$, and that for $i \le X_4^n(\tau^n \wedge \zeta^n)$ we have

$$\frac{2y_3 - y_1}{2y_3(2y_3-1)} \le \frac{1}{2y_3} \le \frac{1}{\chi_3(t^*)n}, \tag{2.19}$$

as well as

$$\frac{y_1}{2y_3(2y_3-1)} \le \frac{1}{2y_3} \le \frac{1}{\chi_3(t^*)n}$$

where $\chi_3(t^*) > 0$, we get the result about the asymptotic of S^n.

We now turn to the case $k \ge 2$. Let us label y_2 nodes of degree at least 2 in an arbitrary manner with labels from the set $\{1, ..., y_2\}$. For $l \in \{1, ..., y_2\}$, let D_l^n be the degree of a node l, and let ϕ_l be the indicator that the fixed half-edge is paired with the half-edge adjacent to a node l. For $i \le X_4^n(\tau^n \wedge \zeta^n)$, we have

$$\begin{aligned}
\mathbb{P}(\Psi^n(i) = k|\mathcal{F}_i^n) &= \sum_{l=1}^{y_2} \mathbb{P}(\phi_l = 1, D_l^n = k|\mathcal{F}_i^n) \\
&= \sum_{l=1}^{y_2} \mathbb{P}(\phi_l = 1|D_l^n = k, \mathcal{F}_i^n)\mathbb{P}(D_l^n = k|\mathcal{F}_i^n) \\
&= y_2 \cdot \frac{k}{2y_3 - 1} \cdot \mathbb{P}(D_1^n = k|\mathcal{F}_i^n) \\
&=: \frac{y_2}{2y_3} \cdot \frac{z^k}{(k-1)!f(z)} + J_k^n(i)
\end{aligned} \tag{2.20}$$

where the last equality holds by Lemma 22 for

$$J_k^n(i) = \left(\frac{y_2}{2y_3-1} - \frac{y_2}{2y_3}\right) \cdot \frac{z^k}{(k-1)!f(z)} + \frac{y_2}{2y_3-1} \cdot k \cdot R_k^n(i).$$

Let us note that

$$\begin{aligned}\left(\frac{y_2}{2y_3-1} - \frac{y_2}{2y_3}\right) \cdot \frac{z^k}{(k-1)!f(z)} &= \frac{y_2}{2y_3(2y_3-1)} \cdot \frac{z^k}{(k-1)!f(z)} \\ &\leq \frac{1}{2y_3-1} \cdot \frac{z^k}{(k-1)!f(z)} \\ &\leq \frac{1}{\chi_3(t^*)n} \cdot \frac{z^k}{(k-1)!f(z)} \\ &\leq \frac{k}{\chi_3(t^*)n} \\ &\leq \frac{\log n}{\chi_3(t^*)n},\end{aligned}$$

where the second inequality follows as for $i \leq X_4^n(\tau^n \wedge \zeta^n)$ we have $X_3^n(i) \geq \frac{\chi_3(t^*)}{2}n$. We also have

$$\frac{y_2}{2y_3-1} \leq \frac{n}{2y_3-1} \leq \frac{1}{\chi_3(t^*)},$$

and so the asymptotic for J^n is the consequence of the asymptotic for R^n from Lemma 22. □

Lemma 31. *For $i \leq X_4^n(\tau^n \wedge \zeta^n)$, we have*

$$\mathbb{P}(\Psi^n(i) \geq 3 | \mathcal{F}_i^n) = \frac{y_2 \cdot z}{2y_3} + L^n(i),$$

where as $n \to \infty$

$$\sup_{0 \leq i \leq X_4^n(\tau^n \wedge \zeta^n)} \frac{n}{\log^3 n} |L^n(i)| \xrightarrow{p} 0.$$

Proof. As a consequence of Lemma 30, we have

$$\begin{aligned}\mathbb{P}(\Psi^n(i) \geq 3 | \mathcal{F}_i^n) &= 1 - \mathbb{P}(\Psi^n(i) = 1 | \mathcal{F}_i^n) - \mathbb{P}(\Psi^n(i) = 2 | \mathcal{F}_i^n) \\ &= 1 - \frac{y_1}{2y_3} - \frac{y_2}{2y_3}\frac{z^2}{f(z)} - S^n(i) - J_2^n(i).\end{aligned}$$

As (2.8) gives us $\frac{z(e^z-1)}{f(z)} \cdot \frac{y_2}{2y_3} = 1 - \frac{y_1}{2y_3}$, this is further equal to

$$\begin{aligned}\mathbb{P}(\Psi^n(i) \geq 3 | \mathcal{F}_i^n) &= \frac{y_2}{2y_3} \cdot \frac{z(e^z-1)}{f(z)} - \frac{y_2}{2y_3} \frac{z^2}{f(z)} - S^n(i) - J_2^n(i) \\ &=: \frac{y_2 \cdot z}{2y_3} + L^n(i),\end{aligned}$$

(recall that $f(z) = e^z - z - 1$). The asymptotics for S^n and J_2^n complete the proof. $\square$

The following lemma deals with the distribution of degrees of the neighbours of a fixed node, conditional on its degree.

Lemma 32. *We work on the event $\mathcal{A}^n$, i.e. the initial graph has maximum degree at most $\log n$. Let us fix the graph we have after i steps of the algorithm, and let $(Y_1^n(i), Y_2^n(i), Y_3^n(i)) = (y_1, y_2, y_3)$ and z be the state of the jump chain of Z^n after i jumps. Let us condition on the event that a fixed node in this graph has degree $k_1+k_2+k_3$, where since we work on $\mathcal{A}^n$ we must have $k_1+k_2+k_3 \leq \log n$. Conditioned on $\mathcal{F}_i^n$, such a node has k_1 half-edges leading to nodes of degree 1, k_2 half-edges leading to nodes of degree 2 and k_3 half-edges leading to nodes of degree at least 3 with probability*

$$\frac{(k_1+k_2+k_3)!}{k_1!k_2!k_3!} \left(\frac{y_1}{2y_3}\right)^{k_1} \left(\frac{y_2 z^2}{2y_3 f(z)}\right)^{k_2} \left(\frac{y_2 z}{2y_3}\right)^{k_3} + H_{k_1,k_2,k_3}^n(i) \tag{2.21}$$

where

$$\frac{n}{\log^3 n} \sup_{0 \leq i \leq X_4^n(\tau^n \wedge \zeta^n)} \sup_{(k_1,k_2,k_3) \in \mathcal{K}^n} |H_{k_1,k_2,k_3}^n(i)| \xrightarrow{p} 0,$$

as $n \to \infty$.

Proof. Let us first provide a heuristic explanation of how conditioning on a fixed node's degree affects the rest of the graph. When we condition on a fixed node having degree $2 \leq k_1+k_2+k_3$, the vector of degrees of the remaining $y_2 - 1$ nodes of degree at least 2 has multinomial distribution with $2y_3 - y_1 - (k_1+k_2+k_3)$ trials and $y_2 - 1$ categories with equal probabilities of success conditioned on each category having at least 2 successes. Having in mind that for $i \leq X_4^n(\tau^n \wedge \zeta^n)$ we have $Y_3^n(i) \geq \frac{\chi_3(t^*)}{2} n$ and $k_1+k_2+k_3 \leq \log n$, we expect this conditioning to introduce only minor changes to the approximation of the degree distribution. Consequently, we expect the degrees of the nodes connected to a given node to be asymptotically independent. Note that the leading term in (2.21) is the probability mass function of a multinomial distribution, with the probabilities of a node having a degree 1, 2 or at least 3 being equal to the leading terms of the appropriate probabilities from Lemma 30 and Lemma 31. We now formally develop the argument.

As before, for a fixed half-edge in the graph we have after i steps of the algorithm, let $\Psi^n(i)$ be the degree of the node adjacent to the half-edge paired with it in the configuration. Let us condition on the event $\mathcal{D}^n$ that a number $l \leq 2\log n$ of half-edges are already paired. Note that this implies that at most $2\log n$ vertices are adjacent to an already paired half-edge. Lemma 30 and Lemma 31 deal with the distribution of $\Psi^n(i)$ without conditioning, and so we adjust them in the following way. The part of the graph which is unaffected by the conditioning on the event $\mathcal{D}^n$, contains at least $y_1 + y_2 - 2\log n$ nodes and at least $2y_3 - 2 \cdot (2\log n)^2$ half-edges (recall that we work on $\mathcal{A}^n$, and so each "affected node" has a degree at most $\log n$). Let us note that Lemma 30 and Lemma 31 hold on that subgraph. Thus, it remains to impose the correction arising from the fact that we might connect to part of the graph which is "affected by the conditioning". Given that for $i \leq X_4^n(\tau^n \wedge \zeta^n)$ we have $Y_3^n(i) \geq \frac{\chi_3(t^*)}{2}n$, for n large enough we then have $Y_3^n(i) - (2\log n)^2 \geq \frac{\chi_3(t^*)}{3}n$, and so the absolute value of the correction is at most $\frac{24(\log n)^2}{\chi_3(t^*)n}$. Hence, by Lemma 30 we have

$$\mathbb{P}(\Psi^n(i) = 1|\mathcal{D}^n, \mathcal{F}_i^n) = \frac{y_1}{2y_3} + \tilde{H}^n(i), \tag{2.22}$$

where

$$|\tilde{H}^n(i)| \leq |S^n(i)| + \frac{24(\log n)^2}{\chi_3(t^*)n},$$

for S^n from Lemma 30. As a consequence of the asymptotic for S^n, as $n \to \infty$ we have

$$\frac{n}{(\log n)^3} \sup_{0 \leq i \leq X_4^n(\tau^n \wedge \zeta^n)} |\tilde{H}^n(i)| \xrightarrow{p} 0.$$

Similarly,

$$\mathbb{P}(\Psi^n(i) = 2|\mathcal{D}^n, \mathcal{F}_i^n) = \frac{y_2}{2y_3} \cdot \frac{z^2}{f(z)} + \hat{H}^n(i), \tag{2.23}$$

where

$$|\hat{H}^n(i)| \le |J_2^n(i)| + \frac{24(\log n)^2}{\chi_3(t^*)n},$$

for J_2^n from Lemma 30. As $n \to \infty$, the asymptotic for J_2^n gives us

$$\frac{n}{(\log n)^3} \sup_{0\le i\le X_4^n(\tau^n\wedge\zeta^n)} |\hat{H}^n(i)| \xrightarrow{p} 0.$$

Finally, by Lemma 31 we have

$$\mathbb{P}(\Psi^n(i) \ge 3|\mathcal{D}^n, \mathcal{F}_i^n) = \frac{y_2 \cdot z}{2y_3} + \bar{H}^n(i), \tag{2.24}$$

where

$$|\bar{H}^n(i)| \le |L^n(i)| + \frac{24(\log n)^2}{\chi_3(t^*)n},$$

for L^n from Lemma 31, and so the asymptotic for L^n leads us to

$$\frac{n}{(\log n)^3} \sup_{0\le i\le X_4^n(\tau^n\wedge\zeta^n)} |\bar{H}^n(i)| \xrightarrow{p} 0,$$

as $n \to \infty$.

Up to now, we considered the degree of the node adjacent to the half-edge paired with the half-edge we have fixed. Now we turn our attention to the neighbours of the fixed node of degree $k_1 + k_2 + k_3$. Let us label the half-edges adjacent to the fixed node by the labels from $\{1, \dots, k_1 + k_1 + k_3\}$ in an arbitrary manner. For $l \in \{1, \dots, k_1 + k_1 + k_3\}$, let Ψ_l denote a degree of a node adjacent to a half-edge paired with l. By the successive use of the approximations given by (2.22), (2.23) and (2.24), we get

$$\begin{aligned}
&\mathbb{P}(\Psi_1 = 1, \dots, \Psi_{k_1} = 1, \Psi_{k_1+1} = 2, \dots, \Psi_{k_1+k_2} = 2, \Psi_{k_1+k_2+1} \ge 3, \dots, \Psi_{k_1+k_2+k_3} \ge 3|\mathcal{F}_i^n) \\
&= \mathbb{P}(\Psi_1 = 1|\Psi_2 = 1 \dots, \Psi_{k_1} = 1, \Psi_{k_1+1} = 2, \dots, \Psi_{k_1+k_2+k_3} \ge 3, \mathcal{F}_i^n) \\
&\quad \cdot \mathbb{P}(\Psi_2 = 1 \dots, \Psi_{k_1} = 1, \Psi_{k_1+1} = 2, \dots, \Psi_{k_1+k_2+k_3} \ge 3|\mathcal{F}_i^n)
\end{aligned}$$

$$
\begin{aligned}
&= \left(\frac{y_1}{2y_3} + \tilde{H}_1^n(i)\right) \\
&\quad \cdot \mathbb{P}(\Psi_2 = 1, \dots, \Psi_{k_1} = 1, \Psi_{k_1+1} = 2, \dots, \Psi_{k_1+k_2} = 2, \Psi_{k_1+k_2+1} \geq 3, \dots, \Psi_{k_1+k_2+k_3} \geq 3 | \mathcal{F}_i^n) \\
&= \left(\frac{y_1}{2y_3} + \tilde{H}_1^n(i)\right) \dots \left(\frac{y_1}{2y_3} + \tilde{H}_{k_1}^n(i)\right) \\
&\quad \cdot \mathbb{P}(\Psi_{k_1+1} = 2, \dots, \Psi_{k_1+k_2} = 2, \Psi_{k_1+k_2+1} \geq 3, \dots, \Psi_{k_1+k_2+k_3} \geq 3 | \mathcal{F}_i^n) \\
&= \left(\frac{y_1}{2y_3} + \tilde{H}_1^n(i)\right) \cdots \left(\frac{y_1}{2y_3} + \tilde{H}_{k_1}^n(i)\right) \cdot \left(\frac{y_2}{2y_3} \cdot \frac{z^2}{f(z)} + \hat{H}_{k_1+1}^n\right) \cdots \left(\frac{y_2}{2y_3} \cdot \frac{z^2}{f(z)} + \hat{H}_{k_1+k_2}^n\right) \\
&\quad \cdot \mathbb{P}(\Psi_{k_1+k_2+1} \geq 3, \dots, \Psi_{k_1+k_2+k_3} \geq 3 | \mathcal{F}_i^n) \\
&= \left(\frac{y_1}{2y_3} + \tilde{H}_1^n(i)\right) \cdots \left(\frac{y_1}{2y_3} + \tilde{H}_{k_1}^n(i)\right) \cdot \left(\frac{y_2}{2y_3} \cdot \frac{z^2}{f(z)} + \hat{H}_{k_1+1}^n\right) \cdots \left(\frac{y_2}{2y_3} \cdot \frac{z^2}{f(z)} + \hat{H}_{k_1+k_2}^n\right) \\
&\quad \cdot \left(\frac{y_2 z}{2y_3} + \bar{H}_{k_1+k_2+1}^n\right) \cdots \left(\frac{y_2 z}{2y_3} + \bar{H}_{k_1+k_2+k_3}^n\right) \\
&= \left(\frac{y_1}{2y_3}\right)^{k_1} \left(\frac{y_2 z^2}{2y_3 f(z)}\right)^{k_2} \left(\frac{y_2 z}{2y_3}\right)^{k_3} + H_{k_1,k_2,k_3}^n(i).
\end{aligned}
$$

In each step we are conditioning on the event that at most $2(k_1 + k_2 + k_3) \leq 2\log n$ half-edges are already paired, and this justifies the inductive use of (2.22), (2.23) and (2.24). As a consequence of the asymptotics for $\tilde{H}^n$, $\hat{H}^n$ and $\bar{H}^n$, as $n \to \infty$, we get

$$
\frac{n}{(\log n)^4} \sup_{0 \leq i \leq X_4^n(\tau^n \wedge \zeta^n)} \sup_{(k_1,k_2,k_3) \in \mathcal{K}^n} |H_{k_1,k_2,k_3}^n(i)| \xrightarrow{p} 0,
$$

as $n \to \infty$. The additional $\log n$ term in the rescaling is a consequence of $k_1 + k_2 + k_3 \leq \log n$ and so the errors H_{k_1,k_2,k_3}^n have the form of a sum of at most $\log n$ quantities of a size which is negligible compared to $\frac{(\log n)^3}{n}$.

It remains to note that there are $\frac{(k_1+k_2+k_3)!}{k_1!k_2!k_3!}$ ways to choose k_1 half-edges which lead to nodes of degree 1, k_2 half-edges which lead to nodes of degree 2 and k_3 half-edges which lead to nodes of degree at least 3. $\square$

We are now ready to determine the jump probabilities.

Proof of Theorem 28. Let us recall Lemma 29 which identifies possible transitions of the jump chain of X^n in Phase 1. As before, let $\alpha(i)$, $\kappa(i)$, $\omega(i)$ denote the number of half-edges adjacent to the partner which lead to a node of degree 1, 2 and at least 3, respectively. Having in mind (2.16), for each $(k_1, k_2, k_3) \in \mathcal{K}^n$ we have

$$\mathbb{P}(Y_1^n(i+1) = y_1 - k_1 + k_2, Y_2^n(i+1) = y_2 - 1 - k_2, Y_3^n(i+1) = y_3 - k_1 - k_2 - k_3 | \mathcal{F}_i^n)$$
$$= \mathbb{P}(\alpha(i) = k_1, \kappa(i) = k_2, \omega(i) = k_3 | \mathcal{F}_i^n) + \mathcal{E}^n(i),$$

where by (2.17) we have

$$\sup_{0 \le i \le X_4^n(\tau^n \wedge \zeta^n)} \mathcal{E}^n(i) \le \frac{(c+c^2)(\log n)^2}{n}.$$

This is further equal to

$$\begin{aligned}
&\mathbb{P}(\alpha(i) = k_1, \kappa(i) = k_2, \omega(i) = k_3 | \alpha(i) + \kappa(i) + \omega(i) = k_1 + k_2 + k_3, \mathcal{F}_i^n) \\
&\cdot \mathbb{P}(\alpha(i) + \kappa(i) + \omega(i) = k_1 + k_2 + k_3 | \mathcal{F}_i^n) + \mathcal{E}^n(i) \\
&= \left(\frac{(k_1+k_2+k_3-1)!}{(k_1-1)!k_2!k_3!} \cdot \left(\frac{y_1}{2y_3}\right)^{k_1-1} \cdot \left(\frac{y_2}{2y_3}\frac{z^2}{f(z)}\right)^{k_2} \cdot \left(\frac{y_2}{2y_3}z\right)^{k_3} + H^n_{k_1,k_2,k_3}(i) \right) \\
&\quad \cdot \left(\frac{y_2}{2y_3} \cdot \frac{z^{k_1+k_2+k_3}}{(k_1+k_2+k_3-1)! f(z)} + J^n_{k_1+k_2+k_3}(i) \right) + \mathcal{E}^n(i) \\
&=: \frac{y_2}{2y_3}\frac{z}{f(z)} \left(\frac{y_1 z}{2y_3}\right)^{k_1-1} \left(\frac{y_2 z^3}{2y_3 f(z)}\right)^{k_2} \left(\frac{y_2 z^2}{2y_3}\right)^{k_3} \cdot \frac{1}{(k_1-1)!k_2!k_3!} + E^n_{k_1,k_2,k_3}(i),
\end{aligned}$$

where the first equality is a consequence of Lemma 32 and Lemma 30. In particular, note that one of the partner's half-edges is paired with the leaf with probability 1, and so it remains to pair the rest of $k_1 + k_2 + k_3 - 1$ half-edges. The asymptotic for the errors in Lemmas 30 and 32 together with $\sup_{0 \le i \le X_4^n(\tau^n \wedge \zeta^n)} \mathcal{E}^n(i) \le \frac{(c+c^2)(\log n)^2}{n}$ leads to

$$\frac{n}{\log^3 n} \sup_{0 \le i \le X_4^n(\tau^n \wedge \zeta^n)} \sup_{(k_1,k_2,k_3) \in \mathcal{K}^n} \left| E^n_{k_1,k_2,k_3}(i) \right| \xrightarrow{p} 0,$$

as desired. We also have

$$\mathbb{P}(Y_1^n(i+1) = y_1 - 2, Y_2^n(i+1) = y_2, Y_3^n(i+1) = y_3 - 1 | \mathcal{F}_i^n) = \frac{y_1 - 1}{2y_3 - 1},$$

and bearing in mind that $\frac{y_1-1}{2y_3-1} = \frac{y_1}{2y_3} - \frac{2y_3 - y_1}{2y_3(2y_3-1)}$, (2.19) gives us the result about the asymptotic for B^n and this completes the proof. □

For $(k_1, k_2, k_3) \notin \mathcal{K}^n$, and any $i \leq X_4^n(\tau^n \wedge \zeta^n)$, let

$$E^n_{k_1,k_2,k_3}(i) = 0. \tag{2.25}$$

The following property of the errors from Theorem 28 will turn out as useful.

Lemma 33. *For the errors from Theorem 28, as $n \to \infty$ we have*

$$\sqrt{n}\sum_{k_1}^{\infty}\sum_{k_2=0}^{\infty}\sum_{k_3=0}^{\infty}\Bigg(\Big((-k_1+k_2)^2+(-1-k_2)^2+(-k_1-k_2-k_3)^2\Big) \cdot \sup_{0\leq i\leq X_4^n(\tau^n\wedge\zeta^n)} |E_{k_1,k_2,k_3}(i)|\Bigg) + \sqrt{n}\cdot \sup_{0\leq i\leq X_4^n(\tau^n\wedge\zeta^n)} |B^n(i)| \xrightarrow{p} 0.$$

Proof. We have

$$\begin{aligned}\sqrt{n}\sum_{k_1}^{\infty}\sum_{k_2=0}^{\infty}\sum_{k_3=0}^{\infty}&\Bigg(\Big((-k_1+k_2)^2+(-1-k_2)^2+(-k_1-k_2-k_3)^2\Big) \cdot \sup_{0\leq i\leq X_4^n(\tau^n\wedge\zeta^n)} |E_{k_1,k_2,k_3}(i)|\Bigg)\\ = \sqrt{n}\cdot \sum_{(k_1,k_2,k_3)\in\mathcal{K}^n}&\Big(\big((-k_1+k_2)^2+(-1-k_2)^2+(-k_1-k_2-k_3)^2\big) \cdot \sup_{0\leq i\leq X_4^n(\tau^n\wedge\zeta^n)} |E_{k_1,k_2,k_3}(i)|\Big),\end{aligned}$$

where the equality holds by (2.25). Let us recall that $\mathcal{K}^n = \{(k_1, k_2, k_3) \in \mathbb{Z}^3 : k_1 \geq 1, k_2, k_3 \geq 0, \log n \geq k_1 + k_2 + k_3 \geq 2\}$, and so the number of terms in the sum is less than $(\log n)^3$. Thus, the previous is at most

$$\sqrt{n}\cdot(\log n)^3\cdot 3\cdot(\log n)^2\cdot \sup_{0\leq i\leq X_4^n(\tau^n\wedge\zeta^n)}\ \sup_{(k_1,k_2,k_3)\in\mathcal{K}^n} |E_{k_1,k_2,k_3}(i)|,$$

Let us recall that Theorem 28 guarantees that

$$\frac{n}{(\log n)^3}\sup_{0\leq i\leq X_4^n(\tau^n\wedge\zeta^n)}\ \sup_{(k_1,k_2,k_3)\in\mathcal{K}^n} |E^n_{k_1,k_2,k_3}(i)| \xrightarrow{p} 0,$$

and this, together with the asymptotic for B^n, completes the proof. $\square$

2.3 The central limit theorem for the Markov chain in Phase 1

2.3.1 The framework for proving a CLT for a Markov chain

Our main tool is a collection of results from Ethier and Kurtz [21], Chapter 11 which we slightly adjust to our setting. These results belong to the general framework for the analysis of density dependent population processes we briefly commented on in Chapter 1. In this chapter, the main objects of our study are continuous time Markov processes. In Chapter 3, we exploit this framework in a slightly different setting, as the main object of our study there are stochastic processes evolving in a discrete time, which are either Markov or represent a coordinate of a higher-dimensional Markov process.

Let $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ denote the space of càdlàg functions from $\mathbb{R}_+$ to $\mathbb{R}$, equipped with the Skorokhod topology. Assume that we have a collection of nonnegative functions β_l and r_l^n for $l \in \mathbb{Z}^d$, $n \in \mathbb{N}$ defined on $E \subseteq \mathbb{R}^d$. Let

$$E_n = E \cap \left\{ \frac{k}{n} : k \in \mathbb{Z}^d \right\},$$

and let $x \in E_n$ and $\beta_l(x) + r_l^n(x) > 0$, imply $x + l/n \in E_n$.

Let $(U^n(t))_{t \geq 0}$, where for $t \geq 0$ we have $U^n(t) = (U_1^n(t), \ldots, U_d^n(t))^T$, be a continuous time Markov process whose state space is a subset of $\mathbb{Z}^d$. For each $l \in \mathbb{Z}^d$ let the rate of the transition from $k \in \mathbb{Z}^d$ to $k + l$ be given by $n(\beta_l(k/n) + r_l^n(k/n))$. Let us fix t such that U^n has made only finitely many jumps up to t with probability 1. For $l \in \mathbb{Z}^d$, let P_l^n be independent Poisson processes with rate 1. Then, there is a realization of the process ([21], Chapter 6, Theorem 4.1) that satisfies

$$\frac{U^n(t)}{n} = \frac{U^n(0)}{n} + \sum_{l \in \mathbb{Z}^d} \frac{l \cdot P_l^n\Big(n \int_0^t \Big(\beta_l\Big(\frac{U^n(s)}{n}\Big) + r_l^n\Big(\frac{U^n(s)}{n}\Big)\Big) ds\Big)}{n}. \tag{2.26}$$

The weak law of large numbers suggests that the second term on the right hand side of (2.26) converges to its mean in probability as $n \to \infty$. If we assume that the r_l^n are small enough (we will be more precise about this in a moment), and that $\frac{U^n(0)}{n}$ converges in probability to some constant $u(0)$, then it makes sense to compare $\frac{U^n(t)}{n}$ with the solution to the following integral equation (assuming that such a solution exists)

$$u(t) = u(0) + \int_0^t F(u(s))ds \qquad t \geq 0, \tag{2.27}$$

where the *drift function* $F : E \to \mathbb{R}^d$ is defined by

$$F(x) = (F_1(x), \ldots, F_d(x))^T = \sum_{l \in \mathbb{Z}^d} l\beta_l(x). \tag{2.28}$$

Theorem 34. *Suppose that*

$$\sum_{l \in \mathbb{Z}^d} |l| \sup_{x \in E} \beta_l(x) < \infty \tag{2.29}$$

where $|\cdot|$ *denotes the supremum norm, and that there exists* $M > 0$ *such that for* $x, y \in E$ *we have*

$$|F(x) - F(y)| \leq M|x - y|. \tag{2.30}$$

Also, assume that

$$\limsup_{n \to \infty} \sum_{l \in \mathbb{Z}^d} |l| \sup_{x \in E} |r_l^n(x)| = 0, \tag{2.31}$$

and that $\frac{U^n(0)}{n} \xrightarrow{P} u(0)$ *as* $n \to \infty$*. Then for every* $t \geq 0$ *and every* $\delta > 0$*, as* $n \to \infty$ *we have*

$$\mathbb{P}\left(\sup_{s \leq t} \left| \frac{U^n(s)}{n} - u(s) \right| > \delta \right) \to 0.$$

Proof. This is a mild generalization of Theorem 2.1 from [21], Chapter 11. In their original setting, for all $n \in \mathbb{N}$, $l \in \mathbb{Z}^d$ and $x \in E$ they have $r_l^n(x) = 0$. Bearing in mind (2.26) and (2.28), we have

$$\begin{aligned}
\varepsilon^n(t) &= \sup_{0 \leq s \leq t} \left| \tfrac{U^n(s)}{n} - \tfrac{U^n(0)}{n} - \int_0^s F\left(\tfrac{U^n(w)}{n}\right) dw \right| \\
&\leq \sup_{0 \leq s \leq t} \sum_{l \in \mathbb{Z}^d} \tfrac{|l|}{n} \left| P_l^n\left(n \int_0^s \left(\beta_l\left(\tfrac{U^n(w)}{n}\right) + r_l^n\left(\tfrac{U^n(w)}{n}\right)\right) dw \right) - n \int_0^s \left(\beta_l\left(\tfrac{U^n(w)}{n}\right) + r_l^n\left(\tfrac{U^n(w)}{n}\right)\right) dw \right| \\
&\quad + \sum_{l \in \mathbb{Z}^d} |l| t \sup_{x \in E} |r_l^n(x)|.
\end{aligned}$$

The first term on the right hand side of the inequality tends to zero in probability by a weak law of large numbers argument, in the same way as (2.14) holds in the proof of Theorem 2.1 from [21], Chapter 11. The condition (2.31) guarantees that the second term on the right hand side of the inequality also tends to zero in probability, and so we have $\varepsilon^n(t) \xrightarrow{p} 0$ as $n \to \infty$. Using (2.27), we have

$$\begin{aligned}\sup_{0\le s\le t}\Big|\frac{U^n(s)}{n} - u(s)\Big| &\le \varepsilon^n(t) + \Big|\frac{U^n(0)}{n} - u(0)\Big| + \sup_{0\le s\le t}\Big|\int_0^s \Big(F\Big(\frac{U^n(w)}{n}\Big) - F(u(w))\Big)dw\Big| \\ &\le \varepsilon^n(t) + \Big|\frac{U^n(0)}{n} - u(0)\Big| + M\int_0^t \sup_{0\le q\le w}\Big|\frac{U^n(q)}{n} - u(q)\Big|dw.\end{aligned}$$

We have already proved that $\varepsilon^n(t) \xrightarrow{p} 0$. By assumption, the second term on the right hand side of the last inequality also tends to zero in probability. Thus, for a fixed $\delta > 0$ on an event of probability tending to 1, we have

$$\sup_{0\le s\le t}\Big|\frac{U^n(s)}{n} - u(s)\Big| \le \delta e^{-Mt} + M\int_0^t \sup_{0\le q\le w}\Big|\frac{U^n(q)}{n} - u(q)\Big|dw,$$

and so Gronwall's lemma implies that the probability of the event

$$\sup_{0\le s\le t}\Big|\frac{U^n(s)}{n} - u(s)\Big| \le \delta$$

tends to 1 as $n \to \infty$. This completes the proof. $\square$

Having in mind (2.26) and the previous theorem, the central limit theorem suggests we should explore the behaviour of the fluctuations of the process around the deterministic approximation. Thus, for $t \ge 0$ we introduce $V^n(t) = (V_1^n(t), \dots, V_d^n(t))^T$ which is defined by

$$V^n(t) = \sqrt{n}\Big(\frac{U^n(t)}{n} - u(t)\Big). \tag{2.32}$$

Bearing in mind (2.26), (2.27) and (2.28), we have

$$\begin{aligned}V^n(t) &= V^n(0) \\ &+ \sum_{l\in\mathbb{Z}^d} \frac{l}{\sqrt{n}}\Big(P_l^n\Big(n\int_0^t \Big(\beta_l\Big(\tfrac{U^n(s)}{n}\Big) + r_l^n\Big(\tfrac{U^n(s)}{n}\Big)\Big)ds\Big) - n\int_0^t\Big(\beta_l\Big(\tfrac{U^n(s)}{n}\Big) + r_l^n\Big(\tfrac{U^n(s)}{n}\Big)\Big)ds\Big) \\ &+ \int_0^t \sqrt{n}\Big(F\Big(\tfrac{U^n(s)}{n}\Big) - F(u(s))\Big)ds + \sum_{l\in\mathbb{Z}^d}\int_0^t \sqrt{n}\cdot l\cdot r_l^n\Big(\tfrac{U^n(s)}{n}\Big)ds,\end{aligned} \tag{2.33}$$

where $V^n(0)$ and the Poisson processes P_l^n are independent. Let us assume that for some random vector $V(0)$ taking values in $\mathbb{R}^d$, as $n \to \infty$ we have

$$V^n(0) \xrightarrow{d} V(0).$$

Theorem 34 guarantees the closeness of $\frac{U^n}{n}$ and the function u, and thus under the central limit theorem rescaling we have imposed, we expect the second term on the right hand side of (2.33) to converge in distribution to a Gaussian process in $\mathbb{R}^d$ (assuming that appropriate regularity conditions are satisfied). Let us assume that as $n \to \infty$ the last term on the right hand side of (2.33) converges in probability to a zero-vector in $\mathbb{R}^d$. Then, it seems reasonable to explore the relationship between $V^n(t)$ and the solution to the stochastic differential equation (in integral form)

$$V(t) = V(0) + \sum_{l \in \mathbb{Z}^d} l W_l \Big(\int_0^t \beta_l(u(s)) ds \Big) + \int_0^t \partial F(u(s)) V(s) ds, \tag{2.34}$$

(assuming that appropriate regularity conditions are satisfied and such a solution exists), where

$$\partial F(x) = ((\partial_j F_i(x))), \quad i, j = 1, \ldots, d,$$

and W_l, $l \in \mathbb{Z}^d$ are independent Brownian motions in $\mathbb{R}$, and the random vector $V(0)$ is independent of everything else.

Lemma 35. *Suppose that, in addition to the conditions of Theorem 34, we have*

$$\sum_{l \in \mathbb{Z}^d} |l|^2 \sup_{x \in E} \beta_l(x) < \infty,$$

and that β_l and ∂F are continuous functions. Then, there is a unique strong solution to (2.34).

Proof. As $\partial F(x)$, $x \in E$ is a continuous function, there is a unique solution to the partial differential equation

$$\frac{\partial}{\partial t} \Phi(t, s) = \partial F(u(t)) \Phi(t, s), \quad \Phi(s, s) = I, \tag{2.35}$$

on $0 \leq s \leq t < \infty$. Then, as in [21], Chapter 11, Section 2, the SDE is explicitly solvable via an integrating factor, and the solution is given by

$$V(t) = \Phi(t,0)V(0) + \int_0^t \Phi(t,s)d\mathcal{W}(s)$$

where

$$\mathcal{W}(s) = \sum_{l\in\mathbb{Z}^d} lW_l\Big(\int_0^s \beta_l(u(q))dq\Big). \tag{2.36}$$

Thus, we have a unique strong solution. □

Theorem 36. *Suppose that, in addition to the conditions of Theorem 34, we have*

$$\sum_{l\in\mathbb{Z}^d} |l|^2 \sup_{x\in E} \beta_l(x) < \infty, \tag{2.37}$$

and

$$\sqrt{n}\cdot \sum_{l\in\mathbb{Z}^d} |l|^2 \sup_{0\leq s\leq t} \left| r_l^n\left(\frac{U^n(s)}{n}\right)\right| \xrightarrow{p} 0 \quad \text{as } n\to\infty. \tag{2.38}$$

Let us also assume that β_l and ∂F are continuous, and let $(V(s), 0\leq s\leq t)$ be the unique strong solution to (2.34). Then, $(V_n(s), 0 \leq s \leq t) \xrightarrow{d} (V(s), 0 \leq s \leq t)$ uniformly as $n\to\infty$.

Before we proceed to the proof of Theorem 36, we prove the following lemma.

Lemma 37. *Let us assume that conditions of Theorem 36 are fulfilled. For $0 \leq s \leq t$, let*

$$\hat{G}^n(s) = \sum_{l\in\mathbb{Z}^d} \tfrac{l}{\sqrt{n}}\Big(P_l^n\Big(n\int_0^s \Big(\beta_l\Big(\tfrac{U^n(w)}{n}\Big) + r_l^n\Big(\tfrac{U^n(w)}{n}\Big)\Big)dw\Big) - n\int_0^s\Big(\beta_l\Big(\tfrac{U^n(w)}{n}\Big) + r_l^n\Big(\tfrac{U^n(w)}{n}\Big)\Big)dw\Big),$$

and

$$G^n(s) = \sum_{l\in\mathbb{Z}^d} \tfrac{l}{\sqrt{n}}\Big(P_l^n\Big(n\int_0^s \beta_l(u(w))dw\Big) - n\int_0^s \beta_l(u(w))dw\Big).$$

Then,

$$\sup_{0\le s\le t}|\hat{G}^n(s)-G^n(s)|\xrightarrow{p}0,$$

as $n\to\infty$, where as usual $|\cdot|$ denotes the supremum norm.

Proof. In [21], Chapter 11 the case $r_l^n(x)=0$ for $n\in\mathbb{N}$, $l\in\mathbb{Z}^d$, $x\in E$ is analysed, and these results give us that

$$\sup_{0\le s\le t}\left|\sum_{l\in\mathbb{Z}^d}\frac{l}{\sqrt{n}}\left(P_l^n\left(n\int_0^s\beta_l\left(\tfrac{U^n(w)}{n}\right)dw\right)-n\int_0^s\beta_l\left(\tfrac{U^n(w)}{n}\right)dw\right)-G^n(s)\right|\xrightarrow{p}0.$$

It remains to note that the condition (2.38) ensures that

$$\sup_{0\le s\le t}\left|\sum_{l\in\mathbb{Z}^d}\frac{l}{\sqrt{n}}\left(P_l^n\left(n\int_0^s\beta_l\left(\tfrac{U^n(w)}{n}\right)dw\right)-n\int_0^s\beta_l\left(\tfrac{U^n(w)}{n}\right)dw\right)-\hat{G}^n(s)\right|\xrightarrow{p}0,$$

and this completes the proof. □

Proof of Theorem 36. This is an adjustment of Theorem 2.3 from [21], Chapter 11. As already mentioned, in their setting they have $r_l^n(x)=0$ for all $n\in\mathbb{N}$, $l\in\mathbb{Z}^d$, $x\in E$ and $V(0)$ is not random.

For $0\le s\le t$, let $\hat{G}^n(s)$ and $G^n(s)$ be as in Lemma 37 and

$$\epsilon_1^n(s)=\hat{G}^n(s)-G^n(s).$$

Lemma 37 guarantees that $\sup_{0\le s\le t}|\epsilon_1^n(s)|\xrightarrow{p}0$ as $n\to\infty$. Let

$$\epsilon_2^n(t)=\int_0^t\sqrt{n}\left(F\left(\frac{U^n(s)}{n}\right)-F(u(s))-n^{-1/2}\partial F(u(s))V^n(s)\right)ds.$$

Bearing in mind that $V^n(s)=\sqrt{n}\left(\frac{U^n(s)}{n}-u(s)\right)$, by the mean value theorem we have

$$\epsilon_2^n(t)=\int_0^t\left(\int_0^1\partial F\left(u(s)+\xi\left(\frac{U^n(s)}{n}-u(s)\right)\right)d\xi-\partial F(u(s))\right)V^n(s)ds.$$

Theorem 34 gives us that $\sup_{s\leq t}|\frac{U^n(s)}{n}-u(s)|\xrightarrow{p}0$, and together with continuity of ∂F this leads us to $\sup_{0\leq s\leq t}|\epsilon_2^n(s)|\xrightarrow{p}0$. Also, for

$$\epsilon_3^n(t)=\int_0^t\sqrt{n}\sum_{l\in\mathbb{Z}^d}l\cdot r_l^n\Big(\frac{U^n(s)}{n}\Big)ds,$$

condition (2.38) implies that $\sup_{0\leq s\leq t}|\epsilon_3^n(s)|\xrightarrow{p}0$. Consequently, for

$$\epsilon^n(t)=\epsilon_1^n(t)+\epsilon_2^n(t)+\epsilon_3^n(t),$$

we have

$$\sup_{0\leq s\leq t}|\epsilon^n(s)|\xrightarrow{p}0. \tag{2.39}$$

Using the representation (2.33), we have

$$V^n(t)=V^n(0)+G^n(t)+\epsilon^n(t)+\int_0^t\partial F(u(s))V^n(s)ds.$$

By the same reasoning as in the proof of Lemma 35, we have

$$V^n(t)=\Phi(t,0)V^n(0)+\int_0^t\Phi(t,s)d(G^n(s)+\epsilon^n(s)), \tag{2.40}$$

where $\Phi(t,s)$, $0\leq s\leq t$, is the solution to the partial differential equation from Lemma 35. Let us recall that by Lemma 35 we have

$$V(t)=\Phi(t,0)V(0)+\int_0^t\Phi(t,s)d\mathcal{W}(s), \tag{2.41}$$

where $\mathcal{W}(s)=\sum_{l\in\mathbb{Z}^d}lW_l(\int_0^s\beta_l(u(q))dq)$, and W_l, $l\in\mathbb{Z}^d$ are independent Brownian motions in $\mathbb{R}$, independent of $V(0)$.

Let Θ be the functional from $\mathbb{R}^d\times\mathbb{D}(\mathbb{R}_+,\mathbb{R}^d)$ into $\mathbb{D}(\mathbb{R}_+,\mathbb{R}^d)$ given by

$$\Theta(b,h)(s)=\Phi(s,0)b+\int_0^s\Phi(s,q)dh(q),$$

for $s \geq 0$. Condition (2.37) together with (2.39) implies that

$$(G^n(s) + \epsilon^n(s), 0 \leq s \leq t) \xrightarrow{d} (\mathcal{W}(s), 0 \leq s \leq t),$$

uniformly as $n \to \infty$. The independence of $V^n(0)$ and each P_l^n guarantees the independence of $V^n(0)$ and $G^n(s)$, $0 \leq s \leq t$. Thus, we have

$$(V^n(0), G^n(s) + \epsilon^n(s), 0 \leq s \leq t) \xrightarrow{d} (V(0), \mathcal{W}(s), 0 \leq s \leq t)$$

uniformly as $n \to \infty$, where $V(0)$ and $\mathcal{W}$ are independent. It remains to note that Θ is continuous, and by the continuous mapping theorem we have

$$(\Theta(V^n(0), G^n(\cdot) + \epsilon^n(\cdot))(s), 0 \leq s \leq t) \xrightarrow{d} (\Theta(V(0), \mathcal{W}(\cdot))(s), 0 \leq s \leq t)$$

uniformly as $n \to \infty$, and having in mind (2.40) and (2.41), this completes the proof. $\square$

Theorem 36 proves the convergence in distribution up until a *fixed time* t. However, one is often interested in the properties of a Markov chain at the time when it leaves a certain set, that is at a *random time*. Again, the use of results from [21] enables us to provide the appropriate extension of Theorem 36. Here we state an extension of Theorem 36 in the form convenient for our analysis. For the d-dimensional Markov process $U^n(t) = (U_1^n(t), \ldots, U_d^n(t))^T$, $t \geq 0$, let

$$\mathcal{T}^n = \inf\{t \geq 0 : U_1^n(t) \leq 0\},$$

be the time when the first coordinate of U^n drops below zero. Also, for the deterministic function $u(t) = (u_1(t), \ldots, u_d(t))^T$ given by (2.27) as before, let

$$a = \inf\{t \geq 0 : u_1(t) \leq 0\},$$

be the time when its first coordinate drops below zero.

Theorem 38. *Suppose that for the Markov process U^n given by (2.26) and a function $u : \mathbb{R}_+ \to \mathbb{R}^d$ given by (2.27) the conditions of Theorem 36 hold. Let us also assume* *that* $a < \infty$ *and that for a function* F *given by (2.28) we have*

$$F_1(u(a)) < 0. \tag{2.42}$$

Then, as $n \to \infty$ *we have*

$$\sqrt{n}(\mathcal{T}^n - a) \xrightarrow{d} -\frac{V_1(a)}{F_1(u(a))},$$

as well as

$$\sqrt{n}\Big(\frac{U^n(\mathcal{T}^n)}{n} - u(a)\Big) \xrightarrow{d} V(a) - \frac{V_1(a)}{F_1(u(a))}F(u(a)). \tag{2.43}$$

Theorem 38 is a consequence of Theorem 36 in the same way as Theorem 4.1 in [21], Chapter 11, Section 4 follows from Theorem 2.3 there. Condition (2.42) tells us that the function u_1 has a negative derivative at the time when it first hits zero, and so it does not come back to the positive half-plane "too soon". This property guarantees that the process U_1^n also hits zero some time close to a. Let us also note that the second term on the right hand side of (2.43) stands for the correction proportional to the asymptotic difference between $\mathcal{T}^n$ and a.

2.3.2 Initial fluctuations

Equipped with the tools from Section 2.3.1, we now come back to the analysis of the Markov chain X^n. In this section, we prove the following theorem about the initial fluctuations of X^n around the fluid limit approximation.

Theorem 39. *As* $n \to \infty$*, the random vector*

$$\Big(\frac{X_1^n(0) - n\chi_1(0)}{\sqrt{n}}, \frac{X_2^n(0) - n\chi_2(0)}{\sqrt{n}}, \frac{X_3^n(0) - n\chi_3(0)}{\sqrt{n}}, \frac{X_4^n(0)}{\sqrt{n}}\Big)$$

converges in distribution to a multivariate normal distribution with mean zero and the covariance matrix Σ *given by*

$$\begin{pmatrix} c^2e^{-2c}+ce^{-c}-ce^{-2c}-c^3e^{-2c} & -ce^{-c}+ce^{-2c}+c^3e^{-2c} & 0 & 0 \\ -ce^{-c}+ce^{-2c}+c^3e^{-2c} & (e^{-c}+ce^{-c})(1-e^{-c}-ce^{-c})-c^3e^{-2c} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Before we proceed to the proof of the theorem, it will be useful to define the following notion.

Definition 40 (Definition 2.1 from [27]). *Let* $\mathcal{Q}$ *and* $\mathcal{R}$ *be random variables defined on the same probability space, and let* $\mathcal{R}$ *have a discrete distribution. We say that* $\mathcal{Q}$ *is stochastically decreasing with respect to* $\mathcal{R}$ *if for any real* q *and* $b_1 \le b_2$ *we have*

$$\mathbb{P}(\mathcal{Q} \le q | \mathcal{R} = b_1) \le \mathbb{P}(\mathcal{Q} \le q | \mathcal{R} = b_2).$$

We will use the following lemma from Janson [27] in our proof (we state here its convenient version).

Lemma 41 (Theorem 2.3 from [27]). *Let us suppose that* $(\mathcal{Q}^n, \mathcal{R}^n)$, $n \ge 1$ *are pairs of random vectors* $\mathcal{Q}^n = (\mathcal{Q}^n_1, \mathcal{Q}^n_2)$ *taking values in* $\mathbb{R}^2$ *and random variables* $\mathcal{R}^n$ *taking values in* $\mathbb{Z}$, *defined on the same probability space. Let us assume that* $\mathcal{Q}^n$ *is stochastically decreasing in each component with respect to* $\mathcal{R}^n$ *and that for some sequences of real numbers and vectors* $a_n > 0$, $b_n \in \mathbb{R}^2$, $c_n > 0$, $d_n \in \mathbb{R}$, *as* $n \to \infty$ *we have*

$$\left(\frac{\mathcal{Q}^n - b_n}{a_n}, \frac{\mathcal{R}^n - d_n}{c_n}\right) \xrightarrow{d} (\mathcal{Q}, \mathcal{R}),$$

for a 3*-dimensional normal random vector* $(\mathcal{Q}, \mathcal{R})$ *with* $\mathcal{Q} = (\mathcal{Q}_1, \mathcal{Q}_2)$ *such that* $\mathrm{Var}(\mathcal{R}) > 0$. *Assume that* r_n *is a sequence in* $\mathbb{R}$ *such that* $\frac{r_n - d_n}{c_n} \to \xi$ *for some real* ξ *and let* $\tilde{\mathcal{Q}}^n = \left(\tilde{\mathcal{Q}}^n_1, \tilde{\mathcal{Q}}^n_2\right)$ *be a random vector whose distribution equals to the conditional distribution* $\mathcal{Q}^n | \mathcal{R}^n = r_n$. *Then, for* $\eta = (\eta_1, \eta_2)$, *where* $\eta_1 = \frac{\mathrm{Cov}(\mathcal{Q}_1, \mathcal{R})}{\mathrm{Var}(\mathcal{R})}$ *and* $\eta_2 = \frac{\mathrm{Cov}(\mathcal{Q}_2, \mathcal{R})}{\mathrm{Var}(\mathcal{R})}$, *we have*

$$\frac{\tilde{\mathcal{Q}}^n - b_n}{a_n} \xrightarrow{d} \mathcal{Q} + (\xi - \mathcal{R})\eta.$$

The limiting vector has a normal distribution with the expectation $\mathbb{E}[\mathcal{Q}] + (\xi - \mathbb{E}[\mathcal{R}])\eta$ *and the covariance matrix determined by*

$$\mathrm{Cov}(\tilde{\mathcal{Q}}_i, \tilde{\mathcal{Q}}_j) = \mathrm{Cov}(\mathcal{Q}_i, \mathcal{Q}_j) - \mathrm{Cov}(\mathcal{Q}_i, \mathcal{R})\mathrm{Cov}(\mathcal{Q}_j, \mathcal{R})/\mathrm{Var}(\mathcal{R}).$$

for $i, j \in \{1, 2\}$.

The essential ingredient of the proof of Theorem 39 will be the following lemma.

Lemma 42. *As* $n \to \infty$*, the vector consisting of the suitably rescaled number of nodes of degree* 0 *and at most* 1

$$\left(\frac{n - X_1^n(0) - X_2^n(0) - n(1 - \chi_1(0) - \chi_2(0))}{\sqrt{n}}, \frac{n - X_2^n(0) - n(1 - \chi_2(0))}{\sqrt{n}}\right),$$

in the multigraph with n *nodes and* $M = \lfloor \frac{cn}{2} \rfloor$ *edges, converges in distribution to a multivariate normal variable with mean zero and the covariance matrix given by*

$$\begin{pmatrix} e^{-c}(1 - e^{-c}) - ce^{-2c} & e^{-c} - e^{-2c} - ce^{-2c} - c^2e^{-2c} \\ e^{-c} - e^{-2c} - ce^{-2c} - c^2e^{-2c} & (e^{-c} + ce^{-c})(1 - e^{-c} - ce^{-c}) - c^3e^{-2c} \end{pmatrix}.$$

Proof. We use the Poissonization trick by assuming that initially we have a random number of edges, and then exploit stochastic monotonicity of certain random variables with respect to the number of edges, in order to obtain the desired convergence. Namely, let us assume that the degree D_k of each node $k \in \{1, \dots, n\}$ has Poiss(c) distribution, and these are independent. Let us note that

$$\left(\sum_{k=1}^{n} \mathbb{1}_{\{D_k=0\}}, \sum_{k=1}^{n} \left(\mathbb{1}_{\{D_k=0\}} + \mathbb{1}_{\{D_k=1\}}\right)\right)$$

is stochastically decreasing with respect to the total number of half-edges $\sum_{k=1}^n D_k$. This is because adding one more half-edge can only increase the nodes' degrees. The central limit theorem gives us that as $n \to \infty$

$$\left(\frac{\sum_{k=1}^n \mathbb{1}_{\{D_k=0\}} - ne^{-c}}{\sqrt{n}}, \frac{\sum_{k=1}^n \mathbb{1}_{\{D_k=0\}} + \sum_{k=1}^n \mathbb{1}_{\{D_k=1\}} - n(e^{-c} + ce^{-c})}{\sqrt{n}}, \frac{\sum_{k=1}^n D_k - nc}{\sqrt{n}}\right)$$

converges in distribution to a multivariate normal distribution with mean zero and the covariance matrix

$$\begin{pmatrix} e^{-c}(1-e^{-c}) & e^{-c}-e^{-2c}-ce^{-2c} & -ce^{-c} \\ e^{-c}-e^{-2c}-ce^{-2c} & (e^{-c}+ce^{-c})(1-e^{-c}-ce^{-c}) & -c^2e^{-c} \\ -ce^{-c} & -c^{-2}e^{-c} & c \end{pmatrix}.$$

Let us note that

$$\left(\frac{\sum_{k=1}^n \mathbb{1}_{\{D_k=0\}} - ne^{-c}}{\sqrt{n}}, \frac{\sum_{k=1}^n \mathbb{1}_{\{D_k=0\}} + \sum_{k=1}^n \mathbb{1}_{\{D_k=1\}} - n(e^{-c}+ce^{-c})}{\sqrt{n}} \middle| \sum_{k=1}^n D_k = 2\lfloor \tfrac{cn}{2} \rfloor \right)$$

has the same distribution as the random vector

$$\left(\frac{n - X_1^n(0) - X_2^n(0) - n(1-\chi_1(0)-\chi_2(0))}{\sqrt{n}}, \frac{n - X_2^n(0) - n(1-\chi_2(0))}{\sqrt{n}}\right)$$

we are interested in. Stochastic monotonicity then enables us to apply Theorem 41 with $\xi = 0$, and this completes the proof. $\square$

Proof of Theorem 39. This is a direct consequence of Lemma 42 and the fact that initially we have $\lfloor \frac{cn}{2} \rfloor$ edges in the multigraph and the algorithm hasn't made any steps so far, and so $\frac{X_3^n(0) - n\chi_3(0)}{\sqrt{n}} \xrightarrow{p} 0$, and $\mathbb{P}(X_4^n(0) = 0) = 1$. $\square$

2.3.3 The drift function

In order to be able to use the framework for proving the central limit theorem for a Markov chain we have presented in Section 2.3.1, we have to define the appropriate drift function. Recall that in Section 2.3.1 we looked at the appropriately rescaled Markov chain. Thus, our interest lies in properties of the Markov chain

$$\frac{X^n(\lfloor nt \rfloor)}{n}, \quad \text{for } 0 \le t \le \tau^n.$$

Theorem 20 tells us that the for $t < \tau^n$, $\frac{X^n(\lfloor nt \rfloor)}{n}$ is sharply concentrated around its deterministic approximation, and so with high probability as $n \to \infty$ it remains in

$$E = \{x \in (\mathbb{R}_+ \cup \{0\})^4 : \tfrac{\vartheta(t^*)}{2} < z_x < c, x_2 > 0, x_1 + x_2 \le 1, x_1 + 2x_2 \le 2x_3\}, \quad (2.44)$$

where $x = (x_1, x_2, x_3, x_4)$ and z_x is given by (2.8).

Remark 43. *Let us note that the auxiliary process Z^n introduced by (2.9), remains unaffected by the imposed space scaling in the following sense. If we assume that for some $0 < t < \tau^n$, $\frac{X^n(\lfloor nt \rfloor)}{n}$ is in a state $(x_1, x_2, x_3, x_4) \in E$, then $Z^n(\lfloor nt \rfloor)$ is in the state z_x determined by (2.8), as $z_{\frac{X^n(\lfloor nt \rfloor)}{n}} = z_{X^n(\lfloor nt \rfloor)}$.*

Before we define the appropriate drift function, let us recall jump probabilities from Theorem 28, and for $x \in E$ let

$$\begin{aligned} p_1(x) &= \frac{x_1}{2x_3}, \\ p_2(x) &= \frac{x_2 (z_x)^2}{2x_3 f(z_x)}, \\ p_3(x) &= \frac{x_2 z_x}{2x_3}. \end{aligned} \tag{2.45}$$

Let us also recall the set $\mathcal{K}$ of triplets of integers introduced by (2.14). For $(k_1, k_2, k_2) \in \mathcal{K}$ and $l = (-k_1 + k_2, -1 - k_2, -k_1 - k_2 - k_3, 1)$, we define the function β_l by

$$\beta_l(x) = \frac{x_2}{2x_3} \frac{z_x}{f(z_x)} (p_1(x) z_x)^{k_1 - 1} (p_2(x) z_x)^{k_2} (p_3(x) z_x)^{k_3} \cdot \frac{1}{(k_1 - 1)! k_2! k_3!},$$

for $x \in E$. Also for $x \in E$, define

$$\beta_{(-2,0,-1,1)}(x) = \frac{x_1}{2x_3}.$$

For all other $l \in \mathbb{Z}^4$, we let $\beta_l(x) = 0$ for all $x \in E$.

Having in mind Theorem 28, let us note that functions β_l, $l \in \mathbb{Z}^4$, are defined so that $\beta_l(x)$ approximates the (rescaled) rate of transition of $\frac{X^n(n\cdot)}{n}$ from a state $x \in E$ to a state $x + l/n$. Finally, for $x \in E$, we define the drift function by

$$F(x) = \sum_{l \in \mathbb{Z}^4} l \beta_l(x).$$

We have $F(x) = (F_1(x), F_2(x), F_3(x), F_4(x))^T$, where

$$\begin{aligned} F_1(x) &= -1 - \frac{x_1}{2x_3} + \frac{x_2^2 z_x^4 e^{z_x}}{(2x_3 f(z_x))^2} - \frac{x_1 x_2 z_x^2 e^{z_x}}{(2x_3)^2 f(z_x)}, \\ F_2(x) &= -1 + \frac{x_1}{2x_3} - \frac{x_2^2 z_x^4 e^{z_x}}{(2x_3 f(z_x))^2}, \\ F_3(x) &= -1 - \frac{x_2 z_x^2 e^{z_x}}{2x_3 f(z_x)}, \\ F_4(x) &= 1. \end{aligned} \tag{2.46}$$

We can now reveal the nature of function $\chi(t)$ introduced in Section 2.2.1. Namely, $\chi(t)$ is the unique solution to the differential equation $\chi'(t) = F(\chi(t))$ with the initial condition $\chi(0) = (ce^{-c}, 1 - e^{-c} - ce^{-c}, c/2, 0)$. Having in mind that $F(x)$ approximates the expected infinitesimal increment of $\frac{X^n\lfloor nt \rfloor}{n}$ conditioned on being in state x, we see that $\chi(t)$ represents the expected trajectory of $\frac{X^n\lfloor nt \rfloor}{n}$.

We now collect some properties we will find useful later.

Lemma 44. *The functions β_l, $l \in \mathbb{Z}^4$, are continuous in $x \in E$. Also, we have*

$$\sum_{l \in \mathbb{Z}^4} |l|^2 \sup_{x \in E} \beta_l(x) < \infty.$$

Proof. Having in mind that

$$E = \{(x_1, x_2, x_3, x_4)^T \in (\mathbb{R}_+ \cup \{0\})^4 : \tfrac{\vartheta(t^*)}{2} < z_x < c, x_2 > 0, x_1 + x_2 \leq 1, x_1 + 2x_2 \leq 2x_3\},$$

it is straightforward to check that β_l are continuous on E. As $x_3 \geq x_2 > 0$, we have

$$\sup_{x \in E} \frac{x_1 z_x}{2x_3} \leq c, \quad \sup_{x \in E} \frac{x_2 (z_x)^2}{2x_3} \leq c^2,$$

as well as

$$\begin{aligned} \sup_{x \in E} \frac{x_2 (z_x)^3}{2x_3 f(z_x)} &\leq \sup_{x \in E} \frac{x_2}{2x_3} \cdot \frac{(z_x)^3}{\frac{1}{2}(z_x)^2} \leq c, \\ \sup_{x \in E} \frac{x_2}{2x_3} \frac{z_x}{f(z_x)} &\leq \sup_{x \in E} \frac{x_2}{2x_3} \cdot \frac{z_x}{\frac{1}{2}(z_x)^2} \leq \frac{2}{\vartheta(t^*)}, \end{aligned}$$

and this leads us to

$$
\begin{aligned}
&\sum_{l\in\mathbb{Z}^4} |l|^2 \sup_{x\in E} \beta_l(x) \\
&\leq \sum_{(k_1,k_2,k_3)\in\mathcal{K}} \Bigg(\Big((-k_1+k_2)^2 + (-1-k_2)^2 + (-k_1-k_2-k_3)^2 + 1 \Big) \\
&\qquad\qquad\qquad \cdot \sup_{x\in E} \beta_{(-k_1+k_2,-1-k_2,-k_1-k_2-k_3,1)}(x) \Bigg) \\
&\quad + 6 \cdot \sup_{x\in E} \beta_{(-2,0,-1,1)}(x) \\
&\leq \sum_{k_1=1}^{\infty} \sum_{k_2=0}^{\infty} \sum_{k_3=0}^{\infty} \Bigg(\Big((-k_1+k_2)^2 + (-1-k_2)^2 + (-k_1-k_2-k_3)^2 + 1 \Big) \\
&\qquad\qquad\qquad \cdot \frac{2}{\vartheta(t^*)} \frac{c^{k_1-1}}{(k_1-1)!} \cdot \frac{c^{k_2}}{(k_2)!} \cdot \frac{c^{2k_3}}{k_3!} \Bigg) \\
&\quad +6 \\
&< \infty.
\end{aligned}
$$

$\square$

2.3.4 The central limit theorem

We prove the following theorem.

Theorem 45. *Let $t < t^*$. We have*

$$\left(\frac{X^n(\lfloor ns \rfloor) - n\chi(s)}{\sqrt{n}}, 0 \leq s \leq t \right) \xrightarrow{d} (V(s), 0 \leq s \leq t),$$

uniformly as $n \to \infty$, where $V(s) = (V_1(s), V_2(s), V_3(s), V_4(s))$, $0 \leq s \leq t$ is the unique solution to the stochastic differential equation (in integral form)

$$V(s) = V(0) + \sum_{l\in\mathbb{Z}^4} l W_l \Big(\int_0^s \beta_l(\chi(w)) dw \Big) + \int_0^s \partial F(\chi(w)) V(w) dw, \tag{2.47}$$

where W_l are independent Brownian motions, $V(0)$ has the same distribution as the limit in Theorem 39 (independent of everything else), β_l are defined in Section 2.3.3 and F is defined by (2.46).

For the end of Phase 1, as $n \to \infty$ we have the following convergence in distribution

$$\frac{X^n(\tau^n) - n\chi(t^*)}{\sqrt{n}} \xrightarrow{d} V(t^*) - \frac{V_1(t^*)}{F_1(\chi(t^*))} F(\chi(t^*)).$$

We start with the following lemma.

Lemma 46. *The function $\partial F(x) = ((\partial_j F_i(x)))$, $i, j \in \{1, 2, 3, 4\}$, $x \in E$ is continuous.*

Proof. We calculate $\partial F(x)$ in Appendix 2.7.1. As $x_2 > 0$ and $x_2 \leq x_3$ imply $x_3 > 0$, for $x \in E$ we have $x_3 > 0$. Also, for $u > 0$ we have $e^{2u} - 2e^u - u^2 e^u + 1 > 0$ (this is easily seen as the variance of the Poisson random variable with parameter u conditioned on being at least 2 is given by $\frac{u(e^{2u}-2e^u-u^2e^u+1)}{f^2(u)}$). Thus, $\partial F(x)$, $x \in E$ is well-defined and continuous. □

Lemma 47. *The function F satisfies Lipschitz condition (2.30) on E.*

Proof. It is straightforward to check that for $x \in E$ there is an upper bound on $|\partial F(x)|$ (see Appendix 2.7.1 for $\partial F(x)$) and this completes the proof. □

Lemma 48. *The stochastic differential equation (2.47) from Theorem 45 has a unique strong solution.*

Proof. By Lemma 46, $\partial F(x)$, $x \in E$ is a continuous function. As $\chi(t)$, $t \leq t^*$ is continuous, $\partial F(\chi(t))$ is also continuous. Thus, there is a unique solution to

$$\frac{\partial}{\partial t}\Phi(t, s) = \partial F(\chi(t))\Phi(t, s), \quad \Phi(s, s) = I, \tag{2.48}$$

on $0 \leq t \leq t^*$, $0 \leq s \leq t$. Then, by Lemma 35 the unique strong solution is given by

$$V(t) = \Phi(t, 0)V(0) + \int_0^t \Phi(t, s) d\mathcal{W}(s) \tag{2.49}$$

where

$$\mathcal{W}(s) = \sum_{l \in \mathbb{Z}^4} l W_l\Big(\int_0^s \beta_l(\chi(q))dq\Big). \tag{2.50}$$

□

We are now ready to prove Theorem 45.

Proof of Theorem 45. Let us recall that $\mathcal{A}^n$ denotes the event that the maximum degree of the multigraph is at most $\log n$. This event happens with high probability as $n \to \infty$.

For $t < t^*$, let us fix $\epsilon > 0$ such that

$$\sup_{0 \le s \le t} \left| \frac{X^n(\lfloor ns \rfloor)}{n} - \chi(s) \right| < \epsilon$$

implies $\lfloor nt \rfloor < \tau^n \wedge \zeta^n$ and $\frac{X^n(\lfloor ns \rfloor)}{n} \in E$ for $s \le t$ (recall (2.44) for the definition E). For such ϵ, let $\mathcal{B}^n$ denote the event that $\sup_{0 \le s \le t} \left| \frac{X^n(\lfloor ns \rfloor)}{n} - \chi(s) \right| < \epsilon$. Theorem 20 guarantees that the event $\mathcal{B}^n$ happens with high probability as $n \to \infty$.

For the result about the convergence in distribution up to time t, it is enough to work on the event $\mathcal{A}^n \cap \mathcal{B}^n$. On $\mathcal{A}^n \cap \mathcal{B}^n$, for $(k_1, k_2, k_3) \in \mathcal{K}^n$ and $l = (-k_1 + k_2, -1 - k_2, -k_1 - k_2 - k_3, 1)$, Theorem 28 gives that the rate of a jump of $\frac{X^n(\lfloor ns \rfloor)}{n}$, $0 \le s \le t$, from a state $\frac{x}{n} \in E$ to a state $\frac{x+l}{n} \in E$ is given by

$$n\left(\beta_l\left(\frac{x}{n}\right) + E^n_{k_1,k_2,k_3}(\lfloor ns \rfloor)\right)$$

where β_l are defined in Section 2.3.3. Also, by Theorem 28 , $\frac{X^n(\lfloor ns \rfloor)}{n}$, $0 \le s \le t$, jumps from $\frac{x}{n} \in E$ to $\frac{x+(-2,0,-1,1)}{n} \in E$ with the rate

$$n\left(\beta_{(-2,0,-1,1)}\left(\frac{x}{n}\right) + B^n(\lfloor ns \rfloor)\right).$$

Let us now check that conditions of Theorem 36 are fulfilled on this event. Lemma 33 verifies that the condition (2.38) for the errors holds, and so our approximation of the transition rates is precise enough for the result about the convergence under the central limit theorem scaling. Lemma 44 proves the necessary conditions on the functions β_l. Lemma 46 gives us that $\partial F(x)$ is continuous and Lemma 47 that F satisfies Lipschitz condition (2.30). Lemma 48 guarantees the existence of the unique strong solution to the SDE from the statement of the theorem. Finally, Theorem 39 guarantees the convergence in distribution of the initial fluctuations. Thus, on $\mathcal{A}^n \cap \mathcal{B}^n$ the conditions of Theorem 36 hold, and as this event happens with high probability, Theorem 36 gives us

$$\Big(\frac{X^n(\lfloor ns\rfloor) - n\chi(s)}{\sqrt{n}}, 0 \le s \le t\Big) \xrightarrow{d} (V(s), 0 \le s \le t),$$

uniformly as $n \to \infty$.

We now turn to the case of the random time τ^n which denotes the end of Phase 1. Let $\mathcal{D}^n$ denote the event that the process $\frac{X^n(\lfloor n\cdot\rfloor)}{n}$ remains in E all the way up until τ^n. Theorem 20 guarantees the convergence of $\frac{X^n(\lfloor n\cdot\rfloor)}{n}$ to the deterministic approximation all the way up until the end of Phase 1, and so $\mathcal{D}^n$ happens with high probability as $n \to \infty$. Let also $\mathcal{G}^n$ denote the event that $\tau^n \le \zeta^n$. By Remark 21 this event also happens with high probability. Thus, in order to obtain the result about the convergence in distribution we can work on $\mathcal{A}^n \cap \mathcal{D}^n \cap \mathcal{G}^n$. On this event, Lemma 33 verifies that the approximation of the (rescaled) transition rates of $\frac{X^n(\lfloor n\cdot\rfloor)}{n}$ by functions β_l is good enough for the result about the convergence under the central limit theorem scaling, all the way up until the end of Phase 1. The rest of conditions of Theorem 36 are already checked in the proof of the first part of the theorem. As τ^n is itself a random variable, we need to apply the appropriate extension of Theorem 36, stated in Theorem 38. In order to be able to do so, it remains to check that $F_1(\chi(t^*)) < 0$. As already explained, this condition says that when the deterministic function $\chi_1(t)$ hits zero, it does not immediately "come back" to the positive half-plane. This property guarantees that the process $\frac{X_1^n(n\cdot)}{n}$ itself has to hit zero some time close to t^*. Indeed, we have

$$\begin{aligned}
F_1(\chi(t^*)) &= -1 + \frac{(\chi_2(t^*))^2(\vartheta(t^*))^4 e^{\vartheta(t^*)}}{(2\chi_3(t^*)f(\vartheta(t^*)))^2} \\
&= 1 + \frac{(\vartheta(t^*))^2 e^{\vartheta(t^*)}}{(e^{\vartheta(t^*)}-1)^2} \\
&= \Big(\frac{\vartheta(t^*)/2}{\sinh(\vartheta(t^*)/2)}\Big)^2 - 1 \\
&< 0,
\end{aligned}$$

where the second equality comes from (2.10). Again, on the event $\mathcal{A}^n \cap \mathcal{D}^n \cap \mathcal{G}^n$ of high probability all necessary conditions are fulfilled, and this is enough to prove the convergence in distribution at the end of Phase 1. Thus, Theorem 38 completes the proof. □

2.4 The maximum matching number of a random multigraph

In this section, we obtain a central limit theorem for the maximum matching number of the random multigraph with n vertices and $M = \lfloor \frac{cn}{2} \rfloor$ edges for $c > e$.

Theorem 49. *For $c > e$ let ν_c^n denote the size of a maximum matching of the random multigraph with n vertices and $M = \lfloor \frac{cn}{2} \rfloor$ edges. Let γ_1 and γ_2 be the smallest and the largest root, respectively, of the equation $x = c\exp\{-ce^{-x}\}$. As $n \to \infty$,*

$$\frac{\nu_c^n - n\Big(1 - \frac{\gamma_2+\gamma_1+\gamma_2\gamma_1}{2c}\Big)}{\sqrt{n}}$$

converges in distribution to a normal variable with mean 0, and the variance $\sigma(c) > 0$ given by (2.54) below.

Proof. Let us recall that the size of a matching is given by the number of its edges. In each step of the algorithm, we match one edge. Let us recall that the results from [2] give us that for $c > e$ only $\Theta_{\log n}(n^{1/5})$ nodes from the Karp-Sipser core remain unmatched, with high probability as $n \to \infty$. Thus we have

$$\frac{\nu_c^n - (X_4^n(\tau^n) + \frac{1}{2}X_2^n(\tau^n))}{\sqrt{n}} \xrightarrow{p} 0. \tag{2.51}$$

Before we proceed, let us note that, as $\chi_1(t^*) = 0$, we have

$$-\frac{F_4(\chi(t^*))}{F_1(\chi(t^*))} - \frac{1}{2}\frac{F_2(\chi(t^*))}{F_1(\chi(t^*))} = \frac{1}{2}.$$

Hence, as a consequence of Theorem 45, as $n \to \infty$ we have

$$\frac{X_4^n(\tau^n) - n\chi_4(t^*)}{\sqrt{n}} + \frac{1}{2}\cdot\frac{X_2^n(\tau^n) - n\chi_2(t^*)}{\sqrt{n}} \xrightarrow{d} V_4(t^*) + \frac{1}{2}V_2(t^*) + \frac{1}{2}V_1(t^*). \tag{2.52}$$

Let us note that

$$\left(1 - \frac{\gamma_2 + \gamma_1 + \gamma_2\gamma_1}{2c}\right) = \chi_4(t^*) + \frac{1}{2}\chi_2(t^*),$$

(see Appendix 2.7.3). Thus, (2.51) and (2.52) give us

$$\frac{\nu_c^n - n\left(1 - \frac{\gamma_2+\gamma_1+\gamma_2\gamma_1}{2c}\right)}{\sqrt{n}} \xrightarrow{d} V_4(t^*) + \frac{1}{2}V_2(t^*) + \frac{1}{2}V_1(t^*).$$

Let us now recall Lemma 48 which characterizes the limiting distribution. In particular, by (2.49) we have

$$V(t^*) = \Phi(t^*, 0)V(0) + \int_0^{t^*} \Phi(t^*, s)d\mathcal{W}(s), \tag{2.53}$$

where $\mathcal{W}$ is a process defined by (2.50), independent of $V(0)$. Having in mind that $V(0)$ has multidimensional normal distribution, $V(t^*)$ also has multidimensional normal distribution. Thus, the limiting random variable $V_4(t^*) + \frac{1}{2}V_2(t^*) + \frac{1}{2}V_1(t^*)$ is normal with mean zero and the variance

$$\begin{aligned}\sigma(c) = &\frac{1}{4}\mathrm{Var}(V_1(t^*)) + \frac{1}{4}\mathrm{Var}(V_2(t^*)) + \mathrm{Var}(V_4(t^*))\\ &+ \frac{1}{2}\mathrm{Cov}(V_1(t^*), V_2(t^*)) + \mathrm{Cov}(V_1(t^*), V_4(t^*)) + \mathrm{Cov}(V_2(t^*), V_4(t^*)).\end{aligned} \tag{2.54}$$

In Appendix 2.7.2 we check that $\sigma(c) > 0$. □

2.5 The maximum matching number of $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ for $c > e$

Theorem 49 provides the central limit theorem for the maximum matching number of a random multigraph with n nodes and $M = \lfloor \frac{cn}{2} \rfloor$ edges for $c > e$, as $n \to \infty$. In order to carry over the result about the convergence in distribution to the simple graph, we need a careful refinement. We will use the approach taken in Janson and Luczak [29] and prove that the distribution of the suitably rescaled maximum matching number is asymptotically independent of the event that the graph is simple. Loosely speaking, as the total number of half-edges taking part in forming loops and multiple edges converges in distribution (recall Lemma 24), we need to show that fixing a number of loops and multiple edges does not affect the distribution of the rescaled maximum matching number.

Our main tool in this section will be Proposition 7.1 from [29] which we now state.

Proposition 50 (Janson and Luczak [29], Proposition 7.1)**.** *Assume that A^n, $n \geq 1$ and A are random variables, and that for each n, $(I^n_\alpha)_{\alpha\in\mathfrak{I}^n}$, is a finite family of indicator random variables defined on the same probability space as A^n. Let $\Lambda^n = \sum_{\alpha\in\mathfrak{I}^n} I^n_\alpha$ and let $\mathcal{E}^n$ be the event*

$$\mathcal{E}^n = \{I^n_\alpha = 0 \textit{ for every } \alpha \in \mathfrak{I}^n\} = \{\Lambda^n = 0\}.$$

Assume that as $n \to \infty$ we have

1) $A^n \xrightarrow{d} A$,

2) For any fixed $l \geq 1$, if we for each n select l indices $\alpha^n_1, \ldots \alpha^n_l \in \mathfrak{I}^n$ such that we have $\mathbb{P}\Big(I^n_{\alpha^n_1} = \cdots = I^n_{\alpha^n_l} = 1\Big) > 0$, then

$$\Big(A^n | I^n_{\alpha^n_1} = \cdots = I^n_{\alpha^n_l} = 1\Big) \xrightarrow{d} A,$$

3) $\Lambda^n \xrightarrow{d} \Lambda$ *where Λ is a random variable such that $\mathbb{P}(\Lambda = 0) > 0$ and the distribution of Λ is determined by its moments,*

4) $\limsup_{n\to\infty} \mathbb{E}\big[(\Lambda^n)^j\big] < \infty$ *for each $j \geq 1$.*

Then, $A^n | \mathcal{E}^n \xrightarrow{d} A$.

In order to be in the setting of Proposition 50, we introduce the following. Let $\mathfrak{I}^n_1$ be the family of the elements of the form $(\{h_1, h_2\}, v)$, where $\{h_1, h_2\}$ is an unordered pair of half-edges, and v is a vertex in a multigraph with n vertices and $M = \lfloor \frac{cn}{2} \rfloor$ edges. Let also $\mathfrak{I}^n_2$ be the family of the elements of the form $(h_1, h_2, v, h_3, h_4, u)$ where h_1, h_2, h_3, h_4 are half-edges and v and u are distinct vertices. For $\alpha = (\{h_1, h_2\}, v) \in \mathfrak{I}^n_1$, let I^n_α be the event that h_1 and h_2 are half-edges which form a loop on v. Similarly, for $\alpha = (h_1, h_2, v, h_3, h_4, u) \in \mathfrak{I}^n_2$, let I^n_α be the event that h_1 and h_2 are adjacent to v, h_3 and h_4 are adjacent to u, and this two pairs of half-edges form an edge of multiplicity 2 between v and u (that is, we have either that h_1 is paired with h_3 and h_2 is paired with h_4, or that h_1 is paired with h_4 and h_2 is paired with h_3).

Lemma 51. *For a fixed $l \geq 1$, let us select l indices $\alpha_1^n, \cdots, \alpha_l^n \in \mathfrak{I}_1^n \cup \mathfrak{I}_2^n$ such that $P(I_{n\alpha_1^n} = \cdots = I_{n\alpha_l^n} = 1) > 0$. Let $\nu^n_{c,\alpha_1^n,\cdots,\alpha_l^n}$ be the maximum matching number of a random multigraph with n nodes and $M = \lfloor \frac{cn}{2} \rfloor$ edges conditioned on having the edges corresponding to indices $\alpha_1^n \cdots, \alpha_l^n$ fixed. Then, as $n \to \infty$*

$$\frac{\nu^n_{c,\alpha_1^n,\cdots,\alpha_l^n} - n\Big(1 - \frac{\gamma_2+\gamma_1+\gamma_2\gamma_1}{2c}\Big)}{\sqrt{n}}$$

converges in distribution to the limiting random variable from Theorem 49.

Proof. Let us note that each element of $\mathfrak{I}_1^n$ determines one edge in the mutigraph, and each element of $\mathfrak{I}_2^n$ determines two edges. Thus, we are conditioning on having k edges fixed, where $l \leq k \leq 2l$. Let $\nu^n(M-l)$ denote the maximum matching number of a random multigraph with n nodes and $M-l$ edges. We have

$$\frac{\nu^n(M-2l)}{\sqrt{n}} \prec_{\mathrm{sd}} \frac{\nu^n_{c,\alpha_1^n,\cdots,\alpha_l^n}}{\sqrt{n}} \prec_{\mathrm{sd}} \frac{\nu^n(M)+2l}{\sqrt{n}},$$

where $\prec_{\mathrm{sd}}$ denotes stochastic domination. This arises as follows. Firstly, let us note that if we remove fixed k edges from a random multigraph with n nodes and M edges (leaving nodes intact), the resulting graph equals in distribution a random multigraph with n nodes and $M-k$ edges. Having in mind that $l \leq k \leq 2l$, we have $\nu^n(M-2l) \prec_{\mathrm{sd}} \nu^n_{c,\alpha_1^n,\cdots,\alpha_l^n}$. The removal of a single edge can decrease the maximum matching number by at most one, and so $\nu^n_{c,\alpha_1^n,\cdots,\alpha_l^n} \prec_{\mathrm{sd}} \nu^n(M)+2l$. It remains to note that for a fixed l, the difference between $\nu^n(M)$ and $\nu^n(M-2l)$ for $M = \lfloor \frac{cn}{2} \rfloor$ is not visible on order $\sqrt{n}$, and so $\frac{\nu^n(M)+2l-\nu^n(M-2l)}{\sqrt{n}} \xrightarrow{d} 0$. A Sandwich theorem-type argument and Theorem 49 complete the proof. $\square$

Proof of Theorem 19 i). We will apply Proposition 50 to the sequence of random variables

$$\frac{\nu^n_c - n\Big(1 - \frac{\gamma_2+\gamma_1+\gamma_2\gamma_1}{2c}\Big)}{\sqrt{n}},$$

and the family $(I^n_{\alpha^n})_{\alpha^n \in \mathfrak{J}^n}$, for $\mathfrak{J}^n = \mathfrak{J}^n_1 \cup \mathfrak{J}^n_2$. By Theorem 49, we have that Condition 1) of Proposition 50 is fulfilled. Condition 2) holds as a consequence of Lemma 51. Conditions 3) and 4) are fulfilled as Lemma 24 gives us that $\sum_{\alpha \in \mathfrak{J}^n} I^n_{\alpha^n}$ converges in distribution to a Poisson random variable, and that all moments converge to the moments of the limiting Poisson random variable. Then, Proposition 50 gives us that conditioning on the event $\sum_{\alpha^n \in \mathfrak{J}^n} I^n_{\alpha^n} = 0$ does not change the asymptotic distribution of the sequence of random variables we are considering. It remains to note that $\sum_{\alpha^n \in \mathfrak{J}^n} I^n_{\alpha^n} = 0$ is equivalent to the multigraph being simple. Such a multigraph equals in distribution to the Erdős-Rényi random graph $G(n, M = \lfloor \frac{cn}{2} \rfloor)$. This completes the proof. □

2.6 The maximum matching number of $G(n, p = \frac{c}{n})$ for $c > e$

In the previous section, we proved a central limit theorem for the maximum matching number of the Erdős-Rényi $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ random graph for $c > e$. Results from Pittel [45] enable us to translate this result to the $G(n, p = \frac{c}{n})$ model with only modest additional work.

Proposition 52 (Pittel [45], Lemma 2)**.** *Let $\mathfrak{Q}^n$ be a random variable defined on the same probability space as $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ and suppose that for some functions $m(c)$, $\theta(c)$, $c > 0$, the random variable $\frac{\mathfrak{Q}^n - nm(c)}{\sqrt{n}}$ converges in distribution to a normal with mean zero and the variance $\theta(c)$ as $n \to \infty$. Assume also that $m(c)$ is continuously differentiable and $\theta(c)$ is continuous for $c > 0$. Then, for the model $G(n, p = \frac{c}{n})$, the random variable $\frac{\mathfrak{Q}^n - nm(c)}{\sqrt{n}}$, defined on the same probability space, converges in distribution to a normal with mean zero and the variance*

$$2c(m'(c))^2 + \theta(c).$$

We start with the following lemma.

Lemma 53. *For $c > e$, γ_2 and γ_1, the smallest and the largest root, respectively, of the equation $x = c \exp\{-ce^{-x}\}$, are continuously differentiable as functions of c.*

Proof. As we have $\gamma_1 = ce^{-\gamma_2}$ and $\gamma_2 = ce^{-\gamma_1}$ (see (49) in [2]), we get

$$\frac{d\gamma_1}{dc} = e^{-\gamma_2} \cdot \frac{1-\gamma_2}{1-\gamma_1\gamma_2},$$

and

$$\frac{d\gamma_2}{dc} = e^{-\gamma_1} - ce^{-(\gamma_1+\gamma_2)} \cdot \frac{1-\gamma_2}{1-\gamma_1\gamma_2}.$$

As noted in [2], for $c > e$ the relation $\gamma_2 e^{-\gamma_2} = \gamma_1 e^{-\gamma_1}$ implies $\gamma_1 < 1 < \gamma_2$ and $\gamma_1\gamma_2 < 1$, and so the derivatives are well defined and continuous. $\square$

Lemma 54. *The variance $\sigma(c)$ given by (2.54) is a continuous function of c for $c > e$.*

Proof. Having in mind that $\vartheta(t^*) = \gamma_2 - \gamma_1$ (see (48) in [2]), we get that $\vartheta(t^*)$ is a continuous function of c, and so $\chi_1(t^*)$, $\chi_2(t^*)$, $\chi_3(t^*)$, $\chi_4(t^*)$, $F_1(\chi(t^*))$, $F_2(\chi(t^*))$ and $F_4(\chi(t^*))$ are as well. The continuity of $\vartheta(t^*)$ together with the continuity of ∂F we proved in Lemma 46, gives us that $\Phi(t^*, 0)$ determined by (2.48) is continuous as a function of $c > e$. Thus, having in mind (2.53), the covariance matrix of $V(t^*)$ is continuous as a function of $c > e$, and this completes the proof. $\square$

Proof of Theorem 19 ii). Proposition 52 provides the conditions which enable one to carry over the results about the convergence in distribution from the model $G(n, M = \lfloor \frac{cn}{2} \rfloor)$ to $G(n, p = \frac{c}{n})$. Lemma 53 gives that the function $1 - \frac{\gamma_2+\gamma_1+\gamma_2\gamma_1}{2c}$ is continuously differentiable for $c > e$. Lemma 54 proves that the variance $\sigma(c)$, $c > e$ is a continuous function of c. Having in mind our Theorem 19 i), conditions of Proposition 52 are fulfilled. Let us recall that $1 - \frac{\gamma_2+\gamma_1+\gamma_2\gamma_1}{2c} = \chi_4(t^*) + \frac{1}{2}\chi_2(t^*)$ (see Appendix 2.7.3). By Proposition 52,

$$\frac{\tilde{\mu}_c^n - n\left(1 - \frac{\gamma_2+\gamma_1+\gamma_2\gamma_1}{2c}\right)}{\sqrt{n}}$$

converges in distribution to a normal random variable with mean 0 and the variance

$$\tilde{\sigma}(c) = 2c \cdot \left(\frac{\partial}{\partial c}\chi_4(t^*) + \frac{\partial}{\partial c}\frac{\chi_2(t^*)}{2}\right)^2 + \sigma(c). \tag{2.55}$$

As $\sigma(c) > 0$ (see Appendix 2.7.2), we have $\tilde{\sigma}(c) > 0$ for $c > e$. $\square$

2.7 Appendix

2.7.1 Calculating $\partial F(x)$

For the purpose of the proof of Lemma 46, here we perform the straightforward but tedious calculation of $\partial F(x)$. We have

$$\begin{aligned}\frac{\partial F_1}{\partial x_1} =& -\frac{1}{2x_3} - \frac{x_2}{4x_3^2} \cdot \frac{4z_x^3e^{2z_x} - z_x^4e^{2z_x} - 3z_x^4e^{z_x} - 4z_x^3e^{z_x} - z_x^5e^{z_x}}{f(z_x)(e^{2z_x} - 2e^{z_x} - z_x^2e^{z_x} + 1)} \\ &- \frac{x_2z_x^2e^{z_x}}{4x_3^2f(z)} + \frac{x_1z_xe^{z_x}(2e^{z_x} - 2z_x - z_x^2 - 2)}{4x_3^2(e^{2z_x} - 2e^{z_x} - z_x^2e^{z_x} + 1)},\end{aligned}$$

$$\begin{aligned}\frac{\partial F_1}{\partial x_2} =& \frac{x_2z_x^4e^z}{2x_3^2f^2(z)} - \frac{(2x_3 - x_1)(-z_x^4e^{2z_x} + 4z_x^3e^{2z_x} - z_x^5e^{z_x} - 3z_x^4e^{z_x} - 4z_x^3e^{z_x})}{4x_3^2f(z_x)(e^{2z_x} - 2e^{z_x} - z_x^2e^{z_x} + 1)} \\ &- \frac{x_1z_x^2e^{z_x}}{4x_3^2f(z_x)} - \frac{x_1(x_1 - 2x_3)z_xe^{z_x}(2e^{z_x} - z_x^2 - 2z_x - 2)}{4x_3^2x_2(e^{2z_x} - 2e^{z_x} - z_x^2e^{z_x} + 1)},\end{aligned}$$

$$\begin{aligned}\frac{\partial F_1}{\partial x_3} =& \tfrac{x_1}{2x_3^2} + \frac{x_2(-z_x^4e^{2z_x} + 4z_x^3e^{2z_x} - z_x^5e^{z_x} - 3z_x^4e^{z_x} - 4z_x^3e^{z_x})}{2x_3^2f(z_x)(e^{2z_x} - 2e^{z_x} - z_x^2e^{z_x} + 1)} \\ &- \frac{x_2^2z_x^4e^{z_x}}{2x_3^3f^2(z_x)} + \frac{2x_1x_2z_x^2e^{z_x}}{4x_3^3f(z_x)} - \tfrac{x_1z_xe^{z_x}(2e^{z_x} - z_x^2 - 2z_x - 2)}{2x_3^2(e^{2z_x} - 2e^{z_x} - z_x^2e^{z_x} + 1)},\end{aligned}$$

$$\frac{\partial F_1}{\partial x_4} = 0,$$

$$\frac{\partial F_2}{\partial x_1} = \frac{1}{2x_3} + \frac{x_2}{4x_3^2} \cdot \frac{4z_x^3e^{2z_x} - z_x^4e^{2z_x} - 3z_x^4e^{z_x} - 4z_x^3e^{z_x} - z_x^5e_x^z}{f(z)(e^{2z_x} - 2e^{z_x} - z_x^2e^{z_x} + 1)},$$

$$\frac{\partial F_2}{\partial x_2} = -\frac{x_2z_x^4e^{z_x}}{2x_3^2f^2(z_{z_x})} + \frac{(2x_3 - x_1)(-z_x^4e^{2z_x} + 4z_x^3e^{2z_x} - z_x^5e^{z_x} - 3z_x^4e^{z_x} - 4z_x^3e^{z_x})}{4x_3^2f(z_x)(e^{2z_x} - 2e^{z_x} - z_x^2e^{z_x} + 1)},$$

$$\frac{\partial F_2}{\partial x_3} = -\frac{x_1}{2x_3^2} - \frac{x_2(-z_x^4e^{2z_x} + 4z_x^3e^{2z_x} - z_x^5e^{z_x} - 3z_x^4e^{z_x} - 4z_x^3e^{z_x})}{2x_3^2f(z)(e^{2z_x} - 2e^{z_x} - z_x^2e^{z_x} + 1)} + \frac{x_2^2z_x^4e^{z_x}}{2x_3^3f^2(z_x)},$$

$$\frac{\partial F_2}{\partial x_4} = 0,$$

$$\frac{\partial F_3}{\partial x_1} = \frac{z_x e^{z_x}(2e^{z_x} - 2z_x - z_x^2 - 2)}{2x_3(e^{2z_x} - 2e^{z_x} - z_x^2 e^{z_x} + 1)},$$

$$\frac{\partial F_3}{\partial x_2} = -\frac{z_x^2 e^{z_x}}{2x_3 f(z_x)} - \frac{(x_1 - 2x_3)z_x e^{z_x}(2e^{z_x} - z_x^2 - 2z_x - 2)}{2x_3 x_2(e^{2z_x} - 2e^{z_x} - z_x^2 e^{z_x} + 1)},$$

$$\frac{\partial F_3}{\partial x_3} = -\frac{z_x e^{z_x}(2e^{z_x} - z_x^2 - 2z_x - 2)}{x_3(e^{2z_x} - 2e^{z_x} - z_x^2 e^{z_x} + 1)} + \frac{x_2 z_x^2 e^{z_x}}{2x_3^2 f(z_x)},$$

$$\frac{\partial F_4}{\partial x_1} = \frac{\partial F_4}{\partial x_2} = \frac{\partial F_4}{\partial x_3} = \frac{\partial F_4}{\partial x_4} = 0.$$

2.7.2 The variance $\sigma(c)$ is greater than zero

The purpose of this section is to prove the following lemma.

Lemma 55. *The variance $\sigma(c)$ of the limiting random variable from Theorem 49, given by (2.54), is strictly positive for $c > e$.*

Proof. Let us recall that the limiting random variable from Theorem 49 is given by $V_4(t^*) + \frac{1}{2}V_2(t^*) + \frac{1}{2}V_1(t^*)$. This random variable has normal distribution with mean zero and variance $\sigma(c)$. Thus, in order to complete the proof, it is enough to prove that $V_4(t^*) + \frac{1}{2}V_2(t^*) + \frac{1}{2}V_1(t^*)$ is not equal in distribution to zero.

Let us recall that by (2.49), we have

$$V(t^*) = \Phi(t^*, 0)V(0) + \int_0^{t^*} \Phi(t^*, s)d\mathcal{W}(s),$$

where $\mathcal{W}$ is a process defined by (2.50), independent of $V(0)$. Let $A(t^*) = \Phi(t^*, 0)V(0)$ and $B(t^*) = \int_0^{t^*} \Phi(t^*, s)d\mathcal{W}(s)$. The independence of $V(0)$ and $\mathcal{W}$ implies that $A(t^*) = (A_1(t^*), A_2(t^*), A_3(t^*), A_4(t^*))^T$ and $B(t^*) = (B_1(t^*), B_2(t^*), B_3(t^*), B_4(t^*))^T$ are independent. Then,

$$V_4(t^*) + \frac{1}{2}V_2(t^*) + \frac{1}{2}V_1(t^*) = A_4(t^*) + \frac{1}{2}A_2(t^*) + \frac{1}{2}A_1(t^*) + B_4(t^*) + \frac{1}{2}B_2(t^*) + \frac{1}{2}B_1(t^*),$$

and having in mind that $A(t^*)$ and $B(t^*)$ are independent, $V_4(t^*)+\frac{1}{2}V_2(t^*)+\frac{1}{2}V_1(t^*) =_d 0$ is equivalent to both

$$A_4(t^*) + \frac{1}{2}A_2(t^*) + \frac{1}{2}A_1(t^*) =_d 0$$

and

$$B_4(t^*) + \frac{1}{2}B_2(t^*) + \frac{1}{2}B_1(t^*) =_d 0.$$

Hence, it is enough to show that $A_4(t^*) + \frac{1}{2}A_2(t^*) + \frac{1}{2}A_1(t^*)$ is not equal to zero in distribution. Now, as $A(t^*) = \Phi(t^*, 0)V(0)$, we have

$$\begin{aligned}
A_1(t^*) &= \Phi_{11}(t^*, 0)V_1(0) + \Phi_{12}(t^*, 0)V_2(0) + \Phi_{13}(t^*, 0)V_3(0) + \Phi_{14}(t^*, 0)V_4(0), \\
A_2(t^*) &= \Phi_{21}(t^*, 0)V_1(0) + \Phi_{22}(t^*, 0)V_2(0) + \Phi_{23}(t^*, 0)V_3(0) + \Phi_{24}(t^*, 0)V_4(0), \\
A_4(t^*) &= \Phi_{41}(t^*, 0)V_1(0) + \Phi_{42}(t^*, 0)V_2(0) + \Phi_{43}(t^*, 0)V_3(0) + \Phi_{44}(t^*, 0)V_4(0).
\end{aligned}$$

Let us note that $\frac{\partial F_4}{\partial x_1} = \frac{\partial F_4}{\partial x_2} = \frac{\partial F_4}{\partial x_3} = \frac{\partial F_4}{\partial x_4} = 0$ implies $\Phi_{41}(t^*, 0) = \Phi_{42}(t^*, 0) = \Phi_{43}(t^*, 0) = \Phi_{44}(t^*, 0) = 0$, and so $A_4(t^*) =_d 0$. Also, recall that in Section 2.3.2 we get $V_3(0) =_d 0$ and $V_4(0) =_d 0$. Thus, it suffices to show that

$$(\Phi_{11}(t^*, 0) + \Phi_{21}(t^*, 0))V_1(0) + (\Phi_{12}(t^*, 0) + \Phi_{22}(t^*, 0))V_2(0),$$

is not equal to zero in distribution. As we have $(\Phi_{11}(t^*, 0) + \Phi_{21}(t^*, 0)) \neq 0$ and $(\Phi_{12}(t^*, 0) + \Phi_{22}(t^*, 0)) \neq 0$, this is equivalent to proving that there is no event of probability 1 such that

$$V_2(0) = -\frac{(\Phi_{11}(t^*, 0) + \Phi_{21}(t^*, 0))}{(\Phi_{12}(t^*, 0) + \Phi_{22}(t^*, 0))}V_1(0).$$

However, if this were the case we would have

$$\mathrm{Cov}(V_1(0), V_2(0)) = -\frac{(\Phi_{11}(t^*, 0) + \Phi_{21}(t^*, 0))}{(\Phi_{12}(t^*, 0) + \Phi_{22}(t^*, 0))}\mathrm{Var}(V_1(0)),$$

as well as

$$\mathrm{Var}(V_2(0)) = \left(\frac{(\Phi_{11}(t^*, 0) + \Phi_{21}(t^*, 0))}{(\Phi_{12}(t^*, 0) + \Phi_{22}(t^*, 0))}\right)^2 \mathrm{Var}(V_1(0)),$$

and it is straightforward to check that this does not hold by looking at the covariance matrix Σ from Theorem 39. $\square$

2.7.3 On the fluid limit approximation of the maximum matching number

In this section, we verify that

$$\left(1 - \frac{\gamma_2 + \gamma_1 + \gamma_2\gamma_1}{2c}\right) = \chi_4(t^*) + \frac{1}{2}\chi_2(t^*).$$

In [2], it is proved that

$$\begin{aligned}
\vartheta(t^*) &= \gamma_2 - \gamma_1, \\
\gamma_2 &= c\beta(\vartheta(t^*)), \\
\gamma_1 &= c\beta(\vartheta(t^*))e^{-\vartheta(t^*)},
\end{aligned}$$

(see (48) there). Consequently, we have

$$\begin{aligned}
\chi_4(t^*) + \frac{1}{2}\chi_2(t^*) &= \frac{1}{c}(c(1 - \beta(\vartheta(t^*))) - \frac{1}{2}\log^2\beta(\vartheta(t^*))) \\
&\quad + \frac{1}{2}(1 - \vartheta(t^*)e^{-\vartheta(t^*)} - e^{-\vartheta(t^*)})\beta(\vartheta(t^*)) \\
&= \frac{1}{c}(c - \gamma_2 - \frac{1}{2}\gamma_1^2) + \frac{1}{2}(1 - \vartheta(t^*)e^{-\vartheta(t^*)} - e^{-\vartheta(t^*)})\beta(\vartheta(t^*)) \\
&= \frac{1}{c}(c - \gamma_2 - \frac{1}{2}\gamma_1^2) + \frac{1}{2c}(\gamma_2 - \gamma_2 \cdot \vartheta(t^*) \cdot e^{-\vartheta(t^*)} - \gamma_1) \\
&= 1 - \frac{\gamma_2 + \gamma_1 + \gamma_2\gamma_1}{2c}.
\end{aligned}$$

Chapter 3

The spread of fire on a random multigraph

Abstract

We study a model for the destruction of a random network by fire. Suppose that we are given a multigraph of minimum degree at least 2 having real-valued edge-lengths. We pick a uniform point from along the length and set it alight; the edges of the multigraph burn at speed 1. If the fire reaches a vertex of degree 2, the fire gets directly passed on to the neighbouring edge; a vertex of degree at least 3, however, passes the fire either to all of its neighbours or none, each with probability 1/2. If the fire goes out before the whole network is burnt, we again set fire to a uniform point. We are interested in the number of fires which must be set in order to burn the whole network, and the number of points which are burnt from two different directions. We analyse these quantities for a random multigraph having n vertices of degree 3 and $\alpha(n)$ vertices of degree 4, where $\alpha(n)/n \to 0$ as $n \to \infty$, with i.i.d. standard exponential edge-lengths. Depending on whether $\alpha(n) >> \sqrt{n}$ or not, we prove that as $n \to \infty$ these quantities converge jointly in distribution when suitably rescaled to either a pair of constants or to (complicated) functionals of Brownian motion.

We use our analysis of this model to make progress towards a conjecture of Aronson, Frieze and Pittel concerning the number of vertices which remain unmatched when we use the Karp-Sipser algorithm to find a matching on the Erdős-Rényi random graph.

Keywords: random graphs, configuration model, exploration process, Markov chain, fluid limit method, diffusion limit, Karp-Sipser algorithm.

3.1 Introduction

The content of this chapter is based on joint work with my supervisor Professor Christina Goldschmidt. We would like to thank Professor Ruth Williams for directing us to her paper [33] which very much facilitated our proof of the invariance principle.

3.1.1 The model

Suppose that we are given a finite connected multigraph with strictly positive real-valued edge-lengths. We introduce a model for the destruction of such a network by fire. The edges are flammable. First, a point is picked uniformly (i.e. according to the normalised Lebesgue measure) and set alight. (With probability 1, this point will lie in the interior of an edge.) The fire passes at speed 1 along the edge (in both directions) until it reaches a node. A node of degree $d \geq 3$ will pass the fire onto all of its other neighbouring edges with probability 1/2 or stop the fire with probability 1/2. If it stops the fire, it becomes a vertex of degree $d-1$. A vertex of degree 2 necessarily passes a fire arriving along from one of its neighbouring edges onto the other one. The fire spreads until it either goes out or has burnt the whole network. If it goes out before the whole network is burnt, a new uniform point is picked and set alight, and the process continues as before.

We are interested in two aspects of this process:

1. How many new fires must be set in order to burn the whole network?

2. How many times does it happen that a point is burnt from two different directions?

We refer to the second phenomenon as a *clash*.

Let $\alpha : \mathbb{N} \to \mathbb{N}$ be a function such that $\alpha(n)/n \to 0$ as $n \to \infty$. We will study these questions in the setting where the base network is a random multigraph with n vertices of degree 3 and $\alpha(n)$ vertices of degree 4, sampled according to the configuration model we have introduced in Chapter 1. *Throughout this chapter, we will implicitly assume*

$3n + 4\alpha(n)$ *to be even.* We take the edge-lengths to be independent and identically distributed standard exponential random variables.

Let F^n be the number of fires we must set in order to burn the whole network. Let C^n be the number of clashes. We will study the limiting behaviour of these quantities as $n \to \infty$. It turns out that both scale as $\sqrt{n}$ as long as $\alpha(n) = O(\sqrt{n})$ and otherwise as $\alpha(n)$. In order to state our results more precisely, we introduce an auxiliary stochastic process.

Let $(B_t)_{t\geq 0}$ be a standard Brownian motion and for $a \geq 0$ let $(X_t^a, L_t^a)_{0\leq t\leq 1}$ be the unique solution to the stochastic differential equation with reflection determined by

$$dX_t^a = -\frac{2}{3}\frac{X_t^a}{(1-t)}dt + \frac{2a}{3}(1-t)^{1/3}dt + dB_t + dL_t^a, \quad 0 \leq t < 1, \tag{3.1}$$

where $(L_t^a)_{0\leq t<1}$ is the local time process of X^a at level 0. This solution may be explicitly written as a functional of the Brownian motion as follows. First define

$$K_t^a = -\inf_{0\leq s\leq t}\left[a - a(1-s)^{2/3} + \int_0^s \frac{dB_u}{(1-u)^{2/3}}\right], \quad 0 \leq t < 1.$$

Then set

$$\begin{aligned} L_t^a &= \int_0^t (1-s)^{2/3} dK_s^a, \quad 0 \leq t < 1, \\ X_t^a &= a(1-t)^{2/3} - a(1-t)^{4/3} + (1-t)^{2/3}\int_0^t \frac{dB_u}{(1-u)^{2/3}} + (1-t)^{2/3}K_t^a, \quad 0 \leq t < 1. \end{aligned}$$

We show in the sequel that both of these quantities have finite almost sure limits as $t \to 1$, which we call L_1^a and X_1^a.

Theorem 56. *(i) Suppose that $\alpha(n)/\sqrt{n} \to a$ as $n \to \infty$, where $a \geq 0$. Then, as $n \to \infty$,*

$$\frac{1}{\sqrt{n}}(F^n, C^n) \xrightarrow{d} \left(\frac{1}{2}L_1^a, \int_0^1 \frac{X_s^a}{3(1-s)}ds\right),$$

where the limiting random variables are almost surely finite.

(ii) Suppose that $\alpha(n) >> \sqrt{n}$*. Then, as* $n \to \infty$*,*

$$\frac{1}{\alpha(n)} F^n \xrightarrow{p} 0, \quad \frac{1}{\alpha(n)} C^n \xrightarrow{p} \frac{1}{4}.$$

3.1.2 Motivation: the Karp-Sipser algorithm on a random graph

In Chapter 1 we introduced the Karp-Sipser algorithm, a greedy algorithm for obtaining a matching proposed by Karp and Sipser [34], and commented on its performance on certain classes of sparse random graphs. We provided a detailed overview of the performance of the Karp-Sipser algorithm on the Erdős–Rényi random graphs $G(N, M = \lfloor \frac{cN}{2} \rfloor)$ and $G(N, p = \frac{c}{N})$, $c > 0$. The problem we study in this chapter is motivated by certain questions related to the performance of the algorithm on this model.

We start by briefly recalling how the Karp-Sipser algorithm works. If there is at least one pendant vertex in the graph, i.e. a vertex of degree 1, choose one uniformly at random and include the edge adjacent to it in the matching. Remove this edge, the two vertices that form it and any other edges adjacent to them. If, on the other hand, there are no pendant vertices in the graph, choose one of the existing edges uniformly at random, and include it in the matching. Remove the chosen edge together with the two vertices that form it, as well as any other edges adjacent to those vertices. Now repeat the procedure on the resulting graph, until there are no edges remaining.

Suppose that we take the graph to be the Erdős–Rényi random graph $G(N, M)$, with N vertices and M edges, where $M = \lfloor cN/2 \rfloor$ and $c > 0$ is a constant. Let D_N be the difference between the size of a maximum matching on $G(N, M)$ and the matching produced by the Karp-Sipser algorithm. Let A_N be the number of vertices remaining in the graph at the point of the first uniform random choice which are left unmatched by the Karp-Sipser algorithm. Then $D_N \leq A_N/2$. Aronson, Frieze and Pittel [2] proved the following theorem.

Theorem 57. *(i) If* $c < e$*, then* $\mathbb{P}(D_N = 0) \to 1$ *as* $N \to \infty$*.*

(ii) If $c > e$*, there exist constants* $C_1, C_2 > 0$ *such that*

$$C_1 N^{1/5} / (\log N)^{75/2} \leq \mathbb{E}[A_N] \leq C_2 N^{1/5} (\log N)^{12}.$$

Aronson, Frieze and Pittel conjecture that, in fact, $N^{-1/5}\mathbb{E}[A_N]$ converges as $N \to \infty$; indeed, one might also reasonably conjecture that $N^{-1/5}A_N$ possesses a limit in distribution as $N \to \infty$. One of our aims is to make progress towards understanding how such a limit in distribution might arise.

Recall that the Karp-Sipser algorithm proceeds in two phases. In the first phase (Phase 1), the algorithm recursively attacks the pendant subtrees in the graph, matching as it goes. Phase 2 starts the first time that the algorithm is forced to pick a uniform edge. At the start of Phase 2, the graph necessarily contains only vertices of degree 2 or more. It will, in general, have several components, of which some may consist of isolated cycles. In these cycle components, Karp-Sipser necessarily yields a maximum matching. So the source of the "mistakes" is the complex components of the graph present at the start of Phase 2. Our original motivation for introducing the model studied in this chapter is to understand the behaviour of the Karp-Sipser algorithm on the type of complex components appearing in Phase 2.

For $c < e$, Phase 1 is essentially the whole story, apart from a few isolated cycles (which cannot cause "mistakes"). On the other hand, for $c > e$, a non-trivial graph remains at the end of Phase 1. The work of Aronson, Frieze and Pittel [2] suggests that, for $c > e$, the part of the Karp-Sipser process which gives the dominant contribution to A_N arises close to the end of Phase 2. Their analysis indicates the following heuristic picture for the structure of the graph at this point: ignoring log-corrections, it is approximately a uniform random graph with $\Theta(N^{3/5})$ vertices of degree 2, $\Theta(N^{2/5})$ vertices of degree 3 and $\Theta(N^{1/5})$ vertices of degree 4, with high probability as $N \to \infty$. In a moment, we will describe the structure of this graph.

Before going any further, let us recall the configuration model [7, 12, 54, 55, 56] we have introduced in Chapter 1. Fix n and a sequence $\mathbf{d} = (d_1, d_2, \ldots, d_n)$, where $d_i \geq 0$ is the degree of vertex i and $\sum_{i=1}^n d_i$ is even. Suppose that there exists at least one multigraph with these degrees. First assign vertex i a number d_i of half-edges. Then choose a uniform random pairing of the half-edges. This generates a random multigraph $CM(\mathbf{d})$. For a fixed multigraph $\mathcal{M} = ([n], E(\mathcal{M}))$ having degrees $\mathbf{d}$, s self-loops and m_e

the multiplicity of $e \in E(\mathcal{M})$, we have

$$\mathbb{P}(CM(\mathbf{d}) = \mathcal{M}) \propto \frac{1}{2^s \prod_{e \in E(\mathcal{M})} m_e!}.$$

In particular, assuming that there exists at least one simple graph with degrees $\mathbf{d}$, conditionally on the event that $CM(\mathbf{d})$ is simple, it is uniformly distributed on the set of simple graphs with those degrees. (See the account in Chapter 7 of van der Hofstad [52] for proofs of these results.)

Now fix $t, u > 0$ and $v \geq 0$, and suppose we start from a degree sequence with $\lfloor tN^{3/5} \rfloor$ vertices of degree 2, $\lfloor uN^{2/5} \rfloor$ of degree 3 and $\lfloor vN^{1/5} \rfloor$ of degree 4. Let G_N be a uniform random graph with these degrees. Let K_N be the *kernel* of G_N, that is the multigraph obtained by contracting paths of vertices of degree 2. It is straightforward to see that K_N is distributed according to the configuration model with $\lfloor uN^{2/5} \rfloor$ vertices of degree 3 and $\lfloor vN^{1/5} \rfloor$ vertices of degree 4. With probability tending to 1 as $N \to \infty$, G_N possesses a giant complex component C_N, containing all of the vertices of degrees 3 and 4 (i.e. the kernel), as well as some random number t'_N of the vertices of degree 2, where $t'_N / N^{3/5} \xrightarrow{p} t$ as $n \to \infty$. On this event of high probability, C_N is a uniform connected graph with its degree sequence. Outside the giant, there is a collection of $O(\log N)$ disjoint cycles, containing the remaining vertices of degree 2, with high probability as $N \to \infty$. [1]

Now consider the following way of constructing a complex component which we claim has approximately the same distribution as C_N. First generate the kernel K_N according to the configuration model with $\lfloor uN^{2/5} \rfloor$ vertices of degree 3 and $\lfloor vN^{1/5} \rfloor$ vertices of degree 4. Then, one-by-one, allocate t'_N vertices of degree 2 to the edges of K_N: at each step, an edge of the current structure is chosen uniformly at random, split into two edges in series, and the vertex is inserted into the middle. Thus the lengths

[1] We have not found a good reference for these statements, which are essentially folklore in the random graphs literature; statements in a similar spirit may be found, for example, in the recent paper of Joos et al [31]. In particular, their Theorem 2 implies that G_N contains a giant component with high probability. By Theorem 4.14 of van der Hofstad [51], K_N is connected with high probability. That the giant component of G_N contains K_N and a proportion 1 of all vertices of degree 2 essentially comes down to the fact that the number of ways of generating a random 2-regular graph is negligible compared to the number of ways of generating a graph with kernel K_N. (We omit the details since our primary aim here is not to make rigorous statements but rather to give heuristics.) Finally, a random 2-regular graph on n vertices has $O(\log n)$ components with high probability as $n \to \infty$ (see, for example, Arratia et al [3]).

of the paths of degree-2 vertices we insert between neighbouring vertices in K_N evolve according to a multicolour Pólya's urn with $(3\lfloor uN^{2/5}\rfloor + 4\lfloor vN^{1/5}\rfloor)/2$ colours and a single ball of each colour to start.[2] Then for large N, the proportions of the t'_N vertices of degree 2 which get allocated to each of the edges of K_N look approximately like a Dirichlet$(1, 1, \ldots, 1)$ vector, with $(3\lfloor uN^{2/5}\rfloor + 4\lfloor vN^{1/5}\rfloor)/2$ co-ordinates.

Let us now consider how the Karp-Sipser algorithm behaves on such a graph. We may neglect the cycle components, as they can give rise to at most $O(\log N)$ unmatched vertices. The algorithm first picks a uniform edge, matches its end-points, and then removes the neighbouring edges. Typically we matched two vertices somewhere inside a long path of degree 2 vertices, and so we are now left with two pendant vertices, each at the end of a path of degree-2 vertices with a higher-degree vertex at its end. If such a path is of odd length, Karp-Sipser will "consume" the degree-2 vertices but leave the higher-degree vertex untouched (except to reduce its degree by 1). If, on the other hand, the path is of even length, the higher-degree vertex gets matched and removed, causing its neighbours to become pendant vertices. Thus, in this case, the algorithm eats further away into the graph. If ever a particular path gets eaten away at from both ends (which can happen since the graph has cycles), there is a chance that some vertex in the path will remain unmatched. Again, whether this in fact happens or not depends on the parity of the path of degree-2 vertices. For large enough N, we expect that such paths will be of odd and even lengths with approximately equal probability. So we expect that paths which get burnt from both ends will give rise to an unmatched vertex with probability 1/2.

As we have already argued, since the number of edges in the multigraph is much smaller than the number of degree-2 vertices, the proportions of degree-2 vertices assigned to each edge of the multigraph will be, for large N, close to Dirichlet$(1, 1, \ldots, 1)$. For fixed $k \geq 2$, a vector $(D_1, D_2, \ldots, D_k)$ with Dirichlet$(1, 1, \ldots, 1)$ distribution may be obtained by sampling $E_1, E_2, \ldots, E_k$, independent and identically distributed standard

[2]By Proposition 7.13 of [52], K_N possesses constant order numbers of self-loops and multiple edges (indeed, the numbers of self-loops and multiple edges converge jointly in distribution as $N \to \infty$ to a pair of independent Poisson(1) random variables); in the urn process, there is negligible probability that we fail to allocate any vertices of degree 2 to the self-loops or to at least two of a set of edges between the same two vertices.

exponential random variables and setting

$$(D_1, D_2, \dots, D_k) = \frac{1}{\sum_{i=1}^k E_i}(E_1, E_2, \dots, E_k). \tag{3.2}$$

Once we have accounted for parity, the lengths of the paths of degree-2 vertices play a role only when we pick a new uniform edge to match, which we do with probability proportional to length. So we can equivalently think of $E_1, \dots, E_k$ as the "lengths" of the paths of degree 2 vertices in our approximate model.

In summary, letting n be the number of vertices of degree 3 and $\alpha(n)$ be the number of vertices of degree 4, we obtain the model described in Section 3.1.1 as an approximation. We do not attempt here to assess the quality of this approximation (and we only make rigorous statements about the model described in Section 3.1.1). Rather our interest is in the mechanism by which the distribution of the number of vertices which remain unmatched at the end of the Karp-Sipser algorithm arises. With the scaling suggested by Aronson, Frieze and Pittel, we would have $n = \Theta(N^{2/5})$ and $\alpha(n) = \Theta(N^{1/5})$ i.e. $\alpha(n) = \Theta(\sqrt{n})$. So we should be in regime (i) of Theorem 56, which supports the conjecture that $N^{-1/5}A_N$ possesses a limit in distribution.

3.1.3 Our analysis

Our model has two convenient features which make it amenable to analysis: the distributional properties of the exponential edge-lengths and the fact that we may sample the multigraph edge by edge at the same time that we burn it. Let us first address the edge-lengths. We will make use of the following result (Proposition 1 of [8]), which follows from standard properties of exponential random variables via the relationship (3.2).

Proposition 58. *(i) Suppose that* $(D_1, \dots, D_k) \sim \mathrm{Dirichlet}_k(1, 1, \dots, 1)$ *and let* I *be a random index from* $\{1, 2, \dots, k\}$ *with conditional distribution*

$$\mathbb{P}(I = i | D_1, \dots, D_k) = D_i, \quad 1 \le i \le k.$$

Let U be independent of everything else with uniform distribution on $[0,1]$. Then defining

$$D'_i = \begin{cases} D_i & \text{for } 1 \le i < I, \\ UD_i & \text{for } i = I, \\ (1-U)D_i & \text{for } i = I+1, \\ D_{i-1} & \text{for } I+1 < i \le k+1, \end{cases}$$

we have that $(D'_1, D'_2, \ldots, D'_{k+1}) \sim \mathrm{Dirichlet}_{k+1}(1,1,\ldots,1)$.

(ii) Suppose that $(D'_1, D'_2, \ldots, D'_{k+1}) \sim \mathrm{Dirichlet}_{k+1}(1,1,\ldots,1)$ and that J is chosen independently and uniformly at random from $\{1,2,\ldots,k\}$. Then defining

$$D_i = \begin{cases} D'_i & \text{for } 1 \le i < J, \\ D'_i + D'_{i+1} & \text{for } i = J, \\ D'_{i+1} & \text{for } J+1 \le i \le k, \end{cases}$$

we have that $(D_1, D_2, \ldots, D_k) \sim \mathrm{Dirichlet}_k(1,1,\ldots,1)$.

Using part (i) of this proposition, we see that if we start from $\mathrm{Dirichlet}(1,1,\ldots,1)$ relative edge lengths and pick a point uniformly from the length measure, then splitting at it yields relative distances from the sampled point to the adjacent vertices of degree at least 3 either side of it which are again part of $\mathrm{Dirichlet}(1,1,\ldots,1)$ vector (with one more coordinate). We will see in what follows that the property that (given the number of edges in the multigraph) the relative edge-lengths are $\mathrm{Dirichlet}(1,1,\ldots,1)$ distributed is preserved. Consider the process at an instant when the fire has just reached a vertex, or when a point has just been set alight. In general there are several edges burning, and by the memoryless property of the exponential distribution, their relative distances to adjacent vertices are still standard exponential divided by the sum of those exponentials. In particular, whenever there are multiple edges burning, the next to reach its vertex is chosen uniformly from among all those present. Since we are not interested in how long the whole process takes but rather in the number of times we must set light to the network and how many clashes we observe, we perform our analysis in discrete time: that is, we consider the multigraph *without* edge-lengths, always set fire to a uniformly-chosen edge (which has the effect of splitting it into two edges, both alight), and we choose which edge next finishes burning uniformly from among those currently alight.

Recall the description of the configuration model from the previous section. We may generate the pairing of the half-edges in any order we like, which makes the configuration model particularly amenable to an exploration process-type analysis. In particular, given that we have revealed the pairings of a particular collection of half-edges, assuming we keep track of the degree sequence of the rest of the graph, the rest of the graph is again a configuration model with that degree sequence. We exploit this property below. Observe that we have n vertices of degree 3 and $\alpha(n)$ of degree 4 in our configuration model. Our process starts by picking a uniform edge, splitting it in two and setting each resulting edge alight. (Let us refer to this as the beginning of a *wave*, with a new wave beginning every time we set light to a point in the multigraph.) A uniform one of these two edges reaches its vertex next. It samples its vertex from among those of degree 3 with probability $3n/(3n + 4\alpha(n))$ and from those of probability 4 with probability $4\alpha(n)/(3n + 4\alpha(n))$. Thereafter, we may think of edges which are alight as the first half-edge of a pair, whose second half-edge we have yet to sample.

Suppose that we currently have x edges burning, and that there are u vertices with 3 unpaired half-edges and v vertices with 4 unpaired half-edges. We pick the next half-edge to process uniformly from those currently burning, and pick its pair uniformly at random from among those available, including any which are themselves burning. The latter event happens with probability $(x-1)/(3u+4v+x-1)$, in which case there is an edge burning from both ends and we generate a clash. Otherwise, if we connect to a vertex of degree 3, which occurs with probability $3u/(3u+4v+x-1)$, the fire is either passed to the two other half-edges (with probability $1/2$) or stopped. If it is stopped, the vertex of degree 3 becomes a vertex of degree 2. There are two different things that might happen to this vertex of degree 2. With probability $1/(3u + 4v + x - 3)$, its two half-edges are in fact connected to each other to form an isolated cycle. (As observed above, this is a rare event: there are only $O(1)$ self-loops in the whole multigraph.) Any such isolated cycle necessarily yields a clash. With the complementary probability, the two remaining half-edges are not connected to each other but rather to other half-edges.

Since vertices of degree 2 cannot stop fires, in this case we may simply remove the vertex and contract the path of length 2 in which it sat to a single edge. (Using part (ii) of Proposition 58, this results the relative edge-lengths of the edges in the unseen parts of the multigraph still being Dirichlet$(1, 1, \dots, 1)$ distributed). In either case, the remaining unseen part of the multigraph (after deletion of the loop, or contraction of a path of length 2) is still distributed according to the configuration model with the updated degree distribution. Finally, if we connect to a vertex of degree 4, which occurs with probability $4v/(3u + 4v + x - 1)$, then either the fire is passed to all three other neighbours (with probability $1/2$) or to none of them, in which case the result is that we get another vertex of degree 3.

Note that we must treat the very first edge of a wave differently: although we start with two burning half-edges, they cannot be paired to each other (since otherwise there would be an edge in the original graph with no vertex, which is impossible). So let us treat the first step of a wave as consisting of picking a uniform edge, splitting it in two, sampling the vertex to which one of the resulting burning edges is attached and seeing whether it passes the fire on or not. So if at some step, there are no burning half-edges, on the next step there will be either 1, 3 or 4 burning half-edges, corresponding to the events that the first of the two fires was stopped, that it was passed on through a vertex of degree 3, or that it was passed on through a vertex of degree 4.

In this way, we see that at each step of the procedure, we either process one vertex or generate a clash. A vertex of degree 3 is processed precisely once; a vertex of degree 4 is processed once or twice, depending on whether it stops the first fire it encounters or not.

If we track the numbers of vertices of degree 3 and 4 and the number of currently burning half-edges, we have a Markovian evolution, which we may hope to analyse using the tools of stochastic process theory. In particular, in what follows we make extensive use of martingales.

The rest of this chapter is organized as follows. In Section 3.2, we write down explicitly the transition probabilities of our Markov chain, and give some first estimates relevant for the forthcoming analysis. In particular, we identify a coupling which facilitates our analysis of the end of the process. In Section 3.3, we prove fluid limits for the suitably rescaled number of nodes of degree 3 and 4; that is, we show that these processes remain close to deterministic functions on a time-interval which is bounded away from the end of the process (this is an application of the fluid limit method we briefly introduced in Chapter 1, Section 1.3). In Section 3.4, we analyse the case $\alpha(n) >> \sqrt{n}$, and prove a fluid limit result for the (suitably rescaled) numbers of fires and clashes we observe, as long as we are bounded away from the end of the process. Section 3.5 deals with the limiting properties of the numbers of fires and the number of clashes we observe, again as long as we bounded away from the end of the process, for $\alpha(n) = O(\sqrt{n})$. In this case, the limiting process for the number of fires is a reflected diffusion, and the proof is based on an invariance principle for reflecting Markov chains (in the spirit of [21, 33, 50]). Finally, in Section 3.6, we prove that the end of the process does not contribute significantly to any of these quantities, and so the convergence results can be extended into that range too.

3.2 The Markov chain

For $i \geq 0$, let $U^n(i)$ and $V^n(i)$ represent the numbers of nodes of degree 3 and 4 respectively after i steps of the burning procedure. Let $X^n(i)$ be the number of burning half-edges we have after i steps. Let $N^n(i)$ be the counting process of the number of clashes observed up to step i. Set $U^n(0) = n$, $V^n(0) = \alpha(n)$, $X^n(0) = 0$, and $N^n(0) = 0$. We have already argued that the process

$$(U^n(i), V^n(i), X^n(i), N^n(i))_{i\geq 0}$$

evolves in a Markovian manner until time $\zeta_n = \inf\{i \geq 0 : U^n(i) + V^n(i) + X^n(i) = 0\}$, when it stops. Write

$$L^n(k) = 2\sum_{i=0}^{k-1} \mathbb{1}_{\{X^n(i)=0\}}$$

(we will think of this quantity as a local time). Then

$$F^n = \frac{1}{2}L^n(\zeta_n), \quad C^n = N^n(\zeta_n).$$

We will find it convenient to rescale time by n (essentially because we start with $n+\alpha(n)$ vertices and a typical step involves the removal of a vertex of degree 3).

Recall the definition of the process X^a from (3.1). We also let $x : [0,1] \to \mathbb{R}_+$ be the (deterministic) function

$$x(t) = (1-t)^{2/3} - (1-t)^{4/3}, \quad 0 \leq t \leq 1.$$

For $0 \leq t \leq 1$, let

$$N_t^a = \int_0^t \frac{X_s^a}{3(1-s)} ds, \qquad m(t) = \frac{1}{4} - \frac{1}{2}(1-t)^{2/3} + \frac{1}{4}(1-t)^{4/3}.$$

Let $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ denote the space of càdlàg functions from $\mathbb{R}_+$ to $\mathbb{R}$, equipped with the Skorokhod topology. We will study the convergence of the sequence of probability measures on the measurable space given by $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ endowed with its Borel σ-algebra. (In fact, since our limit processes will always be continuous, we will rather obtain convergence with respect to the uniform norm.) The crux of our argument is the following scaling limit theorem.

Theorem 59. *(i) Suppose $\alpha(n)/\sqrt{n} \to a$ as $n \to \infty$, for $a \geq 0$. Then*

$$\frac{1}{\sqrt{n}}(X^n(\lfloor nt \rfloor), L^n(\lfloor nt \rfloor), N^n(\lfloor nt \rfloor), 0 \leq t \leq \zeta_n/n) \xrightarrow{d} (X_t^a, L_t^a, N_t^a, 0 \leq t \leq 1)$$

uniformly as $n \to \infty$.

(ii) Suppose $\alpha(n) >> \sqrt{n}$. Then

$$\frac{1}{\alpha(n)}(X^n(\lfloor nt \rfloor), L^n(\lfloor nt \rfloor), N^n(\lfloor nt \rfloor), 0 \leq t \leq \zeta_n/n) \xrightarrow{d} (x(t), 0, m(t), 0 \leq t \leq 1)$$

uniformly as $n \to \infty$.

It is straightforward to see that this implies Theorem 56.

3.2.1 Transition probabilities

We have already described the possible transitions of our four-dimensional Markov chain in Section 3.1. Let us be a little more explicit about the transition probabilities.

Suppose that $(U^n(i), V^n(i), X^n(i), N^n(i)) = (u, v, x, m) \in \mathbb{Z}_+^4$. If $x = 0$ but $u+v > 0$, then

$$
\begin{aligned}
&(U^n(i+1), V^n(i+1), X^n(i+1), N^n(i+1)) \\
&\quad = \begin{cases}
(u-1, v, 1, m+1) & \text{with probability } \frac{3u}{2(3u+4v)(3u+4v-2)}, \\
(u-1, v, 1, m) & \text{with probability } \frac{3u(3u+4v-3)}{2(3u+4v)(3u+4v-2)}, \\
(u-1, v, 3, m) & \text{with probability } \frac{3u}{2(3u+4v)}, \\
(u+1, v-1, 1, m) & \text{with probability } \frac{4v}{2(3u+4v)}, \\
(u, v-1, 4, m) & \text{with probability } \frac{4v}{2(3u+4v)}.
\end{cases}
\end{aligned}
$$

If $x > 0$ then

$$
\begin{aligned}
&(U^n(i+1), V^n(i+1), X^n(i+1), N^n(i+1)) \\
&\quad = \begin{cases}
(u-1, v, x-1, m+1) & \text{with probability } \frac{3u}{2(3u+4v+x-1)(3u+4v+x-3)}, \\
(u-1, v, x-1, m) & \text{with probability } \frac{3u(3u+4v+x-4)}{2(3u+4v+x-1)(3u+4v+x-3)}, \\
(u-1, v, x+1, m) & \text{with probability } \frac{3u}{2(3u+4v+x-1)}, \\
(u+1, v-1, x-1, m) & \text{with probability } \frac{4v}{2(3u+4v+x-1)}, \\
(u, v-1, x+2, m) & \text{with probability } \frac{4v}{2(3u+4v+x-1)}, \\
(u, v, x-2, m+1) & \text{with probability } \frac{x-1}{(3u+4v+x-1)}.
\end{cases}
\end{aligned}
$$

Writing $(\mathcal{F}_i^n)_{i\geq 0}$ for the natural filtration of the four-dimensional process, we have

$$
\begin{aligned}
\mathbb{E}[U^n(i+1) - U^n(i)|\mathcal{F}_i^n] &= \frac{4V^n(i) - 3U^n(i)}{3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}}}, \\
\mathbb{E}[V^n(i+1) - V^n(i)|\mathcal{F}_i^n] &= \frac{-4V^n(i)}{3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}}},
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}[X^n(i+1) - X^n(i)|\mathcal{F}_i^n] &= 2\mathbb{1}_{\{X^n(i)=0\}} \\
&\qquad + \frac{2V^n(i) - 2X^n(i) + 2\mathbb{1}_{\{X^n(i)>0\}}}{3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}}},
\end{aligned}
\tag{3.3}
$$

as well as

$$\mathbb{E}[N^n(i+1) - N^n(i)|\mathcal{F}_i^n] = \frac{3U^n(i)+2(X^n(i)-\mathbb{1}_{\{X^n(i)>0\}})(3U^n(i)+4V^n(i)+X^n(i)-3)}{2(3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}})(3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}-2)}.$$

We will also need the following conditional second moments:

$$\begin{aligned}
\mathbb{E}\Big[(U^n(i+1) - U^n(i))^2|\mathcal{F}_i^n\Big] &= \frac{3U^n(i) + 2V^n(i)}{3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}}}, \\
\mathbb{E}\Big[(V^n(i+1) - V^n(i))^2|\mathcal{F}_i^n\Big] &= \frac{4V^n(i)}{3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}}}, \\
\mathbb{E}\Big[(X^n(i+1) - X^n(i))^2|\mathcal{F}_i^n\Big] &= \begin{cases} 5 + \frac{14V^n(i)}{3U^n(i)+4V^n(i)} & \text{if } X^n(i) = 0, \\ 1 + \frac{6V^n(i)+3X^n(i)-3}{3U^n(i)+4V^n(i)+X^n(i)-1} & \text{if } X^n(i) > 0, \end{cases} \qquad (3.4) \\
\mathbb{E}\Big[(N^n(i+1) - N^n(i))^2|\mathcal{F}_i^n\Big] &= \mathbb{E}[N^n(i+1) - N^n(i)|\mathcal{F}_i^n].
\end{aligned}$$

3.2.2 A coupling and some first estimates

The process runs for ζ_n steps. On each step, we process a whole edge of the multigraph, except at the start of a wave, when we possibly need two steps to process an edge. There are $(3n + 4\alpha(n))/2$ edges in total and so we get the crude bound

$$\zeta_n \leq 3n + 4\alpha(n). \qquad (3.5)$$

In the sequel, and particularly in Section 3.6, we will make extensive use of a coupling of a modified version of X^n and a reflecting simple symmetric random walk, which we now introduce. First, we divide the burning half-edges into two stacks, of sizes $X_1^n(i)$ and $X_2^n(i)$, such that $X^n(i) = X_1^n(i) + X_2^n(i)$ for all $0 \leq i \leq \zeta_n$. We may give these sub-processes whatever dynamics we choose, as long as their sum behaves as $(X^n(i))_{i \geq 0}$. We proceed as follows.

Whenever $X_1^n(i) > 0$, we select the next half-edge to process from the first stack. If the fire is absorbed, X_1^n is simply reduced by 1. If we connect to a vertex of degree 3 and pass the fire on, X_1^n increases by 1. If we connect to a vertex of degree 4 and pass the fire on, we let each of X_1^n and X_2^n increase by 1. If we create a clash with a half-edge from the first stack, X_1^n is reduced by 2. We may also create a clash with a half-edge from the second stack, in which case X_1^n and X_2^n are both reduced by 1.

If $X_1^n(i) = 0$ but $X_2^n(i) > 0$, then we select the half-edge to process from the second stack, add any new half-edges arising from passing the fire on to a vertex of degree 3 to X_1^n and add 2 of the three burning half-edges arising from passing the fire on to a vertex of degree 4 to X_1^n and the last one to X_2^n. Finally, if $X_1^n(i) = X_2^n(i) = 0$, then we allocate all new half-edges to the first stack, except if we connect to a vertex of degree 4 and pass on the fire, in which case X_1^n jumps to 3 and X_2^n jumps to 1.

We will also track the clashes and split them according to whether they involve a half-edge from the second stack or not, yielding $N^n(i) = N_1^n(i) + N_2^n(i)$ for $i \geq 0$.

We will describe the transition probabilities of this process in detail below. Our aim is to couple X_1^n with a process Y^n in such a way that $X_1^n(i) \leq Y^n(i)$ for all $i \geq 0$ and Y^n is a simple symmetric random walk (SSRW), reflected at 2. In order to keep the notation to a reasonable level, we will describe the transitions of $(X_1^n(i), X_2^n(i), Y^n(i), N_1^n(i), N_2^n(i))_{i\geq 0}$, leaving those of $(U^n(i), V^n(i))_{i\geq 0}$ implicit.

Conditional on $X_1^n(i) = x_1 > 0, X_2^n(i) = x_2 \geq 0, Y^n(i) = y \geq 3, U^n(i) = u, V^n(i) = v$, with $x = x_1 + x_2$,

$$
(X_1^n(i+1) - X_1^n(i), X_2^n(i+1) - X_2(i), Y^n(i+1) - Y^n(i), N_1^n(i+1) - N_1^n(i), N_2^n(i+1) - N_2^n(i))
$$
$$
= \begin{cases}
(-2, 0, -1, +1, 0) & \text{with probability } \frac{x_1-1}{2(3u+4v+x-1)} \\
(-2, 0, +1, +1, 0) & \text{with probability } \frac{x_1-1}{2(3u+4v+x-1)} \\
(-1, -1, -1, 0, +1) & \text{with probability } \frac{x_2}{2(3u+4v+x-1)} \\
(-1, -1, +1, 0, +1) & \text{with probability } \frac{x_2}{2(3u+4v+x-1)} \\
(-1, 0, -1, +1, 0) & \text{with probability } \frac{3u}{2(3u+4v+x-1)(3u+4v+x-3)} \\
(-1, 0, -1, 0, 0) & \text{with probability } \frac{4v}{2(3u+4v+x-1)} + \frac{3u(3u+4v+x-4)}{2(3u+4v+x-1)(3u+4v+x-3)} \\
(+1, 0, +1, 0, 0) & \text{with probability } \frac{3u}{2(3u+4v+x-1)} \\
(+1, +1, +1, 0, 0) & \text{with probability } \frac{4v}{2(3u+4v+x-1)}.
\end{cases}
$$

Conditional on $X_1^n(i) = x_1 > 0, X_2^n(i) = x_2 \geq 0, Y^n(i) = y = 2, U^n(i) = u, V^n(i) = v$, with $x = x_1 + x_2$,

$$(X_1^n(i+1) - X_1^n(i), X_2^n(i+1) - X_2(i), Y^n(i+1) - Y^n(i), N_1^n(i+1) - N_1^n(i), N_2^n(i+1) - N_2^n(i))$$

$$= \begin{cases} (-2,0,+1,+1,0) & \text{with probability } \frac{x_1-1}{(3u+4v+x-1)} \\ (-1,-1,+1,0,+1) & \text{with probability } \frac{x_2}{(3u+4v+x-1)} \\ (-1,0,+1,+1,0) & \text{with probability } \frac{3u}{2(3u+4v+x-1)(3u+4v+x-3)} \\ (-1,0,+1,0,0) & \text{with probability } \frac{4v}{2(3u+4v+x-1)} + \frac{3u(3u+4v+x-4)}{2(3u+4v+x-1)(3u+4v+x-3)} \\ (+1,0,+1,0,0) & \text{with probability } \frac{3u}{2(3u+4v+x-1)} \\ (+1,+1,+1,0,0) & \text{with probability } \frac{4v}{2(3u+4v+x-1)}. \end{cases}$$

Conditional on $X_1^n(i) = 0, X_2^n(i) = x_2 > 0, Y^n(i) = y \geq 3, U^n(i) = u, V^n(i) = v$,

$$(X_1^n(i+1) - X_1^n(i), X_2^n(i+1) - X_2(i), Y^n(i+1) - Y^n(i), N_1^n(i+1) - N_1^n(i), N_2^n(i+1) - N_2^n(i))$$

$$= \begin{cases} (0,-2,-1,0,+1) & \text{with probability } \frac{x_2-1}{2(3u+4v+x_2-1)} \\ (0,-2,+1,0,+1) & \text{with probability } \frac{x_2-1}{2(3u+4v+x_2-1)} \\ (0,-1,-1,+1,0) & \text{with probability } \frac{3u}{2(3u+4v+x_2-1)(3u+4v+x_2-3)} \\ (0,-1,-1,0,0) & \text{with probability } \frac{4v}{2(3u+4v+x_2-1)} + \frac{3u(3u+4v+x_2-4)}{2(3u+4v+x_2-1)(3u+4v+x_2-3)} \\ (+2,-1,+1,0,0) & \text{with probability } \frac{3u}{2(3u+4v+x_2-1)} \\ (+2,0,+1,0,0) & \text{with probability } \frac{4v}{2(3u+4v+x_2-1)}. \end{cases}$$

Conditional on $X_1^n(i) = 0, X_2^n(i) = x_2 > 0, Y^n(i) = y = 2, U^n(i) = u, V^n(i) = v$,

$$(X_1^n(i+1) - X_1^n(i),\ X_2^n(i+1) - X_2(i),\ Y^n(i+1) - Y^n(i), N_1^n(i+1) - N_1^n(i), N_2^n(i+1) - N_2^n(i))$$

$$= \begin{cases} (0,-2,+1,0,+1) & \text{with probability } \frac{x_2-1}{(3u+4v+x_2-1)} \\ (0,-1,+1,+1,0) & \text{with probability } \frac{3u}{2(3u+4v+x_2-1)(3u+4v+x_2-3)} \\ (0,-1,+1,0,0) & \text{with probability } \frac{4v}{2(3u+4v+x_2-1)} + \frac{3u(3u+4v+x_2-4)}{2(3u+4v+x_2-1)(3u+4v+x_2-3)} \\ (+2,-1,+1,0,0) & \text{with probability } \frac{3u}{2(3u+4v+x_2-1)} \\ (+2,0,+1,0,0) & \text{with probability } \frac{4v}{2(3u+4v+x_2-1)}. \end{cases}$$

Conditional on $X_1^n(i) = 0, X_2^n(i) = 0, Y^n(i) = y \geq 3, U^n(i) = u, V^n(i) = v$,

$$(X_1^n(i+1) - X_1^n(i), X_2^n(i+1) - X_2(i), Y^n(i+1) - Y^n(i), N_1^n(i+1) - N_1^n(i), N_2^n(i+1) - N_2^n(i))$$

$$= \begin{cases} (+1, 0, -1, 0, 0) & \text{with probability } \frac{4v}{2(3u+4v)} + \frac{3u(3u+4v-3)}{2(3u+4v)(3u+4v-2)} \\ (+1, 0, -1, +1, 0) & \text{with probability } \frac{3u}{2(3u+4v)(3u+4v-2)} \\ (+3, 0, +1, 0, 0) & \text{with probability } \frac{3u}{2(3u+4v)} \\ (+3, +1, +1, 0, 0) & \text{with probability } \frac{4v}{2(3u+4v)}. \end{cases}$$

Finally, conditional on $X_1^n(i) = 0, X_2^n(i) = 0, Y^n(i) = y = 2, U^n(i) = u, V^n(i) = v$,

$$(X_1^n(i+1) - X_1^n(i), X_2^n(i+1) - X_2(i), Y^n(i+1) - Y^n(i), N_1^n(i+1) - N_1^n(i), N_2^n(i+1) - N_2^n(i))$$

$$= \begin{cases} (+1, 0, +1, 0, 0) & \text{with probability } \frac{4v}{2(3u+4v)} + \frac{3u(3u+4v-3)}{2(3u+4v)(3u+4v-2)} \\ (+1, 0, +1, +1, 0) & \text{with probability } \frac{3u}{2(3u+4v)(3u+4v-2)} \\ (+3, 0, +1, 0, 0) & \text{with probability } \frac{3u}{2(3u+4v)} \\ (+3, +1, +1, 0, 0) & \text{with probability } \frac{4v}{2(3u+4v)}. \end{cases}$$

By construction, Y^n performs a SSRW with upward reflection at 2. For fixed $i \geq 0$, as long as $Y^n(i) \geq 2$ and $X_1^n(i) \leq Y^n(i)$, we obtain $X_1^n(j) \leq Y^n(j)$ for all $j \geq i$.

Finally, we observe that since a vertex of degree 4 contributes to the size of the second stack at most once, $X_2^n(i) \leq V^n(0) - V^n(i)$ for all $i \geq 0$. Similarly, the number of clashes involving at least one half-edge from the second stack is bounded above by the number of vertices of degree 4 processed so far, $N_2^n(i) \leq V^n(0) - V^n(i)$.

Henceforth, we will use the notation $(\mathcal{F}_i^n)_{i \geq 0}$ for the natural filtration of the process

$$(X_1^n(i), X_2^n(i), Y^n(i), U^n(i), V^n(i), N_1^n(i), N_2^n(i))_{i \geq 0}.$$

For future reference, we note that

$$\begin{aligned}
&\mathbb{E}[X_1^n(i+1) - X_1^n(i)|\mathcal{F}_i^n] \\
&\qquad = \mathbb{1}_{\{X_1^n(i)=0, X_2^n(i)>0\}} + 2\mathbb{1}_{\{X_1^n(i)=0, X_2^n(i)=0\}} \\
&\qquad\quad - \frac{2(X_1^n(i) - \mathbb{1}_{\{X_1^n(i)>0\}}) + (X_2^n(i) - \mathbb{1}_{\{X_1^n(i)=0, X_2^n(i)>0\}})}{3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}}}
\end{aligned} \tag{3.6}$$

and

$$\begin{aligned}
&\mathbb{E}[N_1^n(i+1) - N_1^n(i)|\mathcal{F}_i^n] \\
&= \frac{3U^n(i)+2(X_1^n(i)-\mathbb{1}_{\{X_1^n(i)>0\}})(3U^n(i)+4V^n(i)+X^n(i)-3)}{2(3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}})(3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}-2)}.
\end{aligned} \tag{3.7}$$

As a first consequence of our coupling, we show that X^n varies on a smaller scale than n.

Lemma 60. *We have that* $\frac{1}{\alpha(n)\vee\sqrt{n}} \sup_{0\le i\le\zeta_n} X^n(i)$ *is bounded in* L^2*. In particular,*

$$\frac{1}{n}\sup_{0\le i\le\zeta_n} X^n(i) \xrightarrow{p} 0,$$

as $n \to \infty$*.*

Proof. We have $X^n(0) = 0$; let $Y^n(0) = 2$. Then by the coupling, we have

$$\sup_{0\le i\le\zeta_n} X^n(i) \le \alpha(n) + \sup_{0\le i\le\zeta_n} Y^n(i) \le \alpha(n) + \sup_{0\le i\le 3n+4\alpha(n)} Y^n(i).$$

Now let Z be a standard SSRW, and note that Z is a martingale. Then Y^n has the same law as $(2 + |Z(i)|)_{i\ge0}$. By Doob's L^2 inequality,

$$\mathbb{E}\left[\left(\sup_{0\le i\le 3n+4\alpha(n)} |Z(i)|\right)^2\right] \le 4\mathbb{E}\Big[Z(3n+4\alpha(n))^2\Big] = 4(3n+4\alpha(n)).$$

Hence,

$$\begin{aligned}
\mathbb{E}\left[\left(\sup_{0\le i\le\zeta_n} X^n(i)\right)^2\right] &\le \mathbb{E}\left[\left(\alpha(n) + 2 + \sup_{0\le i\le 3n+4\alpha(n)} |Z(i)|\right)^2\right] \\
&\le (\alpha(n)+2)^2 + 4(3n+4\alpha(n)) + 4(\alpha(n)+2)\sqrt{3n+4\alpha(n)}.
\end{aligned}$$

It follows that there exists a constant $C > 0$ such that for all $n \geq 1$ we have

$$\mathbb{E}\left[\left(\frac{1}{\alpha(n) \vee \sqrt{n}} \sup_{0 \leq i \leq \zeta_n} X^n(i)\right)^2\right] \leq C.$$

Since $\alpha(n)/n \to 0$, we have moreover that

$$\mathbb{E}\left[\left(\frac{1}{n} \sup_{0 \leq i \leq \zeta_n} X^n(i)\right)^2\right] \to 0,$$

and so $\frac{1}{n} \sup_{0 \leq i \leq \zeta_n} X^n(i) \xrightarrow{p} 0$. □

3.3 Fluid limit approximations for the auxiliary processes

In this section we prove that U^n and V^n, after appropriate rescaling, remain concentrated around their expected trajectories. In order to do so, we employ the differential equations method [14, 57] we introduced in Chapter 1, Section 1.3. Here we briefly recall the main idea, closely following Darling and Norris [14], although our presentation is for discrete-time rather than continuous-time processes and is in a somewhat simplified setting. Suppose that we are given a stochastic process P^n evolving in discrete time with finite state-space $\mathcal{S}^n \subseteq \mathbb{Z}$. We assume that P^n is either itself Markov or is a co-ordinate of some higher-dimensional Markov process $\mathbf{P}^n$. Let the natural filtration of $\mathbf{P}^n$ be denoted by $(\mathcal{F}^n_{\mathbf{P}}(i))_{i \geq 0}$. For each $i \geq 0$, the process $(M^n_P(i))_{i \geq 0}$ defined by

$$M^n_P(i) = P^n(i) - P^n(0) - \sum_{j=0}^{i-1} \mathbb{E}[P^n(j+1) - P^n(j) | \mathcal{F}^n_{\mathbf{P}}(j)]$$

is a martingale. Then, for each fixed $t > 0$, we have

$$\frac{P^n(\lfloor nt \rfloor)}{n} = \frac{P^n(0)}{n} + \frac{M^n_P(\lfloor nt \rfloor)}{n} + \sum_{j=0}^{\lfloor nt \rfloor - 1} \frac{\mathbb{E}[P^n(j+1) - P^n(j) | \mathcal{F}^n_{\mathbf{P}}(j)]}{n}.$$

Fix $t_0 > 0$ and let $\mathcal{U} \subseteq \mathbb{R}$. We will make four assumptions.

(1) For some constant $p(0) \in \mathcal{U}$ we have $\left|\frac{P^n(0)}{n} - p(0)\right| \xrightarrow{p} 0$.

(2) Suppose that $\nu : [0, t_0] \times \mathcal{U} \to \mathbb{R}$ is a bi-Lipschitz function with Lipschitz constant K. Suppose that p, the unique solution to the differential equation

$$\frac{dp(t)}{dt} = \nu(t, p(t))$$

with initial condition $p(0)$, is such that, for some $\epsilon > 0$, $p(t)$ lies at distance greater than ϵ from $\mathcal{U}^c$ for all $0 \leq t \leq t_0$.

(3) Let $T_n = \inf\{t \geq 0 : P^n(\lfloor nt \rfloor)/n \notin \mathcal{U}\}$. We assume that

$$\sup_{0 \leq t \leq t_0 \wedge T_n} \left| \sum_{j=0}^{\lfloor nt \rfloor - 1} \frac{\mathbb{E}[P^n(j+1) - P^n(j) | \mathcal{F}_P^n(j)]}{n} - \int_0^t \nu\left(s, \frac{P^n(\lfloor ns \rfloor)}{n}\right) ds \right| \xrightarrow{p} 0.$$

(4) We have

$$\frac{\mathbb{E}\left[M_P^n(\lfloor n(t_0 \wedge T_n) \rfloor)^2\right]}{n^2} \to 0$$

as $n \to \infty$.

Assumption (1) tells us that the initial condition is well-concentrated on the scale of n. Assumption (3) gives that, conditionally on P^n being in state $\lfloor nx \rfloor$ at time $\lfloor ns \rfloor$, the size of the expected increment of P^n is close to $\nu(s, x)$. It is then natural to compare P^n/n to the solution to the differential equation in (2), as long as P^n/n remains within the set $\mathcal{U}$.

Proposition 61. *Under assumptions (1), (2), (3) and (4), we have*

$$\sup_{0 \leq t \leq T_n} \left| \frac{P^n(\lfloor nt \rfloor)}{n} - p(t) \right| \xrightarrow{p} 0$$

as $n \to \infty$.

Proof. Since p solves the given differential equation, we may write it in integral form as

$$p(t) = p(0) + \int_0^t \nu(s, p(s)) ds.$$

Hence,

$$\begin{aligned}\frac{P^n(\lfloor nt\rfloor)}{n} &- p(t)\\ &= \frac{P^n(0)}{n} - p(0) + \frac{M^n_P(\lfloor nt\rfloor)}{n} + \sum_{j=0}^{\lfloor nt\rfloor-1}\frac{\mathbb{E}[P^n(j+1)-P^n(j)|\mathcal{F}^n_P(j)]}{n} - \int_0^t \nu(s,p(s))ds.\end{aligned}$$

By Doob's L^2 inequality, assumption (4) implies that $\sup_{0\le t_0\wedge T_n}\frac{|M^n_P(i)|}{n}\xrightarrow{p} 0$ as $n\to\infty$. Fix $\delta>0$ and let

$$\begin{aligned}\Omega_{n,\delta} = &\left\{\left|\frac{P^n(0)}{n} - p(0)\right| \le \frac{\delta e^{-Kt_0}}{3}\right\}\cap\left\{\sup_{0\le t\le t_0\wedge T_n}\frac{|M^n_P(\lfloor nt\rfloor)|}{n}\le\frac{\delta e^{-Kt_0}}{3}\right\}\\ &\cap\left\{\sup_{0\le t\le t_0\wedge T_n}\left|\sum_{j=0}^{\lfloor nt\rfloor-1}\frac{\mathbb{E}[P^n(j+1)-P^n(j)|\mathcal{F}^n_P(j)]}{n} - \int_0^t\nu\left(s,\frac{P^n(\lfloor ns\rfloor)}{n}\right)ds\right|\le\frac{\delta e^{-Kt_0}}{3}\right\}.\end{aligned}$$

By the assumptions, we have $\mathbb{P}\left(\Omega^c_{n,\delta}\right)\to 0$ as $n\to\infty$. On the event $\Omega_{n,\delta}$, we have that for $0\le t\le t_0\wedge T_n$,

$$\begin{aligned}\left|\frac{P^n(\lfloor nt\rfloor)}{n} - p(t)\right| &\le \delta e^{-Kt_0} + \int_0^t\left|\nu\left(s,\frac{P^n(\lfloor ns\rfloor)}{n}\right)ds - \nu(s,p(s))\right|ds\\ &\le \delta e^{-Kt_0} + K\int_0^t\left|\frac{P^n(\lfloor ns\rfloor)}{n} - p(s)\right|ds,\end{aligned}$$

by the Lipschitz property of ν on $\mathcal{U}$. Hence, by Gronwall's lemma, on the event $\Omega_{n,\delta}$,

$$\sup_{0\le t\le t_0\wedge T_n}\left|\frac{P^n(\lfloor nt\rfloor)}{n} - p(t)\right|\le\delta.$$

The result follows. □

We begin with a preparatory lemma.

Lemma 62. *For $t<1$ and $0\le i\le\lfloor nt\rfloor$ we have*

$$\begin{aligned}\mathbb{E}[U^n(i+1)-U^n(i)|\mathcal{F}^n_i] &= -1 + E^n_{U,1}(i)\\ \mathbb{E}\Big[(U^n(i+1)-U^n(i))^2|\mathcal{F}^n_i\Big] &= 1 + E^n_{U,2}(i)\\ \mathbb{E}[V^n(i+1)-V^n(i)|\mathcal{F}^n_i] &= \frac{-4V^n(i)}{3U^n(i)} + E^n_V(i)\\ \mathbb{E}\Big[(V^n(i+1)-V^n(i))^2|\mathcal{F}^n_i\Big] &= \frac{4V^n(i)}{3U^n(i)} - E^n_V(i)\end{aligned}$$

where $\sup_{0\leq i\leq \lfloor nt\rfloor}|E^n_{U,1}(i)|$, $\sup_{0\leq i\leq \lfloor nt\rfloor}|E^n_{U,2}(i)|$, *and* $\frac{n}{\alpha(n)}\sup_{0\leq i\leq \lfloor nt\rfloor}|E^n_V(i)|$ *all converge to zero in* L^2 *as* $n\to\infty$.

Proof. Let us note that

$$\mathbb{E}[U^n(i+1)-U^n(i)|\mathcal{F}^n_i] = -1+\frac{8V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}{3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}.$$

Having in mind that in each step we remove at most one vertex, for $0\leq i\leq \lfloor nt\rfloor$ we have $U^n(i)\geq n-i$, and so

$$\sup_{0\leq i\leq \lfloor nt\rfloor}\frac{8V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}{3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}\leq \sup_{0\leq i\leq \lfloor nt\rfloor}\frac{8\alpha(n)+X^n(i)}{3(n-i)},$$

which converges to 0 in L^2 as $n\to\infty$ by Lemma 60. Also,

$$\mathbb{E}\Big[(U^n(i+1)-U^n(i))^2|\mathcal{F}^n_i\Big] = 1-\frac{6V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}{3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}$$

and

$$\sup_{0\leq i\leq \lfloor nt\rfloor}\frac{6V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}{3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}\leq \sup_{0\leq i\leq \lfloor nt\rfloor}\frac{6\alpha(n)+X^n(i)}{3(n-i)},$$

which again by Lemma 60 converges to 0 in L^2 as $n\to\infty$.

Turning now to V^n, we see that

$$\mathbb{E}[V^n(i+1)-V^n(i)|\mathcal{F}^n_i] = -\frac{4V^n(i)}{3U^n(i)}+\frac{4V^n(i)}{3U^n(i)}\cdot\frac{4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}{3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}},$$

$$\mathbb{E}\Big[(V^n(i+1)-V^n(i))^2|\mathcal{F}^n_i\Big] = \frac{4V^n(i)}{3U^n(i)}-\frac{4V^n(i)}{3U^n(i)}\cdot\frac{4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}{3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}.$$

Letting $E^n_V(i)=\frac{4V^n(i)}{3U^n(i)}\cdot\frac{4V^n(i)+X^n(i)-\mathbb{1}_{\{+X^n(i)>0\}}}{3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{+X^n(i)>0\}}}$, we have

$$\sup_{0\leq i\leq \lfloor nt\rfloor}|E^n_V(i)|\leq \frac{4\alpha(n)}{3(n-i)}\frac{4\sup_{0\leq i\leq \zeta_n}(X^n(i)+V^n(i))}{3(n-i)}$$

and so Lemma 60 implies that $\sup_{0\leq i\leq \lfloor nt\rfloor}\frac{n}{\alpha(n)}|E^n_V(i)|\to 0$ in L^2 as $n\to\infty$. □

Lemma 63. *Let $t < 1$. Then as $n \to \infty$, we have*

$$\sup_{0 \le i \le \lfloor nt \rfloor} \left| \frac{U^n(i)}{n} - \left(1 - \frac{i}{n}\right) \right| \xrightarrow{p} 0.$$

If, in addition, we assume that $\alpha(n) \to \infty$ as $n \to \infty$, then

$$\sup_{0 \le i \le \lfloor nt \rfloor} \left| \frac{V^n(i)}{\alpha(n)} - \left(1 - \frac{i}{n}\right)^{4/3} \right| \xrightarrow{p} 0.$$

Proof. Having in mind the size of expected increments given by Lemma 62, and $U^n(0) = n$ and $V^n(0) = \alpha(n)$, the candidate for the fluid limit is the solution to the system of differential equations given by

$$\begin{aligned} \frac{du(s)}{ds} &= -1, \\ \frac{dv(s)}{ds} &= -\frac{4v(s)}{3u(s)}, \end{aligned} \tag{3.8}$$

for $0 \le s \le t$ and the initial condition $u(0) = 1$, $v(0) = 1$. The solution is given by $u(s) = 1 - s$, $v(s) = (1 - s)^{4/3}$. We observe that $u(s)$ and $v(s)$ remain bounded away from 0 for all $0 \le s \le t$.

Let

$$T_U^n = \inf\{s \ge 0 : U^n(\lfloor ns \rfloor) = 0\}.$$

Since $U^n(i) \ge n - i$, it is clear that $T_U^n \ge 1$ for all n and so, in particular, $T_U^n > t$. For $i \ge 0$, let

$$M_U^n(i) = U^n(i) - U^n(0) - \sum_{j=0}^{i-1} \mathbb{E}\Big[U^n(j+1) - U^n(j) | \mathcal{F}_j^n\Big],$$

so that $(M_U^n(i))_{i \ge 0}$ is a martingale. By Lemma 62, we have that for $0 \le s \le t$,

$$\left| \sum_{j=0}^{\lfloor ns \rfloor - 1} \frac{\mathbb{E}\Big[U^n(j+1) - U^n(j) | \mathcal{F}_j^n\Big]}{n} + \int_0^s du \right| \le \sup_{0 \le s \le t} \left| \frac{\lfloor ns \rfloor}{n} - s \right| + t \sup_{0 \le i \le \lfloor nt \rfloor} |E_{U,1}^n(i)| \xrightarrow{p} 0,$$

as $n \to \infty$. Moreover,

$$\mathbb{E}\Big[M_U^n(i)^2\Big] = \mathbb{E}\left[\sum_{j=0}^{i-1} \mathrm{var}\Big(U^n(j+1) - U^n(j) | \mathcal{F}_j^n\Big)\right] \le i\left(1 + \mathbb{E}\left[\sup_{0 \le j \le i} |E_{U,2}^n(j)|\right]\right)$$

and so by Lemma 62 we have

$$\frac{\mathbb{E}[M_U^n(\lfloor nt)\rfloor)^2]}{n^2} \leq \frac{\lfloor nt\rfloor(1+\mathbb{E}\Big[\sup_{0\leq i\leq \lfloor nt\rfloor}|E_{U,2}^n(i)|\Big])}{n^2} \to 0. \tag{3.9}$$

Hence, by Proposition 61,

$$\sup_{0\leq s\leq t}\left|\frac{U^n(\lfloor ns\rfloor)}{n} - u(s)\right| \xrightarrow{p} 0.$$

We now turn to V^n. For $i \geq 0$, let

$$M_V^n(i) = V^n(i) - V^n(0) - \sum_{j=0}^{i-1}\mathbb{E}\Big[V^n(j+1) - V^n(j)|\mathcal{F}_j^n\Big],$$

be the standard martingale. Let us deal first with its second moment. We have

$$\mathbb{E}\Big[M_V^n(i)^2\Big] = \mathbb{E}\left[\sum_{j=0}^{i-1}\operatorname{var}\Big(V^n(j+1) - V^n(j)|\mathcal{F}_j^n\Big)\right] \leq i\cdot\sup_{0\leq j\leq i}\left(\left|\frac{4V^n(j)}{3U^n(j)}\right| + |E_V^n(j)|\right).$$

Since $\frac{4V^n(j)}{3U^n(j)} \leq \frac{4\alpha(n)}{3(n-j)}$, Lemma 62 gives us

$$\frac{\mathbb{E}[M_V^n(\lfloor nt\rfloor)^2]}{\alpha(n)^2} \leq \frac{\lfloor nt\rfloor\left(\frac{4\alpha(n)}{3(n-\lfloor nt\rfloor)} + \mathbb{E}\Big[\sup_{0\leq i\leq\lfloor nt\rfloor}|E_V^n(i)|\Big]\right)}{\alpha(n)^2} \to 0. \tag{3.10}$$

Turning now to the drift, we have

$$\begin{aligned}
&\sup_{0\leq s\leq t}\left|\sum_{j=0}^{\lfloor ns\rfloor-1}\frac{\mathbb{E}\Big[V^n(j+1)-V^n(j)|\mathcal{F}_j^n\Big]}{\alpha(n)} + \int_0^s\frac{4v(r)}{3u(r)}dr\right| \\
&\quad\leq \sup_{0\leq s\leq t}\left|\sum_{j=0}^{\lfloor ns\rfloor-1}\frac{4V^n(i)/\alpha(n)}{3U^n(i)} - \int_0^s\frac{4v(r)}{3u(r)}dr\right| + \frac{nt}{\alpha(n)}\sup_{0\leq i\leq\lfloor nt\rfloor}|E_V^n(i)| \\
&\quad\leq \sup_{0\leq s\leq t}\left|\int_0^{\lfloor ns\rfloor/n}\left(\frac{4V^n(\lfloor nr\rfloor)/\alpha(n)}{3U^n(\lfloor nr\rfloor)/n} - \frac{4v(r)}{3u(r)}\right)dr\right| + \sup_{0\leq s\leq t}\left|\int_{\lfloor ns\rfloor/n}^s\frac{4v(r)}{3u(r)}dr\right| \\
&\qquad+ \frac{nt}{\alpha(n)}\sup_{0\leq i\leq\lfloor nt\rfloor}|E_V^n(i)|.
\end{aligned}$$

The penultimate term on the right-hand side clearly tends to 0, and the last term converges to 0 in probability by Lemma 62. We have

$$\begin{aligned}
&\sup_{0\le s\le t}\left|\int_0^{\lfloor ns\rfloor/n}\left(\frac{4V^n(\lfloor nr\rfloor)/\alpha(n)}{3U^n(\lfloor nr\rfloor)/n}-\frac{4v(r)}{3u(r)}\right)dr\right|\\
&\le \int_0^{\lfloor nt\rfloor/n}\frac{4V^n(\lfloor ns\rfloor)/\alpha(n)}{3u(s)}\cdot\left|\frac{u(s)}{U^n(\lfloor ns\rfloor)/n}-1\right|ds+\int_0^{\lfloor nt\rfloor/n}\frac{4}{3u(s)}\left|\frac{V^n(\lfloor ns\rfloor)}{\alpha(n)}-v(s)\right|ds\\
&\le \frac{4}{3(1-t)}\sup_{0\le s\le t}\left|\frac{u(s)}{U^n(\lfloor ns\rfloor)/n}-1\right|+\frac{4}{3(1-t)}\int_0^t\left|\frac{V^n(\lfloor ns\rfloor)}{\alpha(n)}-v(s)\right|ds.
\end{aligned}$$

The first term on the right-hand side converges to 0 in probability. So following the argument in the proof of Proposition 61, we easily obtain

$$\sup_{0\le s\le t}\left|\frac{V^n(\lfloor ns\rfloor)}{\alpha(n)}-v(s)\right|\xrightarrow{p}0.$$

□

3.4 Fluid limit for $\alpha(n) >> \sqrt{n}$

In the case where $\alpha(n) >> \sqrt{n}$, we will also prove fluid limits for X^n and N^n.

Theorem 64. *Fix $t\in(0,1)$. For $\alpha(n) >> \sqrt{n}$, as $n\to\infty$ we have*

$$\begin{aligned}
\sup_{0\le s\le t}\left|\frac{X^n(\lfloor ns\rfloor)}{\alpha(n)}-\left((1-s)^{2/3}-(1-s)^{4/3}\right)\right|&\xrightarrow{p}0,\\
\sup_{0\le s\le t}\left|\frac{N^n(\lfloor ns\rfloor)}{\alpha(n)}-\left(\frac{1}{4}-\frac{1}{2}(1-s)^{2/3}+\frac{1}{4}(1-s)^{4/3}\right)\right|&\xrightarrow{p}0,\\
\frac{L^n(\lfloor nt\rfloor)}{\alpha(n)}&\xrightarrow{p}0.
\end{aligned}$$

To this end, we again study the expected increments.

Lemma 65. *For $t<1$ and $0\le i\le\lfloor nt\rfloor$ we have*

$$\begin{aligned}
\mathbb{E}[X^n(i+1)-X^n(i)|\mathcal{F}_i^n]&=2\mathbb{1}_{\{X^n(i)=0\}}+\frac{2\alpha(n)(1-i/n)^{4/3}-2X^n(i)}{3(n-i)}+E^n_{X,1}(i),\\
\mathbb{E}\Big[(X^n(i+1)-X^n(i))^2|\mathcal{F}_i^n\Big]&=4\mathbb{1}_{\{X^n(i)=0\}}+1+E^n_{X,2}(i),\\
\mathbb{E}[N^n(i+1)-N^n(i)|\mathcal{F}_i^n]=\mathbb{E}\Big[(N^n(i+1)-N^n(i))^2|\mathcal{F}_i^n\Big]&=\frac{X^n(i)}{3(n-i)}+E^n_N(i),
\end{aligned}$$

where

$$\frac{n}{\alpha(n)\vee\sqrt{n}}\sup_{0\le i\le\lfloor nt\rfloor}|E^n_{X,1}(i)|\xrightarrow{p}0,$$

and

$$\sup_{0\le i\le \lfloor nt\rfloor} |E^n_{X,2}(i)| \to 0,$$
$$\frac{n}{\alpha(n)\vee\sqrt{n}} \sup_{0\le i\le \lfloor nt\rfloor} |E^n_N(i)| \to 0,$$

in L^2 *as* $n\to\infty$.

Proof. By (3.3), we have

$$\begin{aligned}
&\mathbb{E}[X^n(i+1)-X^n(i)|\mathcal{F}^n_i] \\
&= 2\mathbb{1}_{\{X^n(i)=0\}} + \frac{2\alpha(n)(1-i/n)^{4/3}-2X^n(i)}{3(n-i)} + \frac{2V^n(i)-2\alpha(n)(1-i/n)^{4/3}+2\mathbb{1}_{\{X^n(i)>0\}}}{3(n-i)} \\
&\quad + \frac{2V^n(i)-2X^n(i)+2\mathbb{1}_{\{X^n(i)>0\}}}{3(n-i)}\cdot\left(\left(1+\tfrac{3U^n(i)-3(n-i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}{3(n-i)}\right)^{-1}-1\right) \\
&=: 2\mathbb{1}_{\{X^n(i)=0\}} + \frac{2\alpha(n)(1-i/n)^{4/3}-2X^n(i)}{3(n-i)} + E^n_{X,1}(i).
\end{aligned}$$

For $0\le i\le \lfloor nt\rfloor$, we have

$$\begin{aligned}
&\frac{n}{\alpha(n)}\sup_{0\le i\le \lfloor nt\rfloor}\left|\frac{2V^n(i)-2\alpha(n)(1-i/n)^{4/3}+2\mathbb{1}_{\{X^n(i)>0\}}}{3(n-i)}\right| \\
&\qquad \le \frac{2}{3(1-t)}\sup_{0\le i\le \lfloor nt\rfloor}\left|\frac{V^n(i)}{\alpha(n)}-\left(1-\frac{i}{n}\right)^{4/3}\right| + \frac{2}{3(1-t)\alpha(n)} \xrightarrow{p} 0,
\end{aligned} \tag{3.11}$$

by the fluid limit for V^n in Lemma 63. Since $(1+x)^{-1}\ge 1-x$ for $x>0$, and $U^n(i)\ge n-i$, we have

$$\begin{aligned}
&\sup_{0\le i\le \lfloor nt\rfloor}\left|\left(1+\tfrac{3U^n(i)-3(n-i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}{3(n-i)}\right)^{-1}-1\right| \\
&\quad\le \sup_{0\le i\le \lfloor nt\rfloor}\left|\frac{3U^n(i)-3(n-i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}{3(n-i)}\right| \\
&\quad\le \tfrac{1}{3(1-t)}\left(3\sup_{0\le i\le \lfloor nt\rfloor}\left|\tfrac{U^n(i)}{n}-\left(1-\tfrac{i}{n}\right)\right| + \tfrac{4\alpha(n)}{n} + \tfrac{1+\sup_{0\le i\le \lfloor nt\rfloor}X^n(i)}{n}\right) \xrightarrow{p} 0
\end{aligned} \tag{3.12}$$

by Lemma 60 (which says that X^n is of smaller order than n), and the fluid limit for U^n from Lemma 63.

By Lemma 60, we have that

$$\frac{1}{\alpha(n) \vee \sqrt{n}} \sup_{0 \le i \le \lfloor nt \rfloor} (X^n(i) + V^n(i))$$

is tight in n. It follows that

$$\frac{n}{\alpha(n) \vee \sqrt{n}} \sup_{0 \le i \le \lfloor nt \rfloor} \frac{2V^n(i) - 2X^n(i) + 2\mathbb{1}_{\{X^n(i)>0\}}}{3(n-i)}$$

is also tight, and so together with (3.11) and (3.12) we get

$$\frac{n}{\alpha(n) \vee \sqrt{n}} \sup_{0 \le i \le \lfloor nt \rfloor} |E^n_{X,1}(i)| \xrightarrow{p} 0.$$

For the second moment, rewriting the expression (3.4) we have

$$\begin{aligned}
&\mathbb{E}\Big[(X^n(i+1) - X^n(i))^2 | \mathcal{F}^n_i\Big] \\
&\quad = 4\mathbb{1}_{\{X^n(i)=0\}} + 1 + \frac{6V^n(i) + 8V^n(i)\mathbb{1}_{\{X^n(i)=0\}} + 3X^n(i) - 3\mathbb{1}_{\{X^n(i)>0\}}}{3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}}} \\
&\quad =: 4\mathbb{1}_{\{X^n(i)=0\}} + 1 + E^n_{X,2}(i).
\end{aligned}$$

But then

$$\sup_{0 \le i \le \lfloor nt \rfloor} |E^n_{X,2}(i)| \le \frac{14\alpha(n) + 3\sup_{0 \le i \le \lfloor nt \rfloor} X^n(i)}{3n(1-t)}$$

which converges to 0 in probability and in L^2, by Lemma 60. Turning now to the increments of N^n, we have

$$\begin{aligned}
&\mathbb{E}[N^n(i+1) - N^n(i) | \mathcal{F}^n_i] \\
&\quad = \mathbb{E}\Big[(N^n(i+1) - N^n(i))^2 | \mathcal{F}^n_i\Big] \\
&\quad = \frac{X^n(i)}{3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}}} \\
&\qquad + \frac{3U^n(i) - 2\mathbb{1}_{\{X^n(i)>0\}}(3U^n(i) + 4V^n(i) + X^n(i) - 3)}{2(3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}})(3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}} - 2)} \\
&\quad = \frac{X^n(i)}{3(n-i)} + \frac{X^n(i)}{3(n-i)}\left(\left(1 + \tfrac{3U^n(i) - 3n(1-i/n) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}}}{3(n-i)}\right)^{-1} - 1\right) \\
&\qquad + \frac{3U^n(i) - 2\mathbb{1}_{\{X^n(i)>0\}}(3U^n(i) + 4V^n(i) + X^n(i) - 3)}{2(3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}})(3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}} - 2)}
\end{aligned}$$

$$=: \frac{X^n(i)}{3(n-i)} + E_N^n(i).$$

For $0 \le i \le \lfloor nt \rfloor$ we have

$$\left| \frac{3U^n(i) - 2\mathbb{1}_{\{X^n(i)>0\}}(3U^n(i) + 4V^n(i) + X^n(i) - 3)}{2(3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}})(3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}} - 2)} \right|$$
$$\le \frac{3}{2(3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}})}$$
$$\le \frac{3}{2n(1-t)}.$$

Hence,

$$\frac{n}{\alpha(n) \vee \sqrt{n}} \sup_{0 \le i \le \lfloor nt \rfloor} |E_N^n(i)|$$
$$\le \frac{3}{2(1-t)(\alpha(n) \vee \sqrt{n})}$$
$$+ \frac{\sup_{0 \le i \le \lfloor nt \rfloor} X^n(i)}{3(1-t)(\alpha(n) \vee \sqrt{n})} \left(\left(1 + \frac{3U^n(i) - 3n(1-i/n) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}}}{3(n-i)}\right)^{-1} - 1 \right).$$

Using (3.12) and the fact that $\frac{1}{\alpha(n) \vee \sqrt{n}} \sup_{0 \le i \le \lfloor nt \rfloor} X^n(i)$ is bounded in L^2, this gives that

$$\frac{n}{\alpha(n) \vee \sqrt{n}} \sup_{0 \le i \le \lfloor nt \rfloor} |E_N^n(i)| \to 0,$$

in L^2, as desired. □

Before we can proceed to the proof of Theorem 64, we need some technical lemmas to deal with the fact that X^n reflects off 0. Let

$$\gamma^n = \inf\{i \ge 0 : X^n(i) \ge V^n(i)\}.$$

Lemma 66. *Fix $t \in (1/2, 1)$. Then there exists $t_0 \in (0, t)$ sufficiently small that*

$$\mathbb{P}(\gamma^n < \lfloor t_0 n \rfloor) \to 0$$

as $n \to \infty$.

Proof. Let $\tilde{\gamma}^n = \inf\{i \geq 0 : X^n(i) \geq 3\alpha(n)/8\}$. Observe that V^n is decreasing and that if $s_0 = 1 - \frac{1}{2^{3/4}} \approx 0.405$ then $(1-s_0)^{4/3} = 1/2 > 3/8$. Then by Lemma 63,

$$\mathbb{P}(V^n(\lfloor ns_0 \rfloor) > 3\alpha(n)/8) \to 1$$

as $n \to \infty$. So if we can show that there exists $s_1 \in (0,t)$ such that

$$\mathbb{P}(\tilde{\gamma}^n < ns_1) \to 0 \tag{3.13}$$

as $n \to \infty$ then, setting $t_0 = s_0 \wedge s_1$, we obtain

$$\mathbb{P}(\gamma^n < nt_0) \to 0.$$

So it remains to prove (3.13) for some $s_1 \in (0,t)$.

The time-inhomogeneity of the process X^n makes explicit calculations awkward. We instead couple X^n with a simpler process A^n. If $A^n(i) = 0$ then $A^n(i+1) = 1$. If $A^n(i) > 0$ then

$$A^n(i+1) - A^n(i) = \begin{cases} -1 & \text{with probability } \frac{1}{2} \\ +1 & \text{with probability } \frac{1}{2} - \frac{4\alpha(n)}{3n(1-t)} \\ +2 & \text{with probability } \frac{4\alpha(n)}{3n(1-t)}. \end{cases}$$

(We implicitly take n sufficiently large that $\frac{4\alpha(n)}{3n(1-t)} < 1/2$.) We set $A^n(0) = 0$. It is straightforward to see that we may couple X^n and A^n in such a way that $X^n(i) \leq 3 + A^n(i)$ for all $0 \leq i \leq \lfloor nt \rfloor$. Let $\lambda_n = \inf\{i \geq 0 : A^n(i) \geq 3\alpha(n)/8 - 3\}$. Then we certainly have $\lambda_n \leq \tilde{\gamma}_n$. So we will find $s_1 \in (0,t)$ such that

$$\mathbb{P}(\lambda_n < ns_1) \to 0.$$

For simplicity, write $b = \lfloor 3\alpha(n)/8 \rfloor - 3$ for the upper barrier and $d = \frac{4\alpha(n)}{3n(1-t)}$ for the drift. The quantity

$$h_1 = \mathbb{P}(A^n \text{ hits } 0 \text{ before } \{b, b+1\} | A^n(0) = 1)$$

is calculated via a standard calculation in Lemma 75 in the Appendix. Substituting the given values of b and d in the (lengthy) expression there, we obtain that

$$1 - h_1 \sim \frac{8\alpha(n)}{3n(1-t)},$$

as $n \to \infty$. By the strong Markov property, the random walk A^n visits 0 a Geometric($1-h_1$) number of times before going above $3\alpha(n)/8 - 3$, and so

$$\frac{1}{\alpha(n)}\mathbb{E}\left[\sum_{i=0}^{\lambda_n - 1} \mathbb{1}_{\{A^n(i)=0\}}\right] \sim \frac{3n(1-t)}{8\alpha(n)^2} \to 0,$$

since $\alpha(n) >> \sqrt{n}$.

Now let

$$M_A^n(i) := A^n(i) - \frac{4\alpha(n)i}{3n(1-t)} - \sum_{j=0}^{i-1}\left(1 - \frac{4\alpha(n)}{3n(1-t)}\right)\mathbb{1}_{\{A^n(j)=0\}}.$$

Then $(M_A^n(i))_{i\geq 0}$ is a martingale. Let $\mathcal{F}_A^n(j)$ be the natural filtration of A^n. We have

$$\text{var}(A^n(j+1) - A^n(j)|\mathcal{F}_A^n(j)) \leq \left(1 + \frac{4\alpha(n)}{n(1-t)}\right)\mathbb{1}_{\{A^n(j)>0\}} \leq 1 + \frac{4\alpha(n)}{n(1-t)},$$

and so

$$\mathbb{E}\Big[(M_A^n(i))^2\Big] \leq i\left(1 + \frac{4\alpha(n)}{n(1-t)}\right) \leq 2i$$

for n sufficiently large. For any $s > 0$, then,

$$\sup_{0\leq i\leq \lambda_n \wedge \lfloor ns \rfloor}\left|\frac{A^n(i)}{\alpha(n)} - \frac{4i}{3n(1-t)}\right| \leq \frac{\sup_{0\leq i\leq \lfloor ns \rfloor}|M_A^n(i)|}{\alpha(n)} + \frac{1}{\alpha(n)}\sum_{j=0}^{\lambda_n - 1}\mathbb{1}_{\{A^n(i)=0\}}.$$

For any $\delta > 0$, it then follows by using Doob's L^2 inequality that

$$\begin{aligned}\mathbb{P}\left(\sup_{0\le i\le \lambda_n\wedge\lfloor ns\rfloor}\left|\frac{A^n(i)}{\alpha(n)}-\frac{4i}{3n(1-t)}\right|>2\delta\right) &\le 4\frac{\mathbb{E}[M_A^n(\lfloor ns\rfloor)^2]}{\delta^2\alpha(n)^2}+\frac{1}{\delta\alpha(n)}\mathbb{E}\left[\sum_{j=0}^{\lambda_n-1}\mathbb{1}_{\{A^n(i)=0\}}\right]\\ &\le \frac{8\lfloor ns\rfloor}{\delta^2\alpha(n)^2}+\frac{1}{\delta\alpha(n)}\mathbb{E}\left[\sum_{j=0}^{\lambda_n-1}\mathbb{1}_{\{A^n(i)=0\}}\right]\to 0,\end{aligned}$$

as $n\to\infty$.

Finally, if we take $s_1 = 3(1-t)/16$ then

$$\frac{4\lfloor ns_1\rfloor}{3n(1-t)}\le\frac{1}{4}<\frac{3}{8}.$$

Then for $\delta\in(0,1/8)$,

$$\mathbb{P}(\lambda_n<ns_1)\le\mathbb{P}\left(\sup_{0\le i\le\lambda_n\wedge\lfloor ns_1\rfloor}\left|\frac{A^n(i)}{\alpha(n)}-\frac{4i}{3n(1-t)}\right|>\delta\right)\to 0$$

as $n\to\infty$. The result follows. □

Proposition 67. *For the $t_0\in(0,t)$ from Lemma 66,*

$$\frac{L^n(\lfloor nt_0\rfloor)}{\alpha(n)}\xrightarrow{p}0,$$

as $n\to\infty$.

Proof. We again make use of a coupling, this time to produce a lower bound: we define a process D^n such that $D^n(i)\le X^n(i)$ for $i\le\gamma^n-1$. Conditionally on $(U^n(i),V^n(i),X^n(i),D^n(i))=(u,v,x,d)$ with $D^n(i)=d>1$, we let

$$D^n(i+1)-D^n(i)=\begin{cases}+2 & \text{with probability } \frac{4(x-1)\mathbb{1}_{\{x>0\}}}{2(3u+4v+(x-1)\mathbb{1}_{\{x>0\}})}\\ +1 & \text{with probability } \frac{3u+4v-4(x-1)\mathbb{1}_{\{x>0\}}}{2(3u+4v+(x-1)\mathbb{1}_{\{x>0\}})}\\ -1 & \text{with probability } \frac{3u+4v}{2(3u+4v+(x-1)\mathbb{1}_{\{x>0\}})}\\ -2 & \text{with probability } \frac{2(x-1)\mathbb{1}_{\{x>0\}}}{2(3u+4v+(x-1)\mathbb{1}_{\{x>0\}})}.\end{cases}$$

Conditionally on $(U^n(i), V^n(i), X^n(i), D^n(i)) = (u, v, x, d)$ with $D^n(i) = d = 1$, we let

$$D^n(i+1) - D^n(i) = \begin{cases} +2 & \text{with probability } \frac{2(x-1)\mathbb{1}_{\{x>0\}}}{2(3u+4v+(x-1)\mathbb{1}_{\{x>0\}})} \\ +1 & \text{with probability } \frac{3u+4v-2(x-1)\mathbb{1}_{\{x>0\}}}{2(3u+4v+(x-1)\mathbb{1}_{\{x>0\}})} \\ -1 & \text{with probability } \frac{3u+4v+2(x-1)\mathbb{1}_{\{x>0\}}}{2(3u+4v+(x-1)\mathbb{1}_{\{x>0\}})}. \end{cases}$$

Conditionally on $(U^n(i), V^n(i), X^n(i), D^n(i)) = (u, v, x, d)$ with $D^n(i) = 0$, let us have $D^n(i+1) = 1$. Since we assume $i \leq \gamma^n - 1$, it is straightforward to see that we may produce a coupling such that $D^n(i) \leq X^n(i)$. Let

$$\tilde{L}^n(i) = 2\sum_{j=0}^{i-1} \mathbb{1}_{\{D^n(j)=0\}},$$

and observe that we have $\tilde{L}^n(\gamma^n \wedge \lfloor nt_0 \rfloor) \geq L^n(\gamma^n \wedge \lfloor nt_0 \rfloor)$.

But then for any $\delta > 0$,

$$\mathbb{P}\left(\frac{L^n(\lfloor nt_0 \rfloor)}{\alpha(n)} > \delta\right) \leq \mathbb{P}\left(\frac{\tilde{L}^n(\gamma^n \wedge \lfloor nt_0 \rfloor)}{\alpha(n)} > \delta, \gamma^n \geq \lfloor nt_0 \rfloor\right) + \mathbb{P}(\gamma^n < \lfloor nt_0 \rfloor).$$

By Lemma 66, we have that $\mathbb{P}(\gamma^n < \lfloor nt_0 \rfloor) \to 0$ and so it suffices to prove that the first term tends to 0.

We have

$$\mathbb{E}[D^n(i+1) - D^n(i) | \mathcal{F}_D^n(i)] = \mathbb{1}_{\{D^n(i)=0\}},$$

where $(\mathcal{F}_D^n(i))_{i\geq 0}$ denotes the natural filtration of (U^n, V^n, X^n, D^n). Also,

$$\begin{aligned} \operatorname{var}(D^n(i+1) - D^n(i) | \mathcal{F}_D^n(i)) = 1 &+ \frac{9(X^n(i) - \mathbb{1}_{\{X^n(i)>0\}})}{3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}}} \mathbb{1}_{\{D^n(i)>1\}} \\ &+ \frac{3(X^n(i) - \mathbb{1}_{\{X^n(i)>0\}})}{3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}}} \mathbb{1}_{\{D^n(i)=1\}}, \end{aligned}$$

where

$$\frac{12(X^n(i) - \mathbb{1}_{\{X^n(i)>0\}})}{3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}}} \leq \frac{4\alpha(n)}{n(1-s)},$$

for $i \leq \gamma^n \wedge \lfloor ns \rfloor$.

Hence,

$$\left(D^n(i) - \sum_{j=0}^{i-1} \mathbb{1}_{\{D^n(j)=0\}}\right)_{i\geq 0}$$

is a mean 0 martingale, with bounded steps and

$$\sum_{j=0}^{i-1} \mathbb{E}\Big[\Big(D^n(j+1) - D^n(j) - \mathbb{1}_{\{D^n(j)=0\}}\Big)^2 \Big| \mathcal{F}_D^n(j)\Big] = i + E_D^n(i)$$

for $i \leq \gamma^n \wedge \lfloor nt_0 \rfloor$, where $\sup_{i\leq\gamma^n\wedge\lfloor nt_0\rfloor} \frac{n}{\alpha(n)}|E_D^n(i)|$ is bounded with high probability. By the martingale functional central limit theorem (Theorem 1.4, Section 7.1 of Ethier and Kurtz [21]), we then have that, on the event $\{\gamma^n \geq \lfloor nt_0 \rfloor\}$,

$$\frac{1}{\sqrt{n}}\left(D^n(\lfloor nu \rfloor) - \sum_{j=0}^{\lfloor nu \rfloor -1} \mathbb{1}_{\{D^n(j)=0\}}\right)_{0\leq u\leq t_0} \xrightarrow{d} (W_u)_{0\leq u\leq t_0},$$

where W is a standard Brownian motion. But we also have that

$$\frac{1}{\sqrt{n}} \sum_{j=0}^{\lfloor nt_0 \rfloor -1} \mathbb{1}_{\{D^n(j)=0\}} = -\inf_{0\leq i\leq \lfloor nt_0 \rfloor} \frac{1}{\sqrt{n}}\left(D^n(i) - \sum_{j=0}^{i-1} \mathbb{1}_{\{D^n(j)=0\}}\right),$$

and so by the continuous mapping theorem we deduce that on $\{\gamma^n \geq \lfloor nt_0 \rfloor\}$,

$$\frac{\tilde{L}^n(\lfloor nt_0 \rfloor)}{\sqrt{n}} = \frac{2}{\sqrt{n}} \sum_{j=0}^{\lfloor nt_0 \rfloor -1} \mathbb{1}_{\{D^n(j)=0\}} \xrightarrow{d} -2 \inf_{0\leq u\leq t_0} W_u.$$

The result then follows since $\alpha(n) >> \sqrt{n}$. $\square$

We now turn to the proof of the main theorem in this section.

Proof of Theorem 64. Lemma 65 gives us that the expected increments of X^n and N^n are

$$\mathbb{E}[X^n(i+1) - X^n(i)|\mathcal{F}_i^n] = 2\mathbb{1}_{\{X^n(i)=0\}} + \frac{\alpha(n)}{n}\left(\frac{2(1-i/n)^{1/3}}{3} - \frac{2X^n(i)/\alpha(n)}{3(1-i/n)}\right) + E_{X,1}^n(i),$$

$$\mathbb{E}[N^n(i+1) - N^n(i)|\mathcal{F}_i^n] = \frac{\alpha(n)}{n}\frac{X^n(i)/\alpha(n)}{3(1-i/n)} + E_N^n(i).$$

This suggests that, as long as X^n does not hit 0 too often, the candidate for the fluid limit should be the solution to

$$\frac{dx(s)}{ds} = \frac{2}{3}(1-s)^{1/3} - \frac{2x(s)}{3(1-s)}, \quad \frac{dm(s)}{ds} = \frac{x(s)}{3(1-s)}, \tag{3.14}$$

with initial conditions $x(0) = 0$, $m(0) = 0$. The unique solution to this system of differential equations is given by $x(s) = (1-s)^{2/3} - (1-s)^{4/3}$ and $m(s) = \frac{1}{4} - \frac{1}{2}(1-s)^{2/3} + \frac{1}{4}(1-s)^{4/3}$ for $s \in [0,1]$. We will use Proposition 67 to help control the local time at the start, and then prove the fluid limit result in the standard manner.

Let

$$T^n = \inf\{s \geq t_0 : X^n(\lfloor ns \rfloor) = 0\}.$$

Then for $t \geq t_0$ we have

$$L^n(\lfloor n(t \wedge T^n) \rfloor) = L^n(\lfloor nt_0 \rfloor),$$

since we may only accumulate local time at 0. For $i \geq 0$, let

$$M_X^n(i) = X^n(i) - \sum_{j=0}^{i-1} \mathbb{E}\Big[X^n(j+1) - X^n(j) | \mathcal{F}_j^n\Big],$$

so that $(M_X^n(i))_{i \geq 0}$ is a martingale. We have

$$\mathbb{E}\Big[M_X^n(i)^2\Big] = \mathbb{E}\left[\sum_{j=0}^{i-1} \mathrm{var}\Big(X^n(j+1) - X^n(j) | \mathcal{F}_j^n\Big)\right] \leq i\left(1 + \mathbb{E}\left[\sup_{0 \leq j \leq i} |E_{X,2}^n(j)|\right]\right),$$

and so

$$\frac{\mathbb{E}[M_X^n(\lfloor nt \rfloor)^2]}{\alpha(n)^2} \leq \frac{\lfloor nt \rfloor \Big(1 + \mathbb{E}\Big[\sup_{0 \leq i \leq \lfloor nt \rfloor} |E_{X,2}^n(i)|\Big]\Big)}{\alpha(n)^2} \to 0,$$

by Lemma 65 and the fact that $\alpha(n) >> \sqrt{n}$. It follows as usual that

$$\frac{\sup_{0 \leq s \leq t \wedge T_n} |M_X^n(\lfloor ns \rfloor)|}{\alpha(n)} \xrightarrow{p} 0. \tag{3.15}$$

Then

$$\sup_{0 \leq s \leq t \wedge T^n} \left| \sum_{j=0}^{\lfloor ns \rfloor - 1} \frac{\mathbb{E}\Big[X^n(j+1) - X^n(j) | \mathcal{F}_j^n\Big]}{\alpha(n)} - \int_0^s \left(\frac{2}{3}(1-u)^{1/3} - \frac{2x(s)}{3(1-s)}\right) du \right|$$

$$
\begin{aligned}
&\le \sup_{0\le s\le t}\left|\frac{1}{n}\sum_{j=0}^{\lfloor ns\rfloor-1}\frac{2}{3}(1-j/n)^{1/3}-\int_0^s\frac{2}{3}(1-u)^{1/3}du\right| \\
&\quad+\sup_{0\le s\le t\wedge T^n}\left|\frac{1}{n}\sum_{j=0}^{\lfloor ns\rfloor-1}\frac{2}{3(1-j/n)}\frac{X^n(j)}{\alpha(n)}-\int_0^s\frac{2x(u)}{3(1-u)}du\right|+\frac{\lfloor nt\rfloor}{\alpha(n)}\sup_{0\le i\le\lfloor n(t\wedge T^n)\rfloor}|E^n_{X,1}(i)| \\
&\quad+\sup_{0\le s\le t\wedge T^n}\sum_{j=0}^{\lfloor ns\rfloor-1}\frac{2\mathbb{1}_{\{X^n(i)=0\}}}{\alpha(n)}.
\end{aligned}
$$

The first term is clearly the difference between a Riemann approximation to an integral and that integral, and tends to 0. The third term converges in probability to 0 by Lemma 65. The forth term converges in probability to 0 by Proposition 67. The second term is bounded above by

$$
\begin{aligned}
&\frac{2}{3(1-t)}\int_0^{\lfloor n(t\wedge T^n)\rfloor/n}\left|\frac{X^n(\lfloor nu\rfloor)}{\alpha(n)}-x(u)\right|du+\int_0^t\left|\frac{2}{3(1-u)}-\frac{2}{3(1-\lfloor nu\rfloor/u)}\right|x(u)du \\
&\quad+\sup_{0\le s\le t}\int_{\lfloor ns\rfloor/n}^s\frac{2x(u)}{3(1-u)}du,
\end{aligned}
$$

where the second and third terms clearly tend to 0 as $n\to\infty$. The rest of the argument then goes through as in the proof of Proposition 61 to show that

$$
\sup_{0\le s\le t\wedge T^n}\left|\frac{X^n(\lfloor ns\rfloor)}{\alpha(n)}-x(s)\right|\xrightarrow{p}0.
$$

Now observe that for fixed $t>t_0$, $\inf_{t_0\le s\le t}x(s)>0$. So for sufficiently small $\delta>0$,

$$
\sup_{0\le s\le t\wedge T^n}\left|\frac{X^n(\lfloor ns\rfloor)}{\alpha(n)}-x(s)\right|\le\delta
$$

implies that $T^n>t$ and $L^n(\lfloor nt\rfloor)=L^n(\lfloor nt_0\rfloor)$. Hence,

$$
\sup_{0\le s\le t}\left|\frac{X^n(\lfloor ns\rfloor)}{\alpha(n)}-x(s)\right|\xrightarrow{p}0\quad\text{and}\quad\frac{L^n(\lfloor nt\rfloor)}{\alpha(n)}\xrightarrow{p}0.
$$

Finally, to get the result for N^n, note that again

$$
M^n_N(i)=N^n(i)-\sum_{j=0}^{i-1}\mathbb{E}\Big[N^n(j+1)-N^n(j)|\mathcal{F}^n_j\Big],
$$

defines a martingale, with

$$\mathbb{E}\Big[M_N^n(i)^2\Big] \le \mathbb{E}\left[\sum_{j=0}^{i-1} \mathbb{E}\Big[N^n(j+1) - N^n(j)|\mathcal{F}_j^n\Big]\right]$$

and so

$$\begin{aligned}\frac{\mathbb{E}[M_N^n(\lfloor nt\rfloor)^2]}{\alpha(n)^2} &\le \frac{1}{n\alpha(n)}\mathbb{E}\left[\sum_{i=0}^{\lfloor nt\rfloor-1}\frac{X^n(i)/\alpha(n)}{3(1-i/n)}\right] + \frac{nt}{\alpha(n)^2}\mathbb{E}\left[\sup_{0\le i\le\lfloor nt\rfloor} E_N^n(i)\right]\\ &\le \frac{t}{3(1-t)\alpha(n)^2}\mathbb{E}\left[\sup_{0\le i\le\lfloor nt\rfloor} X^n(i)\right] + \frac{nt}{\alpha(n)^2}\mathbb{E}\left[\sup_{0\le i\le\lfloor nt\rfloor} E_N^n(i)\right] \to 0,\end{aligned}$$

by Lemmas 60 and 65. Then

$$\begin{aligned}\sup_{0\le s\le t}\left|\frac{N^n(\lfloor ns\rfloor)}{\alpha(n)} - m(s)\right| \le \frac{\sup_{0\le s\le t}|M_N^n(\lfloor ns\rfloor)|}{\alpha(n)} + \frac{\sup_{0\le i\le\lfloor nt\rfloor}|E_N^n(i)|}{\alpha(n)}\\ + \sup_{0\le s\le t}\left|\frac{1}{n}\sum_{j=0}^{\lfloor ns\rfloor-1}\frac{1}{3(1-j/n)}\frac{X^n(j)}{\alpha(n)} - \int_0^s\frac{x(s)}{3(1-s)}ds\right|,\end{aligned}$$

and the usual argument allows us to complete the proof. □

3.5 Diffusion limit for $\alpha(n) = O(\sqrt{n})$

3.5.1 The limiting process

Proposition 68. *Let $a \ge 0$ and let $(B_t, 0 \le t \le 1)$ be a Brownian motion. There is a unique solution (X^a, L^a) to the SDE with reflection at 0 given by*

$$dX_t^a = -\frac{2}{3}\frac{X_t^a}{(1-t)}dt + \frac{2a}{3}(1-t)^{1/3}dt + dB_t + dL_t^a, \tag{3.16}$$

for $0 \le t < 1$ with initial condition $X_0^a = 0$, $L_0^a = 0$, satisfying the conditions (a) $X_t^a \ge 0$ for $0 \le t < 1$, (b) L_t^a is non-decreasing, and (c) $\int_0^t \mathbb{1}_{\{X_s^a > 0\}}dL_s^a = 0$ for $0 \le t < 1$.

Indeed, let

$$K_t^a = -\inf_{0\le s\le t}\left[a - a(1-s)^{2/3} + \int_0^s \frac{dB_u}{(1-u)^{2/3}}\right].$$

Then the solution to (3.16) is given for $0 \le t < 1$ *by*

$$
\begin{aligned}
L_t^a &= \int_0^t (1-s)^{2/3} dK_s^a, \\
X_t^a &= a(1-t)^{2/3} - a(1-t)^{4/3} + (1-t)^{2/3} \int_0^t \frac{dB_u}{(1-u)^{2/3}} + (1-t)^{2/3} K_t^a.
\end{aligned}
$$

Moreover, the following statements are true almost surely:

$$
X_1^a := \lim_{t \to 1-} X_t^a = 0, \qquad \int_0^1 \frac{X_s^a}{3(1-s)} ds < \infty, \qquad L_1^a := \lim_{t \to 1-} L_t^a < \infty.
$$

Finally, for $a \ge 0$ *and* $r \ge a+1$ *we have that*

$$
\sup_{0 \le t < 1} \mathbb{P}\left(\frac{X_t^a}{\sqrt{1-t}} > r \right) \le e^{-(r-a)^2/6}.
$$

Proof. Theorem 2.1.1 from Pilipenko [44] guarantees the existence of a unique solution to the reflected SDE on the time-interval $[0,t]$ for any fixed $t<1$, since the diffusion and drift coefficients satisfy the appropriate Lipschitz and linear growth conditions.

For $0 \le t < 1$, define

$$
\begin{aligned}
Y_t^a &= a - a(1-t)^{2/3} + \int_0^t \frac{dB_u}{(1-u)^{2/3}} - \inf_{0 \le s \le t} \left[a - a(1-s)^{2/3} + \int_0^s \frac{dB_u}{(1-u)^{2/3}} \right] \\
&= a - a(1-t)^{2/3} + \int_0^t \frac{dB_u}{(1-u)^{2/3}} + K_t^a.
\end{aligned}
$$

Then we straightforwardly have that

$$
dY_t^a = \frac{2a}{3}(1-t)^{-1/3} dt + (1-t)^{-2/3} dB_t + dK_t^a,
$$

and, by Skorokhod's lemma (see, for example, Lemma 2.1 of Chapter VI of Revuz and Yor [48]), K^a is the local time at 0 of Y^a.

Now let $X_t^a = (1-t)^{2/3} Y_t^a$ and $L_t^a = \int_0^t (1-s)^{2/3} dK_s^a$. Then, by construction, $Y_t^a \ge 0$ for $0 \le t < 1$ and so $X_t^a \ge 0$ for $0 \le t < 1$. Furthermore,

$$
\begin{aligned}
dX_t^a &= -\frac{2}{3}(1-t)^{-1/3} Y_t^a dt + (1-t)^{2/3} dY_t^a \\
&= -\frac{2}{3} \frac{X_t^a}{(1-t)} dt + \frac{2a}{3}(1-t)^{1/3} dt + dB_t + (1-t)^{2/3} dK_t^a
\end{aligned}
$$

$$= -\frac{2}{3}\frac{X_t^a}{(1-t)}dt + \frac{2a}{3}(1-t)^{1/3}dt + dB_t + dL_t^a.$$

Clearly $L_0^a = 0$ and L_t^a is non-decreasing. Let us show that the local time at 0 of X^a is L^a. By Tanaka's formula, the local time is given by

$$\begin{aligned}
&|X_t^a| - \int_0^t \operatorname{sgn}(X_s^a)dX_s^a \\
&= (1-t)^{2/3}|Y_t^a| + \frac{2}{3}\int_0^t \operatorname{sgn}(Y_s^a)Y_s^a(1-s)^{-1/3}ds - \int_0^t (1-s)^{2/3}\operatorname{sgn}(Y_s^a)dY_s^a \\
&= (1-t)^{2/3}|Y_t^a| + \frac{2}{3}\int_0^t |Y_s^a|(1-s)^{-1/3}ds - \int_0^t (1-s)^{2/3}\operatorname{sgn}(Y_s^a)dY_s^a \\
&= \int_0^t (1-s)^{2/3}d|Y_s^a| - \int_0^t (1-s)^{2/3}\operatorname{sgn}(Y_s^a)dY_s^a,
\end{aligned}$$

by integration by parts. But we also have $K_t^a = |Y_t^a| - \int_0^t \operatorname{sgn}(Y_s^a)dY_s^a$ and so the last line is equal to $\int_0^t (1-s)^{2/3}dK_s^a$, which is L_t^a. It follows that $\int_0^t \mathbb{1}_{\{X_s^a>0\}}dL_s^a = 0$. Hence, (X^a, L^a) solves the reflected SDE.

For $a = 0$, we have

$$X_t^0 = (1-t)^{2/3}\int_0^t \frac{dB_u}{(1-u)^{2/3}} - (1-t)^{2/3}\inf_{0\le s\le t}\int_0^s \frac{dB_u}{(1-u)^{2/3}}$$

and so

$$(X_t^0, 0 \le t < 1) \stackrel{d}{=} \left((1-t)^{2/3}\left|\int_0^t \frac{dB_u}{(1-u)^{2/3}}\right|, 0 \le t < 1\right).$$

(See, for example, Exercise 1.3.1 of [44].)

Let us first show that $\lim_{t\to 1} X_t^0 = 0$. By the Dubins-Schwarz theorem we have

$$\left(\int_0^t \frac{dB_s}{(1-s)^{2/3}}, 0 \le t < 1\right) \stackrel{d}{=} \left(B_{f(t)}, 0 \le t < 1\right), \tag{3.17}$$

where $f(t) = 3(1-t)^{-1/3} - 3$ is the quadratic variation of the process on the left-hand side of (3.17). Recall that

$$(B_t, t \ge 0) \stackrel{d}{=} (tB_{1/t}, t \ge 0).$$

Thus we have

$$\Big(X_t^0, 0 \le t < 1\Big) \stackrel{d}{=} \Big((1-t)^{2/3} f(t) \cdot |B_{1/f(t)}|, 0 \le t < 1\Big).$$

But then

$$(1-t)^{2/3} f(t) \cdot |B_{1/f(t)}| = 3\Big((1-t)^{1/3} - (1-t)^{2/3}\Big) \Bigg| B_{\frac{(1-t)^{1/3}}{3(1-(1-t)^{1/3})}} \Bigg| \to 0$$

almost surely as $t \to 1$. It follows that $\lim_{t\to 1} X_t^0 = 0$ almost surely. Moreover,

$$\int_0^1 \frac{X_t^0}{3(1-t)} dt \stackrel{d}{=} \frac{1}{3} \int_0^1 (1-t)^{-1/3} |B_{f(t)}| dt.$$

The change of variables $u = f(t)$ yields that the right-hand side is equal to

$$9 \int_0^1 \frac{|B_u|}{(u+3)^3} du.$$

Since $B_u =_d \sqrt{u} B_1$ and $\mathbb{E}[|B_u|] = \sqrt{\frac{2u}{\pi}}$ we have

$$\mathbb{E}\left[\int_0^1 \frac{X_t^0}{3(1-t)} dt\right] = 9\sqrt{\frac{2}{\pi}} \int_0^1 \frac{\sqrt{u}}{(u+3)^3} du < \infty,$$

and so

$$\int_0^1 \frac{X_t^0}{3(1-t)} dt < \infty \text{ a.s.}$$

Since we have

$$X_t^0 = -2 \int_0^t \frac{X_s^0}{3(1-s)} ds + B_t + L_t^0,$$

and X_t^0, $\int_0^t \frac{X_s^0}{3(1-s)} dt$ and B_t all possess finite almost sure limits as $t \to 1$, the same must also be true of L_t^0.

To obtain the almost sure finiteness statements for a general $a \ge 0$, note that if we build X^0 and X^a from the same Brownian motion then

$$0 \le X_t^a \le a(1-t)^{2/3} + X_t^0. \tag{3.18}$$

It follows that $\lim_{t\to 1} X_t^a = 0$ almost surely. We also get

$$\int_0^1 \frac{X_s^a}{3(1-s)} ds \le \frac{a}{2} + \int_0^1 \frac{X_s^0}{3(1-s)} ds < \infty$$

almost surely. Finally, by the same argument as in the $a = 0$ case, we must then also have that $L_1^a := \lim_{t\to 1} L_t^a < \infty$ almost surely.

We now turn to the final statement. By (3.18), it is sufficient to show the result for $a = 0$. For any $0 \le t < 1$, we have

$$\begin{aligned}\mathbb{P}\left(\frac{X_t^0}{\sqrt{1-t}} > r\right) &= \mathbb{P}\Big((1-t)^{1/6}|B_{f(t)}| > r\Big) \\ &= \mathbb{P}\Big(|\mathrm{N}(0, 3 - 3(1-t)^{1/3})| > r\Big) \\ &\le \mathbb{P}\Big(|\mathrm{N}(0,1)| > r/\sqrt{3}\Big) \le e^{-r^2/6},\end{aligned}$$

by standard Gaussian tail bounds. □

3.5.2 The invariance principle

Theorem 64 gives that for $\alpha(n) >> \sqrt{n}$ the number of fires rescaled by $\alpha(n)$ possesses a fluid limit. This arises because the dominant contribution to the number of fires comes from connecting to nodes of degree 4. If we had no vertices of degree 4, we would initially connect only to vertices of degree 3, and the number of fires would make jumps of $+1$ or -1 only, each with probability 1/2. This suggests that X^n should behave like a simple symmetric random walk. If $\alpha(n) \sim a\sqrt{n}$, $a > 0$ this random walk acquires a positive drift.

Recall that for $0 \le i \le \zeta^n$, we have

$$L^n(i) = 2\sum_{j=0}^{i-1} \mathbb{1}_{\{X^n(j)=0\}}.$$

We will prove the following scaling limit.

Theorem 69. *Fix $t < 1$ and suppose $\alpha(n)/\sqrt{n} \to a$ for $a \ge 0$. Then*

$$\left(\frac{X^n(\lfloor ns\rfloor)}{\sqrt{n}}, \frac{L^n(\lfloor ns\rfloor)}{\sqrt{n}}, 0 \le s \le t\right) \xrightarrow{d} (X_s^a, L_s^a, 0 \le s \le t),$$

uniformly, where $(X^a_s, L^a_s, 0 \le s \le 1)$ *is the process from Proposition 68.*

We will deduce Theorem 69 from the following general invariance principle for reflected diffusions. Such invariance principles go back to Stroock and Varadhan [50], and we do not believe that this result is novel. But since we could not find a version in the existing literature adapted to our particular setting, we give a proof in the appendix, using ideas from Ethier and Kurtz [21] and Kang and Williams [33].

Theorem 70. *Let* $q : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$, $b : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ *satisfy the following conditions:*

(1) (Lipschitz condition)

$$\exists K > 0 \quad \forall t \ge 0 \quad \forall y_1, y_2 \in \mathbb{R}_+ : |q(t,y_1) - q(t,y_2)| + |b(t,y_1) - b(t,y_2)| \le K|y_1 - y_2|;$$

(2) (linear growth condition)

$$\exists C > 0 \quad \forall t \ge 0 \quad \forall x \in \mathbb{R}_+ : |q(t,y)| + |b(t,y)| \le C(1 + |y|).$$

Let Y^n, B^n *be* $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$*-valued stochastic processes and let* Q^n, L^n *be increasing* $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$*-valued stochastic processes. Let* $\mathcal{F}^n_t = \sigma(Y^n(s), L^n(s), Q^n(s), B^n(s) : s \le t)$ *and let* $\tau^n(r) = \inf\{t \ge 0 : |Y^n(t)| \ge r \text{ or } |Y^n(t-)| \ge r\}$*. Suppose that*

(a) $M^n = Y^n - B^n$ *is an* $(\mathcal{F}^n_t)$*-local martingale,*

(b) $(M^n)^2 - Q^n$ *is an* $(\mathcal{F}^n_t)$*-local martingale,*

and that for all $r > 0$ *and* $T > 0$,

(c) $\lim_{n\to\infty} \mathbb{E}\Big[\sup_{t \le T\wedge\tau^n(r)} |Y^n(t) - Y^n(t-)|^2\Big] = 0$,

(d) $\lim_{n\to\infty} \mathbb{E}\Big[\sup_{t \le T\wedge\tau^n(r)} |B^n(t) - B^n(t-)|^2\Big] = 0$,

(e) $\lim_{n\to\infty} \mathbb{E}\Big[\sup_{t \le T\wedge\tau^n(r)} |Q^n(t) - Q^n(t-)|\Big] = 0$,

(f) $\sup_{t \le T\wedge\tau^n(r)} \Big| B^n(t) - L^n(t) - \int_0^t b(s, Y^n(s))ds \Big| \xrightarrow{p} 0$ *as* $n \to \infty$,

(g) $\sup_{t \le T\wedge\tau^n(r)} \Big| Q^n(t) - \int_0^t q(s, Y^n(s))ds \Big| \xrightarrow{p} 0$ *as* $n \to \infty$,

(h) $Y^n(0) \xrightarrow{p} 0$ *as* $n \to \infty$,

(i) there exists a sequence $(\delta_n)_{n\geq 1}$ *of constants with* $\delta_n \to 0$ *such that* $L^n(0) = 0$, $L^n(t) - L^n(t-) \leq \delta_n$ *and* $L^n(t) = \int_0^t \mathbb{1}_{\{Y^n(s)\leq \delta_n\}} dL^n(s)$.

Then $(Y^n, L^n) \xrightarrow{d} (Y, L)$ *in the Skorokhod topology as* $n \to \infty$*, where* (Y, L) *is the unique solution to the reflected SDE specified by* $Y_0 = L_0 = 0$*,* L *is non-decreasing, and*

$$dY_t = b(t, Y_t)dt + \sigma(t, Y_t)dW_t + dL_t,$$

where $\int_0^t \mathbb{1}_{\{Y_s>0\}} dL_s = 0$*,* W *is a standard Brownian motion and* $\sigma(t, y) = \sqrt{q(t, y)}$.

Proof of Theorem 69. Define

$$\tilde{X}^n(s) = \frac{X^n(\lfloor ns \rfloor)}{\sqrt{n}} \quad \text{and} \quad \tilde{L}^n(s) = \frac{L^n(\lfloor ns \rfloor)}{\sqrt{n}}.$$

Recall from Proposition 68 that $(X^a_s, L^a_s, 0 \leq s \leq t)$ is the solution to the reflected SDE

$$dX^a_s = -\frac{2}{3}\frac{X^a_s}{(1-s)}ds + \frac{2a}{3}(1-s)^{1/3}ds + dB_s + dL^a_s.$$

For $0 \leq s \leq t$ and $x \geq 0$, let $q(s, x) = 1$ and $b(s, x) = -\frac{2}{3}\frac{x}{(1-s)} + \frac{2a}{3}(1-s)^{1/3}$. It follows straightforwardly that for $0 \leq s \leq t$, the functions q and b satisfy the Lipschitz and linear growth conditions (1) and (2) in Theorem 70.

For $i \leq \lfloor nt \rfloor$, let

$$B^n(i) = \sum_{j=0}^{i-1} \mathbb{E}\Big[X^n(j+1) - X^n(j) | \mathcal{F}^n_j\Big],$$

and

$$M^n_X(i) := X^n(i) - B^n(i),$$

as well as

$$\begin{aligned} Q^n(i) &= \sum_{j=0}^{i-1} \mathbb{E}\Big[M^n_X(j+1)^2 - M^n_X(j)^2 | \mathcal{F}^n_j\Big] = \sum_{j=0}^{i-1} \mathbb{E}\Big[(M^n_X(j+1) - M^n_X(j))^2 | \mathcal{F}^n_j\Big] \\ &= \sum_{j=0}^{i-1} \mathbb{E}\Big[(X^n(j+1) - X^n(j))^2 | \mathcal{F}^n_i\Big] - \mathbb{E}\Big[X^n(j+1) - X^n(j) | \mathcal{F}^n_j\Big]^2. \end{aligned}$$

Since everything is bounded, M_X^n and $(M_X^n)^2 - Q^n$ are both $(\mathcal{F}_i^n)$-martingales. Let

$$\tilde{M}(s) = \frac{M_X^n(\lfloor ns \rfloor)}{\sqrt{n}}, \quad \tilde{B}^n(s) = \frac{B^n(\lfloor ns \rfloor)}{\sqrt{n}} \quad \text{and} \quad \tilde{Q}^n(s) = \frac{Q^n(\lfloor ns \rfloor)}{n}$$

be the appropriately rescaled versions of these quantities. Then (a) and (b) hold by construction and the main task facing us is to check that conditions (f) and (g) of the Theorem are fulfilled for $\tilde{M}^n$ and $(\tilde{M}^n)^2 - \tilde{Q}^n$.

By Lemma 65,

$$\begin{aligned} B^n(i) &= 2\sum_{j=0}^{i-1} \mathbb{1}_{\{X^n(j)=0\}} + \sum_{j=0}^{i-1} \frac{2\alpha(n)(1-j/n)^{4/3} - 2X^n(j)}{3(n-j)} + \sum_{j=0}^{i-1} E_{X,1}^n(j) \\ &= L^n(i) + \sum_{j=0}^{i-1} \frac{2\alpha(n)(1-j/n)^{4/3} - 2X^n(j)}{3(n-j)} + \sum_{j=0}^{i-1} E_{X,1}^n(j). \end{aligned}$$

Then

$$\begin{aligned} &\sup_{0\le s\le t} \left| \tilde{B}^n(s) - \tilde{L}^n(s) - \int_0^s \left(\frac{2a}{3}(1-u)^{1/3} - \frac{2\tilde{X}^n(u)}{(1-u)} \right) du \right| \\ &\le \sup_{0\le s\le t} \left| \sum_{j=0}^{\lfloor ns \rfloor -1} \frac{1}{n} \left(\frac{2\alpha(n)}{3\sqrt{n}}(1-j/n)^{1/3} - \frac{2\tilde{X}^n(j/n)}{3(1-j/n)} \right) - \int_0^s \left(\frac{2a}{3}(1-u)^{1/3} - \frac{2\tilde{X}^n(u)}{3(1-u)} \right) du \right| \\ &\quad + \sup_{0\le i\le \lfloor nt \rfloor} \sqrt{n} |E_{X,1}^n(i)|, \end{aligned}$$

which converges to 0 in probability as $n \to \infty$ by Lemma 65. Hence, condition (f) holds.

By Lemma 65,

$$\begin{aligned} Q^n(i) = \sum_{j=0}^{i-1} \Bigg(1 &+ E_{X,2}^n(j) - 4\mathbb{1}_{\{X^n(j)=0\}} \left(\frac{2\alpha(n)(1-j/n)^{4/3} - 2X^n(j)}{3(n-j)} + E_{X,1}^n(j) \right) \\ &- \left(\frac{2\alpha(n)(1-j/n)^{4/3} - 2X^n(j)}{3(n-j)} + E_{X,1}^n(j) \right)^2 \Bigg) \\ = i + \sum_{j=0}^{i-1} \Bigg(&E_{X,2}^n(j) - 4\mathbb{1}_{\{X^n(j)=0\}} \left(\frac{2\alpha(n)(1-j/n)^{4/3} - 2X^n(j)}{3(n-j)} + E_{X,1}^n(j) \right) \\ &- \left(\frac{2\alpha(n)(1-j/n)^{4/3} - 2X^n(j)}{3(n-j)} + E_{X,1}^n(j) \right)^2 \Bigg). \end{aligned}$$

Then

$$\begin{aligned}
&\sup_{0\le s\le t}\left|\tilde{Q}^n(s) - s\right| \\
&= \sup_{0\le s\le t}\left|\frac{\lfloor ns\rfloor}{n} - s + \sum_{j=0}^{\lfloor ns\rfloor-1}\left(\frac{E^n_{X,2}(j)}{n} - \left(\frac{2(\alpha(n)/\sqrt{n})(1-j/n)^{4/3} - 2\tilde{X}(j/n)}{3(n-j)} + \frac{E^n_{X,1}(j)}{\sqrt{n}}\right)^2\right.\right. \\
&\qquad\left.\left. - \frac{4}{\sqrt{n}}\mathbb{1}_{\{\tilde{X}^n(j/n)=0\}}\left(\frac{2(\alpha(n)/\sqrt{n})(1-j/n)^{4/3} - 2\tilde{X}(j/n)}{3(n-j)} + \frac{E^n_{X,1}(j)}{\sqrt{n}}\right)\right)\right|.
\end{aligned}$$

Now, for large enough n we have

$$\begin{aligned}
&\left|\frac{2(\alpha(n)/\sqrt{n})(1-j/n)^{4/3} - 2\tilde{X}(j/n)}{3(n-j)} + \frac{E^n_{X,1}(j)}{\sqrt{n}}\right| \\
&\qquad\le \frac{1}{n}\frac{4a}{3(1-t)^{1/3}} + \frac{1}{n}\frac{2}{3(1-t)}\sup_{0\le s\le t}\tilde{X}^n(s) + \frac{\sup_{0\le i\le\lfloor nt\rfloor}|E^n_{X,1}(i)|}{\sqrt{n}}.
\end{aligned}$$

Since $\sup_{0\le i\le\lfloor nt\rfloor}|E^n_{X,2}(i)| \xrightarrow{p} 0$, $\sup_{0\le i\le\lfloor nt\rfloor}|E^n_{X,1}(i)| \xrightarrow{p} 0$ and $\sup_{0\le s\le t}\tilde{X}^n(s)$ is tight, it follows that

$$\sup_{0\le s\le t}\left|\tilde{Q}^n(s) - s\right| \xrightarrow{p} 0$$

as $n\to\infty$, and so we have that condition (g) of Theorem 70 is fulfilled.

Since the increments of X^n are bounded, conditions (c), (d) and (e) of Theorem 70 are trivially satisfied. As $\tilde{X}^n(0) = 0$, condition (h) is satisfied. Finally, note that condition (i) is fulfilled if we take $\delta_n = \frac{2}{\sqrt{n}}$, since $\tilde{L}^n(s) - \tilde{L}^n(s-) \le \frac{2}{\sqrt{n}}$ and

$$\int_0^s \mathbb{1}_{\{\tilde{X}^n(u)\le\frac{2}{\sqrt{n}}\}}dL^n(u) = 2\sum_{i=0}^{\lfloor ns\rfloor-1}\mathbb{1}_{\{\tilde{X}^n(i)=0\}} = L^n(s).$$

Applying Theorem 70 completes the proof. □

Finally, we are ready to prove the convergence of the suitably rescaled number of clashes we encounter up to a fixed time $t < 1$.

Lemma 71. *For fixed* $0 < t < 1$ *we have*

$$\left(\frac{N^n(\lfloor ns\rfloor)}{\sqrt{n}}, 0\le s\le t\right) \xrightarrow{d} (N^a_s, 0\le s\le t)$$

uniformly as $n\to\infty$.

Proof. The argument is standard, since N^n is a counting process which is bounded by $(3n+4\alpha(n))/2$. We have that

$$M_N^n(k) = N^n(k) - \sum_{i=1}^{k-1} \mathbb{E}[N^n(i+1) - N^n(i)|\mathcal{F}_i^n]$$

defines a martingale. We also have

$$\begin{aligned}\mathbb{E}\Big[M_N^n(k)^2\Big] &= \mathbb{E}\left[\sum_{i=0}^{k-1} \operatorname{var}(N^n(i+1) - N^n(i)|\mathcal{F}_i^n)\right] \\ &\leq \mathbb{E}\left[\sum_{i=0}^{k-1} \mathbb{E}[N^n(i+1) - N^n(i)|\mathcal{F}_i^n]\right] = \mathbb{E}[N^n(k)],\end{aligned}$$

where the inequality holds since $N^n(i+1) - N^n(i) \in \{0,1\}$. Now,

$$\begin{aligned}\frac{1}{\sqrt{n}}\mathbb{E}[N^n(\lfloor nt\rfloor)] &\leq \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt\rfloor -1} \mathbb{E}\left[\frac{3n + 2X^n(i)(3n+4\alpha(n)+2X^n(i))}{2(3n(1-t)-1)(3n(1-t)-3)}\right] \\ &\leq \frac{\sqrt{n}t\Big(3n + 2\mathbb{E}\Big[\sup_{0\leq i\leq \zeta_n} X^n(i)\Big](3n+4\alpha(n)+2n+2\alpha(n))\Big)}{2(3n(1-t)-1)(3n(1-t)-3)} \leq C\end{aligned} \tag{3.19}$$

for some constant C, uniformly in n. In particular, $\mathbb{E}[N^n(\lfloor nt\rfloor)]/n \to 0$. Fix $\delta > 0$. Using Markov's inequality, we obtain that

$$\frac{M_N^n(\lfloor nt\rfloor)}{\sqrt{n}} \xrightarrow{p} 0,$$

as $n \to \infty$.

Now, by Lemma 65 we have

$$\sqrt{n} \sup_{0\leq i\leq \lfloor nt\rfloor} \left| \mathbb{E}[N^n(i+1) - N^n(i)|\mathcal{F}_i^n] - \frac{X^n(i)}{3(n-i)}\right| \xrightarrow{p} 0$$

and so

$$\sup_{0\leq s\leq t} \frac{1}{\sqrt{n}} \left| \sum_{i=1}^{\lfloor ns\rfloor -1} \mathbb{E}[N^n(i+1) - N^n(i)|\mathcal{F}_i^n] - \sum_{i=0}^{\lfloor ns\rfloor -1} \frac{X^n(i)}{3(n-i)}\right| \xrightarrow{p} 0.$$

Finally, we have

$$\frac{1}{\sqrt{n}} \sum_{i=0}^{\lfloor ns\rfloor -1} \frac{X^n(i)}{3(n-i)} = \frac{1}{\sqrt{n}} \int_0^{\lfloor ns\rfloor/n} \frac{X^n(\lfloor nu\rfloor)}{3\Big(1 - \frac{\lfloor nu\rfloor}{n}\Big)} du.$$

But then using Theorem 69 and the continuous mapping theorem, we get

$$\frac{1}{\sqrt{n}}\int_0^{\lfloor ns\rfloor/n}\frac{X^n(\lfloor nu\rfloor)}{3\left(1-\frac{\lfloor nu\rfloor}{n}\right)}du \xrightarrow{d} \int_0^s \frac{X_u^a}{3(1-u)}du = N_s^a,$$

uniformly for $0 \le s \le t$. Hence, as $n \to \infty$

$$\frac{N^n(\lfloor ns\rfloor)}{\sqrt{n}} \xrightarrow{d} N_s^a$$

uniformly for $0 \le s \le t$, as desired. □

3.6 The end of the process

We now understand the scaling behaviour of the process $(X^n(i), L^n(i), N^n(i))_{i\ge 0}$ on the time-interval $[0, \lfloor n(1-\epsilon)\rfloor]$ for any $\epsilon > 0$. It remains to get from $\lfloor n(1-\epsilon)\rfloor$ to ζ_n.

We first show that $V^n(n)$ is small and that $X^n(\lfloor n(1-\epsilon)\rfloor), L^n(\lfloor n(1-\epsilon)\rfloor), N^n(\lfloor n(1-\epsilon)\rfloor)$ are, for small ϵ, good approximations of $X^n(n), L^n(n), N^n(n)$, by careful use of the coupling from Section 3.2.2.

Proposition 72. *For any $\delta > 0$,*

(a) if $\alpha(n) \to \infty$ then $V^n(n)/\alpha(n) \xrightarrow{p} 0$ as $n \to \infty$,

(b) $\lim_{\epsilon\to 0}\limsup_{n\to\infty}\mathbb{P}\left(\frac{1}{\alpha(n)\vee\sqrt{n}}(N^n(n) - N^n(\lfloor n(1-\epsilon)\rfloor)) > \delta\right) = 0,$

(c) $\lim_{\epsilon\to 0}\limsup_{n\to\infty}\mathbb{P}\left(\frac{1}{\alpha(n)\vee\sqrt{n}}\sup_{\lfloor n(1-\epsilon)\rfloor\le i\le n} X^n(i) > \delta\right) = 0,$

(d) $\lim_{\epsilon\to 0}\limsup_{n\to\infty}\mathbb{P}\left(\frac{1}{\alpha(n)\vee\sqrt{n}}(L^n(n) - L^n(\lfloor n(1-\epsilon)\rfloor)) > \delta\right) = 0.$

Proof. (a) We have

$$\mathbb{E}[V^n(\lfloor n(1-\epsilon)\rfloor)] \le 2\epsilon^{4/3}\alpha(n) + \alpha(n)\mathbb{P}\left(V^n(\lfloor(1-\epsilon)n\rfloor) > 2\epsilon^{4/3}\alpha(n)\right)$$

and, by Lemma 63, $\mathbb{P}\left(V^n(\lfloor(1-\epsilon)n\rfloor) > 2\epsilon^{4/3}\alpha(n)\right) \to 0$ as $n \to \infty$. But V^n is decreasing and so

$$\limsup_{n\to\infty}\frac{\mathbb{E}[V^n(n)]}{\alpha(n)} \le \limsup_{n\to\infty}\frac{\mathbb{E}[V^n(\lfloor n(1-\epsilon)\rfloor)]}{\alpha(n)} \le 2\epsilon^{4/3}. \tag{3.20}$$

Since $V^n(n) \geq 0$ and this inequality holds for any $\epsilon > 0$, we must have that the left-hand side is, in fact, equal to 0. Markov's inequality then gives that $V^n(n)/\alpha(n) \xrightarrow{p} 0$.

(b) Recall the coupling and notation from Section 3.2.2. First note that

$$\begin{aligned}
&\mathbb{E}[N^n(n) - N^n(\lfloor n(1-\epsilon)\rfloor)] \\
&\qquad = \mathbb{E}[N_1^n(n) - N_1^n(\lfloor n(1-\epsilon)\rfloor)] + \mathbb{E}[N_2^n(n) - N_2^n(\lfloor n(1-\epsilon)\rfloor)] \\
&\qquad \leq \mathbb{E}[N_1^n(n) - N_1^n(\lfloor n(1-\epsilon)\rfloor)] + \mathbb{E}[V^n(\lfloor n(1-\epsilon)\rfloor)].
\end{aligned} \tag{3.21}$$

From (3.7), the process

$$\left(N_1^n(k) - \sum_{i=0}^{k-1} \frac{3U^n(i)+2(X_1^n(i)-\mathbb{1}_{\{X_1^n(i)>0\}})(3U^n(i)+4V^n(i)+X^n(i)-3)}{2(3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}})(3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}-2)}, k \geq 0\right)$$

is a martingale.

Hence,

$$\begin{aligned}
&\mathbb{E}[N_1^n(n) - N_1^n(\lfloor n(1-\epsilon)\rfloor)] \\
&\quad = \mathbb{E}\left[\sum_{i=\lfloor n(1-\epsilon)\rfloor}^{n-1} \frac{3U^n(i)+2(X_1^n(i)-\mathbb{1}_{\{X_1^n(i)>0\}})(3U^n(i)+4V^n(i)+X^n(i)-3)}{2(3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}})(3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}-2)}\right] \\
&\quad \leq \frac{1}{6}\sum_{i=\lfloor n(1-\epsilon)\rfloor}^{n-1} \frac{1}{(n-i)} + \mathbb{E}\left[\sum_{i=\lfloor n(1-\epsilon)\rfloor}^{n-1} \frac{X_1^n(i) - \mathbb{1}_{\{X_1^n(i)>0\}}}{3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}\right] \\
&\quad \leq \frac{1}{6}(\log(\lceil n\epsilon\rceil)+1) + \mathbb{E}\left[\sum_{i=\lfloor n(1-\epsilon)\rfloor}^{n-1} \frac{X_1^n(i) - \mathbb{1}_{\{X_1^n(i)>0\}}}{3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}\right].
\end{aligned} \tag{3.22}$$

We fix $C > \sqrt{6}$ and define a sequence of stopping times. Let

$$T_0 = \inf\{k \geq \lfloor n(1-\epsilon)\rfloor : X_1^n(k) \geq \lfloor C\sqrt{n-k}\rfloor\},$$

and, inductively, for $i \geq 1$, let

$$\begin{aligned}
S_i &= \inf\{k \geq T_{i-1} : X_1^n(k) = 0\} \wedge n \\
T_i &= \inf\{k \geq S_i : X_1^n(k) = \lfloor C\sqrt{n-k}\rfloor\} \wedge n.
\end{aligned}$$

Then since $U^n(i) \geq n-i$,

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i=\lfloor n(1-\epsilon)\rfloor}^{n-1} \frac{X_1^n(i)-\mathbb{1}_{\{X_1^n(i)>0\}}}{3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}\right]\\
&\leq \mathbb{E}\left[\sum_{i=\lfloor n(1-\epsilon)\rfloor}^{T_0-1} \frac{X_1^n(i)-\mathbb{1}_{\{X_1^n(i)>0\}}}{3(n-i)}\right] + \mathbb{E}\left[\sum_{k=1}^{n}\sum_{i=T_{k-1}}^{S_k-1} \frac{X_1^n(i)-\mathbb{1}_{\{X_1^n(i)>0\}}}{3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}\right]\\
&\quad + \mathbb{E}\left[\sum_{k=1}^{n}\sum_{i=S_k}^{T_k-1} \frac{X_1^n(i)-\mathbb{1}_{\{X_1^n(i)>0\}}}{3(n-i)}\right].
\end{aligned}$$

Either $T_0 = \lfloor n(1-\epsilon)\rfloor$ (in which case the first sum is empty) or on the time interval $[\lfloor n(1-\epsilon)\rfloor, T_0-1]$ we have $X_1^n(i)-1 \leq C\sqrt{n-i}$. Likewise, on $[S_k, T_k-1]$, we have $X_1^n(i)-1 \leq C\sqrt{n-i}$. So

$$\mathbb{E}\left[\sum_{i=\lfloor n(1-\epsilon)\rfloor}^{T_0-1} \frac{X_1^n(i)-\mathbb{1}_{\{X_1^n(i)>0\}}}{3(n-i)}\right] + \mathbb{E}\left[\sum_{k=1}^{n}\sum_{i=S_k}^{T_k-1} \frac{X_1^n(i)-\mathbb{1}_{\{X_1^n(i)>0\}}}{3(n-i)}\right] \leq \frac{C}{3}\sum_{i=\lfloor n(1-\epsilon)\rfloor}^{n-1} \frac{1}{\sqrt{n-i}}.$$

Since

$$\sum_{i=1}^{\lceil n\epsilon\rceil} \frac{1}{\sqrt{i}} \leq \int_0^{\lceil n\epsilon\rceil} \frac{1}{\sqrt{x}}dx \leq 2\sqrt{n\epsilon+1},$$

we get

$$\mathbb{E}\left[\sum_{i=\lfloor n(1-\epsilon)\rfloor}^{T_0-1} \frac{X_1^n(i)-\mathbb{1}_{\{X_1^n(i)>0\}}}{3(n-i)}\right] + \mathbb{E}\left[\sum_{k=1}^{n}\sum_{i=S_k}^{T_k-1} \frac{X_1^n(i)-\mathbb{1}_{\{X_1^n(i)>0\}}}{3(n-i)}\right] \leq \frac{2C}{3}\sqrt{n\epsilon+1}. \tag{3.23}$$

It therefore remains to deal with the term

$$\mathbb{E}\left[\sum_{k=1}^{n}\sum_{i=T_{k-1}}^{S_k-1} \frac{X_1^n(i)-\mathbb{1}_{\{X_1^n(i)>0\}}}{3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}\right].$$

Now note from (3.6) that

$$\begin{aligned}
\Bigg(X_1^n(k) + \sum_{i=0}^{k-1} &\frac{2X_1^n(i)-2\mathbb{1}_{\{X_1^n(i)>0\}}+X_2^n(i)-\mathbb{1}_{\{X_1^n(i)=0,X_2^n(i)>0\}}}{3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}\\
&-\sum_{i=0}^{k-1}\mathbb{1}_{\{X_1^n(i)=0,X_2^n(i)>0\}} - 2\sum_{i=0}^{k-1}\mathbb{1}_{\{X_1^n(i)=X_2^n(i)=0\}}, k\geq 0\Bigg)
\end{aligned}$$

is a martingale. Since X_1^n does not touch 0 on the time interval $[T_{k-1}, S_k - 1]$, we have by the optional stopping theorem that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i=T_{k-1}}^{S_k-1} \frac{X_1^n(i) - \mathbb{1}_{\{X_1^n(i)>0\}}}{3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}\right] \\
&\quad = \frac{1}{2}\mathbb{E}[X_1^n(T_{k-1}) - X_1^n(S_k)] - \frac{1}{2}\mathbb{E}\left[\sum_{i=T_{k-1}}^{S_k-1} \frac{X_2^n(i) - \mathbb{1}_{\{X_1^n(i)=0, X_2^n(i)>0\}}}{3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}\right] \\
&\quad \le \frac{1}{2}\mathbb{E}[X_1^n(T_{k-1})] \\
&\quad \le \frac{C\mathbb{E}[\sqrt{n-T_{k-1}}]}{2} \\
&\quad \le \frac{C\sqrt{n-\mathbb{E}[T_{k-1}]}}{2},
\end{aligned}$$

where the last line follows using Jensen's inequality. So we have

$$\mathbb{E}\left[\sum_{k=1}^{n}\sum_{i=T_{k-1}}^{S_k-1} \frac{X_1^n(i) - \mathbb{1}_{\{X_1^n(i)>0\}}}{3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}\right] \le \frac{C}{2}\sum_{k=0}^{n}\sqrt{n-\mathbb{E}[T_k]}. \tag{3.24}$$

We thus need a lower bound on $\mathbb{E}[T_k]$ for any k. Let us assume that at some time $\lfloor n(1-\epsilon)\rfloor \le \ell < n$, we have $X_1^n(\ell) = 0$. In order to find such a lower bound, we use the coupling from Section 3.2.2, which yields a SSRW reflected above 2, Y^n, such that $Y^n(\ell) = 2$ and $X_1^n(i) \le Y^n(i)$ for all $i \ge \ell$. Recall that

$$(Y^n(i))_{i\ge\ell} \overset{d}{=} (2+|Z(i)|)_{i\ge\ell},$$

where Z is a SSRW with $Z(\ell) = 0$. Then if $\sigma = \inf\{i \ge \ell : 2+|Z(i)| \ge \lfloor C\sqrt{n-i}\rfloor\}$, we have that σ is stochastically smaller than T_k conditioned on $S_k = \ell$. Now, $(Z(i)^2 - i)_{i\ge\ell}$ is a martingale, and so

$$\mathbb{E}\Big[Z(\sigma)^2 - \sigma\Big] = -\ell.$$

But

$$\mathbb{E}\Big[Z(\sigma)^2\Big] = \mathbb{E}\Big[(\lfloor C\sqrt{n-\sigma}\rfloor - 2)^2\Big] \ge \mathbb{E}\Big[(C\sqrt{n-\sigma} - 3)^2\Big] \ge C^2(n-\mathbb{E}[\sigma]) - 6\mathbb{E}\Big[\sqrt{n-\sigma}\Big].$$

Since $n-\sigma \in \mathbb{Z}$, we have the very crude bound $\sqrt{n-\sigma} \le n-\sigma$. Hence,

$$\mathbb{E}[\sigma] - \ell = \mathbb{E}\Big[Z(\sigma)^2\Big] \ge (C^2-6)(n-\mathbb{E}[\sigma]).$$

Then

$$\mathbb{E}[\sigma] \ge \frac{(C^2-6)n+\ell}{C^2-5}.$$

We obtain, by the stochastic domination, that for $k \ge 1$,

$$\mathbb{E}[T_k] \ge \frac{(C^2-6)n+\mathbb{E}[S_k]}{C^2-5}.$$

Since $S_k \ge T_{k-1}$, we get

$$\mathbb{E}[T_k] \ge \frac{(C^2-6)n+\mathbb{E}[T_{k-1}]}{C^2-5},$$

and so

$$n-\mathbb{E}[T_k] \le \frac{n-\mathbb{E}[T_{k-1}]}{C^2-5}.$$

By induction,

$$n-\mathbb{E}[T_k] \le \frac{n-\mathbb{E}[T_0]}{(C^2-5)^k} \le \frac{\lceil n\epsilon\rceil}{(C^2-5)^k},$$

since $T_0 \ge \lfloor n(1-\epsilon)\rfloor$. It follows from (3.24) that

$$\begin{aligned}\mathbb{E}\left[\sum_{k=1}^{n}\sum_{i=T_{k-1}}^{S_k-1}\frac{X_1^n(i)-\mathbb{1}_{\{X_1^n(i)>0\}}}{3U^n(i)+4V^n(i)+X^n(i)-\mathbb{1}_{\{X^n(i)>0\}}}\right] &\le \frac{C}{2}\sum_{k=0}^{\infty}\frac{\sqrt{n\epsilon+1}}{(C^2-5)^{k/2}}\\ &= \frac{C\sqrt{C^2-5}}{2(\sqrt{C^2-5}-1)}\sqrt{n\epsilon+1}.\end{aligned}$$

Putting this together with (3.21), (3.22) and (3.23) we get

$$\begin{aligned}&\mathbb{E}[N^n(n)-N^n(\lfloor n(1-\epsilon)\rfloor)]\\ &\qquad\le \frac{1}{6}(\log(\lceil n\epsilon\rceil)+1)+C\sqrt{n\epsilon+1}+\frac{C\sqrt{C^2-5}}{2(\sqrt{C^2-5}-1)}\sqrt{n\epsilon+1}+\mathbb{E}[V^n(\lfloor(1-\epsilon)n\rfloor)].\end{aligned}$$

Using (3.20), we have

$$\limsup_{n\to\infty} \frac{\mathbb{E}[V^n(\lfloor n(1-\epsilon)\rfloor)]}{\alpha(n)} \leq 2\epsilon^{4/3}.$$

Hence,

$$\limsup_{n\to\infty} \frac{1}{\alpha(n)\vee\sqrt{n}}\mathbb{E}[N^n(n) - N^n(\lfloor n(1-\epsilon)\rfloor)] \leq \left(C + \frac{C\sqrt{C^2-5}}{2(\sqrt{C^2-5}-1)}\right)\sqrt{\epsilon} + 2\epsilon^{4/3}. \tag{3.25}$$

Applying Markov's inequality and taking $\epsilon \to 0$ then yields that for $\delta > 0$,

$$\lim_{\epsilon\to 0}\limsup_{n\to\infty} \mathbb{P}\left(\frac{1}{\alpha(n)\vee\sqrt{n}}(N^n(n) - N^n(\lfloor n(1-\epsilon)\rfloor)) > \delta\right) = 0.$$

(c) We will show that $\frac{1}{\alpha(n)\vee\sqrt{n}}\mathbb{E}\Big[\sup_{\lfloor n(1-\epsilon)\rfloor\leq i\leq n} X^n(i)\Big]$ is small. We again use the coupling of X_1^n with Y^n, but started now with $Y^n(\lfloor n(1-\epsilon)\rfloor) = X_1^n(\lfloor n(1-\epsilon)\rfloor) + 2$. This gives

$$(Y^n(\lfloor n(1-\epsilon)\rfloor + k))_{k\geq 0} \overset{d}{=} (2 + |X_1^n(\lfloor (1-\epsilon)n\rfloor) + Z(k)|)_{k\geq 0}.$$

Since $X^n(i) = X_1^n(i) + X_2^n(i)$ with $X_2^n(i) \leq V^n(\lfloor n(1-\epsilon)\rfloor)$ for $i \geq \lfloor n(1-\epsilon)\rfloor$, we have

$$\sup_{\lfloor n(1-\epsilon)\rfloor\leq i\leq n} X^n(i) \leq V^n(\lfloor n(1-\epsilon)\rfloor) + \sup_{\lfloor n(1-\epsilon)\rfloor\leq i\leq n} Y^n(i)$$

and so we get

$$\begin{aligned}
&\frac{1}{\alpha(n)\vee\sqrt{n}}\mathbb{E}\left[\sup_{\lfloor n(1-\epsilon)\rfloor\leq i\leq n} X^n(i)\right] \\
&\quad\leq \frac{2 + \mathbb{E}[V^n(\lfloor n(1-\epsilon)\rfloor)] + \mathbb{E}[X_1^n(\lfloor n(1-\epsilon)\rfloor)]}{\alpha(n)\vee\sqrt{n}} + \frac{1}{\sqrt{n}}\mathbb{E}\left[\sup_{0\leq i\leq\lceil n\epsilon\rceil} |Z(i)|\right] \\
&\quad\leq \frac{2 + \mathbb{E}[V^n(\lfloor n(1-\epsilon)\rfloor)] + \mathbb{E}[X_1^n(\lfloor n(1-\epsilon)\rfloor)]}{\alpha(n)\vee\sqrt{n}} + 2\sqrt{\epsilon},
\end{aligned}$$

by Doob's L^2 inequality. We have already shown above that

$$\lim_{\epsilon\to 0}\limsup_{n\to\infty} \frac{\mathbb{E}[V^n(\lfloor n(1-\epsilon)\rfloor)]}{\alpha(n)\vee\sqrt{n}} = 0.$$

For $\alpha(n)/\sqrt{n} \to a$, we have

$$\frac{X^n(\lfloor n(1-\epsilon)\rfloor)}{\sqrt{n}} \xrightarrow{d} X^a_{1-\epsilon}$$

as $n \to \infty$. By Lemma 60, $X^n(\lfloor n(1-\epsilon)\rfloor)/\sqrt{n}$ is a uniformly integrable sequence of random variables, and so we obtain

$$\mathbb{E}\Big[\frac{X^n(\lfloor n(1-\epsilon)\rfloor)}{\sqrt{n}}\Big] \to \mathbb{E}\big[X^a_{1-\epsilon}\big]$$

as $n \to \infty$. Hence,

$$\limsup_{n\to\infty} \frac{\mathbb{E}\Big[\sup_{\lfloor n(1-\epsilon)\rfloor \le i \le n} X^n(i)\Big]}{\sqrt{n}} \le 2a\epsilon^{4/3} + \mathbb{E}\big[X^a_{1-\epsilon}\big] + 2\sqrt{\epsilon}.$$

By the final statement of Proposition 68, we have

$$\mathbb{P}\big(X^a_{1-\epsilon} > r\sqrt{\epsilon}\big) \le e^{-(r-a)^2/6}$$

for $r > a$ and it follows, in particular, that $\mathbb{E}\big[X^a_{1-\epsilon}\big] \le C\sqrt{\epsilon}$ for some constant $C > 0$.

For $\alpha(n) >> \sqrt{n}$, we have $X^n(\lfloor n(1-\epsilon)\rfloor)/\alpha(n) \xrightarrow{p} \epsilon^{2/3} - \epsilon^{4/3}$. Again, using the uniform integrability from Lemma 60, we get $\mathbb{E}[X^n(\lfloor n(1-\epsilon)\rfloor)]/\alpha(n) \xrightarrow{p} \epsilon^{2/3} - \epsilon^{4/3}$ and so

$$\limsup_{n\to\infty} \frac{\mathbb{E}\Big[\sup_{\lfloor n(1-\epsilon)\rfloor \le i \le n} X^n(i)\Big]}{\alpha(n)} \le 2\epsilon^{4/3} + \epsilon^{2/3} - \epsilon^{4/3}.$$

Hence, for any $\alpha(n)$,

$$\lim_{\epsilon\to 0}\limsup_{n\to\infty} \frac{\mathbb{E}\Big[\sup_{\lfloor n(1-\epsilon)\rfloor \le i \le n} X^n(i)\Big]}{\alpha(n) \vee \sqrt{n}} = 0 \tag{3.26}$$

and another application of Markov's inequality gives

$$\lim_{\epsilon\to 0}\limsup_{n\to\infty} \mathbb{P}\Bigg(\frac{1}{\sqrt{n}} \sup_{\lfloor n(1-\epsilon)\rfloor \le i \le n} X^n(i) > \delta\Bigg) = 0.$$

(d) Recall that $(M_X^n(k), k \geq 0)$ defined by

$$M_X^n(k) = X^n(k) + \frac{1}{2}V^n(k) - L^n(k) + \sum_{i=0}^{k-1} \frac{2X^n(i) - 2\mathbb{1}_{\{X^n(i)>0\}}}{3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}}} \quad (3.27)$$

is a martingale. It follows that

$$\mathbb{E}[X^n(n) - X^n(\lfloor n(1-\epsilon)\rfloor] = \mathbb{E}[L^n(n) - L^n(\lfloor n(1-\epsilon)\rfloor)] + \frac{1}{2}\mathbb{E}[V^n(\lfloor n(1-\epsilon)\rfloor) - V^n(n)] - \mathbb{E}\left[\sum_{i=\lfloor n(1-\epsilon)\rfloor}^{n-1} \frac{2X^n(i) - 2\mathbb{1}_{\{X^n(i)>0\}}}{3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}}}\right].$$

Now we have

$$\mathbb{E}[N^n(n) - N^n(\lfloor n(1-\epsilon)\rfloor)] \geq \mathbb{E}\left[\sum_{i=\lfloor n(1-\epsilon)\rfloor}^{n-1} \frac{X^n(i) - \mathbb{1}_{\{X^n(i)>0\}}}{3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}}}\right].$$

Hence,

$$\mathbb{E}[L^n(n) - L^n(\lfloor n(1-\epsilon)\rfloor)] \leq \mathbb{E}[X^n(n)] + 2\mathbb{E}[N^n(n) - N^n(\lfloor n(1-\epsilon)\rfloor)].$$

Applying Markov's inequality and using (3.25) and (3.26), we get that for any $\delta > 0$,

$$\lim_{\epsilon\to 0} \limsup_{n\to\infty} \mathbb{P}\left(\frac{1}{\alpha(n) \vee \sqrt{n}}(L^n(n) - L^n(\lfloor n(1-\epsilon)\rfloor)) > \delta\right) = 0. \qquad \square$$

Lemma 73. *We have*

$$U^n(n) \leq N^n(n) + 3V^n(0).$$

Proof. At each step we either connect to a node of degree 3 or 4, or there is a clash. Each node of degree 3 is removed after we connect to it, and each node of degree 4 can be visited at most twice before being removed. Thus, during the first n steps, a node of degree 3 is removed at least $n - N^n(n) - 2V^n(0)$ times. Moreover, at most $V^n(0)$ nodes of degree 4 become nodes of degree 3. Hence,

$$U^n(n) \leq U^n(0) - (n - N^n(n) - 2V^n(0)) + V^n(0) = N^n(n) + 3V^n(0). \qquad \square$$

We finally turn to the behaviour of our processes on the time-interval $[n, \zeta_n]$.

Proposition 74. *As $n \to \infty$,*

(a) $\zeta_n/n \xrightarrow{p} 1$,

(b) $\frac{1}{\alpha(n)\vee\sqrt{n}} \sup_{n\le i\le \zeta_n} X^n(i) \xrightarrow{p} 0$,

(c) $\frac{1}{\alpha(n)\vee\sqrt{n}}(N^n(\zeta_n) - N^n(n)) \xrightarrow{p} 0$,

(d) $\frac{1}{\alpha(n)\vee\sqrt{n}}(L^n(\zeta_n) - L^n(n)) \xrightarrow{p} 0$.

Proof. Write $\bar{\alpha}(n) = \alpha(n) \vee \sqrt{n}$.

(a) As in each step we remove at least two half-edges, by the same argument as gave us (3.5), we have $\zeta_n - n \le 3U^n(n) + 4V^n(n) \le 3U^n(n) + 4\alpha(n)$ and so Lemma 73 implies that

$$\zeta_n - n \le 3N^n(n) + 13\alpha(n). \tag{3.28}$$

If $\alpha(n)/\sqrt{n} \to a$ then $N^n(n)/\sqrt{n} \xrightarrow{d} N_1^a$, and it follows that $\zeta_n/n \xrightarrow{p} 1$ and, indeed, that $(\zeta_n - n)/\sqrt{n}$ is a tight sequence of random variables. If, on the other hand, $\alpha(n) >> \sqrt{n}$, we have $N^n(n)/\alpha(n) \xrightarrow{p} 1/4$, and again we get $\zeta_n/n \xrightarrow{p} 1$ with $(\zeta_n - n)/\alpha(n)$ a tight sequence of random variables.

(b) Consider $\sup_{n\le i\le\zeta_n} X^n(i)$. By the usual coupling with Y^n, we have

$$\sup_{n\le i\le\zeta_n} X^n(i) \le V^n(n) + \sup_{n\le i\le\zeta_n} Y^n(i),$$

where

$$(Y^n(n+i))_{0\le i\le\zeta_n-n} \stackrel{d}{=} (2 + |X_1^n(n) + Z(i)|)_{0\le i\le\zeta_n-n}$$

Fix $\delta > 0$. Then for any $C > 13$, since $\frac{\sup_{0\le i\le C\bar{\alpha}(n)}|Z(i)|}{\sqrt{\bar{\alpha}(n)}}$ is bounded in L^2, we have

$$\begin{aligned}
&\limsup_{n\to\infty} \mathbb{P}\left(\frac{\sup_{0\le i\le\zeta_n-n}|Z(i)|}{\bar{\alpha}(n)} > \delta/2\right) \\
&\quad\le \limsup_{n\to\infty} \mathbb{P}\left(\frac{\sup_{0\le i\le C\bar{\alpha}(n)}|Z(i)|}{\bar{\alpha}(n)} > \delta/2\right) + \limsup_{n\to\infty} \mathbb{P}(\zeta_n - n > C\bar{\alpha}(n)) \\
&\quad\le \mathbb{P}(N_1 > (C-13)/3),
\end{aligned}$$

where for the last inequality we recall (3.28). We have $V^n(n)/\bar{\alpha}(n) \xrightarrow{p} 0$ by Proposition 72 (a) and $X^n(n)/\bar{\alpha}(n) \xrightarrow{p} 0$, so since $C > 13$ was arbitrary, we obtain

$$\frac{1}{\bar{\alpha}(n)} \sup_{n \leq i \leq \zeta_n} X^n(i) \xrightarrow{p} 0.$$

(c) We have

$$\begin{aligned}
&\mathbb{P}\bigg(\frac{1}{\bar{\alpha}(n)}(N^n(\zeta_n) - N^n(n)) > \delta\bigg) \\
&\leq \mathbb{P}\bigg(\frac{1}{\bar{\alpha}(n)}(N^n(\zeta_n) - N^n(n)) > \delta, \frac{N^n(n)}{\bar{\alpha}(n)} \leq \frac{1}{\epsilon}\bigg) + \mathbb{P}\bigg(\frac{N^n(n)}{\bar{\alpha}(n)} > \frac{1}{\epsilon}\bigg) \qquad (3.29) \\
&\leq \frac{1}{\delta \bar{\alpha}(n)} \mathbb{E}\Big[(N^n(\zeta_n) - N^n(n)) \mathbb{1}_{\{\frac{N^n(n)}{\bar{\alpha}(n)} \leq \frac{1}{\epsilon}\}}\Big] + \epsilon \frac{\mathbb{E}[N^n(n)]}{\bar{\alpha}(n)}.
\end{aligned}$$

Let $\eta_n = \inf\{i \geq n : U^n(i) \leq \epsilon \bar{\alpha}(n)\}$. We have that $\zeta_n - \eta_n \leq 3U^n(\eta_n) + 4V^n(n) \leq 3\epsilon\bar{\alpha}(n)$ and $\zeta_n - n \leq 3N^n(n)$. Now

$$\begin{aligned}
&\mathbb{E}\Big[(N^n(\zeta_n) - N^n(n)) \mathbb{1}_{\{\frac{N^n(n)}{\bar{\alpha}(n)} \leq \frac{1}{\epsilon}\}}\Big] \\
&\leq \mathbb{E}\Bigg[\sum_{i=n}^{\zeta_n - 1} \frac{3U^n(i) + 2(X^n(i) - \mathbb{1}_{\{X^n(i)>0\}})(3U^n(i) + 4V^n(i) + X^n(i) - 3)}{2(3U^n(i) + 4V^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}})(3U^n(i) + X^n(i) - \mathbb{1}_{\{X^n(i)>0\}} - 2)} \mathbb{1}_{\{\frac{N^n(n)}{\bar{\alpha}(n)} \leq \frac{1}{\epsilon}\}}\Bigg] \\
&\leq \mathbb{E}\Bigg[\sum_{i=n}^{\zeta_n - 1} \bigg(\frac{X^n(i) + 1}{3U^n(i)} \wedge 1\bigg) \mathbb{1}_{\{\frac{N^n(n)}{\bar{\alpha}(n)} \leq \frac{1}{\epsilon}\}}\Bigg] \\
&\leq \mathbb{E}\Bigg[\sum_{i=n}^{\eta_n} \bigg(\frac{X^n(i) + 1}{3\epsilon\bar{\alpha}(n)} \wedge 1\bigg) \mathbb{1}_{\{\frac{N^n(n)}{\bar{\alpha}(n)} \leq \frac{1}{\epsilon}\}}\Bigg] + 3\epsilon\bar{\alpha}(n) \\
&\leq \frac{1}{\epsilon^2} \mathbb{E}\Bigg[\Bigg(\bigg(1 + \sup_{n \leq i \leq \zeta_n} X^n(i)\bigg) \wedge 3\epsilon\bar{\alpha}(n)\Bigg) \mathbb{1}_{\{\frac{N^n(n)}{\bar{\alpha}(n)} \leq \frac{1}{\epsilon}\}}\Bigg] + 3\epsilon\bar{\alpha}(n).
\end{aligned} \qquad (3.30)$$

By the coupling, we have that the expectation is bounded by

$$\mathbb{E}\Bigg[\bigg(3 + V^n(n) + X^n(n) + \sup_{0 \leq i \leq 3\bar{\alpha}(n)/\epsilon} |Z(i)|\bigg) \wedge 3\epsilon\bar{\alpha}(n)\Bigg].$$

Now

$$\frac{1}{\bar{\alpha}(n)}\bigg(3 + V^n(n) + X^n(n) + \sup_{0 \leq i \leq 3\bar{\alpha}(n)/\epsilon} |Z(i)|\bigg) \wedge 3\epsilon$$

is a bounded random variable which converges to 0 in probability, and so from (3.30) we get

$$\limsup_{n\to\infty}\frac{1}{\bar{\alpha}(n)}\mathbb{E}\Big[(N^n(\zeta_n)-N^n(n))\mathbb{1}_{\{\frac{N^n(n)}{\bar{\alpha}(n)}\le\frac{1}{\epsilon}\}}\Big]\le 3\epsilon.$$

Now note that combining Theorem 64, (3.19) and (3.25), we have that there exists a constant $C>0$ such that

$$\sup_{n\ge 1}\frac{\mathbb{E}[N^n(n)]}{\bar{\alpha}(n)}<C$$

and so using (3.29) we obtain

$$\limsup_{n\to\infty}\mathbb{P}\Big(\frac{1}{\bar{\alpha}(n)}(N^n(\zeta_n)-N^n(n))>\delta\Big)\le\frac{3\epsilon}{\delta}+C\epsilon.$$

Since $\epsilon>0$ was arbitrary, we get

$$\limsup_{n\to\infty}\mathbb{P}\Big(\frac{1}{\bar{\alpha}(n)}(N^n(\zeta_n)-N^n(n))>\delta\Big)=0.$$

(d) Finally, we observe that using the martingale M^n_X defined in (3.27) and the optional stopping theorem, we have

$$\begin{aligned}0&\le\mathbb{E}\Big[(L^n(\zeta_n)-L^n(n))\mathbb{1}_{\{\frac{N^n(n)}{\bar{\alpha}(n)}\le\frac{1}{\epsilon}\}}\Big]\\&\le\mathbb{E}\Big[(X^n(\zeta_n)-X^n(n))\mathbb{1}_{\{\frac{N^n(n)}{\bar{\alpha}(n)}\le\frac{1}{\epsilon}\}}\Big]+2\mathbb{E}\Big[(N^n(\zeta_n)-N^n(n))\mathbb{1}_{\{\frac{N^n(n)}{\bar{\alpha}(n)}\le\frac{1}{\epsilon}\}}\Big]\\&\le 2\mathbb{E}\Big[(N^n(\zeta^n)-N^n(n))\mathbb{1}_{\{\frac{N^n(n)}{\bar{\alpha}(n)}\le\frac{1}{\epsilon}\}}\Big]\end{aligned}$$

since $X^n(\zeta_n)=0$ and $X^n(n)\ge 0$. So then

$$\begin{aligned}&\mathbb{P}\Big(\frac{1}{\bar{\alpha}(n)}(L^n(\zeta_n)-L^n(n))>\delta\Big)\\&\quad\le\mathbb{P}\Big(\frac{1}{\bar{\alpha}(n)}(L^n(\zeta_n)-L^n(n))>\delta,\frac{N^n(n)}{\bar{\alpha}(n)}\le\frac{1}{\epsilon}\Big)+\mathbb{P}\Big(\frac{N^n(n)}{\bar{\alpha}(n)}>\frac{1}{\epsilon}\Big)\\&\quad\le\frac{2}{\delta\bar{\alpha}(n)}\mathbb{E}\Big[(N^n(\zeta_n)-N^n(n))\mathbb{1}_{\{\frac{N^n(n)}{\bar{\alpha}(n)}\le\frac{1}{\epsilon}\}}\Big]+\epsilon\frac{\mathbb{E}[N^n(n)]}{\bar{\alpha}(n)}\end{aligned}$$

and the rest of the argument goes through as before. □

We now have all the elements needed to prove Theorem 59.

Proof of Theorem 59. (i) From Theorem 69 and Lemma 71, for any $\epsilon > 0$ we have

$$\Big(\frac{X^n(\lfloor ns\rfloor)}{\sqrt{n}}, \frac{L^n(\lfloor ns\rfloor)}{\sqrt{n}}, \frac{N^n(\lfloor ns\rfloor)}{\sqrt{n}}, 0 \le s \le 1-\epsilon\Big) \xrightarrow{d} (X^a_s, L^a_s, N^a_s, 0 \le s \le 1-\epsilon).$$

By Proposition 68, we have

$$(X^a_{1-\epsilon}, L^a_{1-\epsilon}, N^a_{1-\epsilon}) \to (0, L^a_1, N^a_1)$$

almost surely as $\epsilon \to 0$. The principle of accompanying laws (Theorem 3.2 of Billingsley [9]) gives that the statement of Proposition 72 is exactly what we need in order to deduce that

$$\Big(\frac{X^n(\lfloor ns\rfloor)}{\sqrt{n}}, \frac{L^n(\lfloor ns\rfloor)}{\sqrt{n}}, \frac{N^n(\lfloor ns\rfloor)}{\sqrt{n}}, 0 \le s \le 1\Big) \xrightarrow{d} (X^a_s, L^a_s, N^a_s, 0 \le s \le 1).$$

Proposition 74 then allows us to conclude that

$$\Big(\frac{X^n(\lfloor ns\rfloor)}{\sqrt{n}}, \frac{L^n(\lfloor ns\rfloor)}{\sqrt{n}}, \frac{N^n(\lfloor ns\rfloor)}{\sqrt{n}}, 0 \le s \le \zeta_n/n\Big) \xrightarrow{d} (X^a_s, L^a_s, N^a_s, 0 \le s \le 1),$$

as desired.

(ii) From Theorem 64, for any $\epsilon > 0$, we have

$$\Big(\frac{X^n(\lfloor ns\rfloor)}{\alpha(n)}, \frac{L^n(\lfloor ns\rfloor)}{\alpha(n)}, \frac{N^n(\lfloor ns\rfloor)}{\alpha(n)}, 0 \le s \le 1-\epsilon\Big) \xrightarrow{d} (x(s), 0, m(s), 0 \le s \le 1-\epsilon).$$

It is straightforward that $x(1-\epsilon) \to 0$ and that $m(1-\epsilon) \to 1/4$ as $\epsilon \to 0$. By another application of the principle of accompanying laws, and Propositions 72 and 74, we may then conclude that

$$\Big(\frac{X^n(\lfloor ns\rfloor)}{\alpha(n)}, \frac{L^n(\lfloor ns\rfloor)}{\alpha(n)}, \frac{N^n(\lfloor ns\rfloor)}{\alpha(n)}, 0 \le s \le \zeta_n/n\Big) \xrightarrow{d} (x(s), 0, m(s), 0 \le s \le 1),$$

as desired. $\square$

3.7 Appendix

3.7.1 Proof of the invariance principle

Proof of Theorem 70. Let $\theta^n(r) = \inf\{t \geq 0 : Q^n(t) > \int_0^t \sup_{|y|\leq r} q(s,y)ds + 1\}$. Set $\tilde{M}_r^n(\cdot) := M^n(\cdot \wedge \theta^n(r) \wedge \tau^n(r))$. Relative compactness of $\{\tilde{M}_r^n\}_{n\geq 1}$ follows as in Theorem 4.1 of Chapter 7 of Ethier and Kurtz. Assumptions (c) and (d) imply that any subsequential weak limit has sample paths in $\mathbb{C}(\mathbb{R}_+, \mathbb{R})$ almost surely. Then all of the conditions in Assumption 4.1 of Kang and Williams [33] hold for the processes stopped at $\tau^n(r)$. It follows by their Theorem 4.2 that the sequence of processes

$$\{(Y^n, M^n, L^n)\}_{n\geq 1}$$

is tight and such that any subsequential weak limit almost surely has continuous sample paths.

Now fix $r_0 > 0$ and let $\{Y^{n_k}(\cdot \wedge \tau^{n_k}(r_0)), M^{n_k}(\cdot \wedge \tau^{n_k}(r_0)), L^{n_k}(\cdot \wedge \tau^{n_k}(r_0))\}_{k\geq 1}$ be a convergent subsequence with limit $Y_{r_0}, M_{r_0}, L_{r_0}$. Let $\tau_{r_0}(r) = \inf\{t \geq 0 : |Y_{r_0}(t)| \geq r\}$. Then for all but countably many $r < r_0$ (i.e. those such that $\mathbb{P}(\lim_{s\to r} \tau_{r_0}(s) = \tau_{r_0}(r)) = 1$), we have

$$\begin{aligned}
&(Y^{n_k}(\cdot \wedge \tau^{n_k}(r)), M^{n_k}(\cdot \wedge \tau^{n_k}(r)), L^{n_k}(\cdot \wedge \tau^{n_k}(r)), \tau^{n_k}(r)) \\
&\xrightarrow{d} (Y_{r_0}(\cdot \wedge \tau_{r_0}(r)), M_{r_0}(\cdot \wedge \tau_{r_0}(r)), L_{r_0}(\cdot \wedge \tau_{r_0}(r)), \tau_{r_0}(r)).
\end{aligned}$$

Conditions (f), (c) and (d) guarantee that $M^{n_k}(\cdot \wedge \tau^{n_k}(r_0))$ is a uniformly integrable martingale. Conditions (b), (g) and (e) guarantee that $(M^{n_k}(\cdot \wedge \tau^{n_k}(r_0)))^2 - Q^{n_k}(\cdot \wedge \tau^{n_k}(r_0))$ is a uniformly integrable martingale. Thus, by Problem 7 from Chapter 7 of Ethier and Kurtz, as in the proof of Theorem 1.4(b) there, we must then have that

$$M_{r_0}(t \wedge \tau_{r_0}(r)) = Y_{r_0}(t \wedge \tau_{r_0}(r)) - L_{r_0}(t \wedge \tau_{r_0}(r)) - \int_0^{t\wedge\tau_{r_0}(r)} b(s, Y_{r_0}(s))ds$$

and

$$M_{r_0}(t \wedge \tau_{r_0}(r))^2 - \int_0^{t\wedge\tau_{r_0}(r)} q(s, Y_{r_0}(s))ds$$

are continuous martingales. It follows that

$$M_{r_0}(t \wedge \tau_{r_0}(r)) = \int_0^{t\wedge\tau_{r_0}(r)} \sigma(s, Y_{r_0}(s))dW(s)$$

for some Brownian motion W. But then we have

$$Y_{r_0}(t \wedge \tau_{r_0}(r)) = \int_0^{t\wedge\tau_{r_0}(r)} b(s, Y_{r_0}(s))ds + \int_0^{t\wedge\tau_{r_0}(r)} \sigma(s, Y_{r_0}(s))dW(s) + L_{r_0}(t \wedge \tau_{r_0}(r)).$$

Moreover, by (i), $L_{r_0}(0) = 0$, L_{r_0} is non-decreasing, and $\int_0^{t\wedge\tau_{r_0}(r)} \mathbb{1}_{\{Y_{r_0}(s)>0\}} dL_{r_0}(s) = 0$. So (Y_{r_0}, L_{r_0}) solves a stopped version of the reflected SDE.

But if (Y, L) is the unique solution of the reflected SDE with Brownian motion W then $(Y(\cdot \wedge \tau(r)), L(\cdot \wedge \tau(r)))$ is the unique solution of the stopped version, where $\tau(r) = \inf\{t \geq 0 : |Y(s)| \geq r\}$. In consequence, $(Y^n(\cdot \wedge \tau^n(r)), L^n(\cdot, \wedge\tau^n(r))) \xrightarrow{d} (Y(\cdot \wedge \tau(r)), L(\cdot \wedge \tau(r))$ for any r such that $\mathbb{P}(\lim_{s\to r} \tau(s) = \tau(r)) = 1$. But $\tau(r) \to \infty$ as $r \to \infty$, since Y has continuous sample paths. Hence, $(Y^n, L^n) \xrightarrow{d} (Y, L)$. □

3.7.2 A hitting probability calculation

Fix $d \in (0, 1/2)$ and an integer $b \geq 1$. Suppose that A is a random walk with step-distribution

$$\begin{aligned}\mathbb{P}(A(i+1) - A(i) = -1|A(i)) &= \frac{1}{2},\\ \mathbb{P}(A(i+1) - A(i) = 1|A(i)) &= \frac{1}{2} - d,\\ \mathbb{P}(A(i+1) - A(i) = 2|A(i)) &= d.\end{aligned}$$

Lemma 75. *The probability, started from 1, that A hits 0 before $\{b, b+1\}$ is given by*

$$1 - \frac{4d(-1-\sqrt{1+8d})^b+4d(-1+\sqrt{1+8d})^b}{(-2)^{b+1}\sqrt{1+8d}+4d(-1-\sqrt{1+8d})^b+4d(-1+\sqrt{1+8d})^b-(-1-\sqrt{1+8d})^{b+1}-(-1+\sqrt{1+8d})^{b+1}}.$$

Proof. For $0 \leq k \leq b+1$, let

$$h_k = \mathbb{P}(A \text{ hits } 0 \text{ before } \{b, b+1\}|A(0) = k).$$

Then for $1 \leq k \leq b-1$,

$$h_k = \frac{1}{2}h_{k-1} + \left(\frac{1}{2} - d\right)h_{k+1} + dh_{k+2}$$

and elementary calculations yield that

$$h_k = \varphi + \beta\lambda_-^k + \gamma\lambda_+^k, \quad 0 \leq k \leq b+1,$$

where

$$\lambda_\pm = \frac{-1 \pm \sqrt{1+8d}}{4d}.$$

The boundary conditions are $h_0 = 1$, $h_b = h_{b+1} = 0$. Solving for the constants gives

$$\begin{aligned}
\varphi &= \frac{\lambda_-^{b+1}\lambda_+^b - \lambda_-^b\lambda_+^{b+1}}{\lambda_-^{b+1}\lambda_+^b - \lambda_-^b\lambda_+^{b+1} - \lambda_-^{b+1} + \lambda_-^b - \lambda_+^{b+1} + \lambda_+^b}, \\
\beta &= \frac{\lambda_+^b - \lambda_+^{b+1}}{\lambda_-^{b+1}\lambda_+^b - \lambda_-^b\lambda_+^{b+1} - \lambda_-^{b+1} + \lambda_-^b - \lambda_+^{b+1} + \lambda_+^b}, \\
\gamma &= \frac{\lambda_-^b - \lambda_-^{b+1}}{\lambda_-^{b+1}\lambda_+^b - \lambda_-^b\lambda_+^{b+1} - \lambda_-^{b+1} + \lambda_-^b - \lambda_+^{b+1} + \lambda_+^b}.
\end{aligned}$$

Simplifying the $k=1$ case, we obtain

$$h_1 = 1 - \frac{4d(-1-\sqrt{1+8d})^b + 4d(-1+\sqrt{1+8d})^b}{(-2)^{b+1}\sqrt{1+8d} + 4d(-1-\sqrt{1+8d})^b + 4d(-1+\sqrt{1+8d})^b - (-1-\sqrt{1+8d})^{b+1} - (-1+\sqrt{1+8d})^{b+1}}. \quad \square$$

Bibliography

[1] Aldous, D. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab. 25*, 2 (1997), 812–854.

[2] Aronson, J., Frieze, A., and Pittel, B. G. Maximum matchings in sparse random graphs: Karp-Sipser revisited. *Random Structures Algorithms 12*, 2 (1998), 111–177.

[3] Arratia, R., Barbour, A. D., and Tavaré, S. *Logarithmic combinatorial structures: a probabilistic approach.* EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich, 2003.

[4] Bal, D., Bennett, P., Bohman, T., and Frieze, A. A greedy algorithm for finding a large 2-matching on a random cubic graph, 2012.

[5] Barbour, A. D. On a functional central limit theorem for markov population processes. *Advances in Applied Probability 6*, 1 (1974), 21–39.

[6] Bauer, M., and Golinelli, O. Core percolation in random graphs: a critical phenomena analysis. *Eur. Phys. J.*, B24 (2001), 339–352.

[7] Bender, E. A., and Canfield, E. R. The asymptotic number of labeled graphs with given degree sequences. *J. Combinatorial Theory Ser. A 24*, 3 (1978), 296–307.

[8] Bertoin, J., and Goldschmidt, C. Dual random fragmentation and coagulation and an application to the genealogy of Yule processes. In *Mathematics and computer science. III*, Trends Math. Birkhäuser, Basel, 2004, pp. 295–308.

[9] Billingsley, P. *Weak convergence of measures: Applications in probability.* Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1971. Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 5.

[10] Blum, N. *A new approach to maximum matching in general graphs.* Springer Berlin Heidelberg, Berlin, Heidelberg, 1990, pp. 586–597.

[11] Bohman, T., and Frieze, A. Karp-Sipser on random graphs with a fixed degree sequence. *Combin. Probab. Comput. 20*, 5 (2011), 721–741.

[12] Bollobás, B. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin. 1*, 4 (1980), 311–316.

[13] Bollobás, B. The evolution of random graphs, 1984.

[14] Darling, R. W. R., and Norris, J. R. Differential equation approximations for Markov chains. *Probab. Surv. 5* (2008), 37–79.

[15] Durrett, R. *Random graphs dynamics.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2007.

[16] Dyer, M., Frieze, A., and Pittel, B. The average performance of the greedy matching algorithm. *Ann. Appl. Probab. 3*, 2 (05 1993), 526–552.

[17] Erdős, P., and Rényi, A. On the strength of connectedness of a random graph. *Acta. Math. Acad. Sci. Hungar.*, 12 (1961), 261–267.

[18] Erdős, P., and Rényi, A. On the existence of a factor of degree one of a connected random graph. *Acta. Math. Acad. Sci. Hungar.*, 17 (1966), 359–368.

[19] Erdős, P., and Rényi, A. On random graphs 1. *Publicationes Mathematicae (Debrecen) 6* (1959), 290–297.

[20] Erdős, P., and Rényi, A. On the evolution of random graphs. In *Publ. Math. Inst. Hung. Acad. Sci* (1960), pp. 17–61.

[21] Ethier, S. N., and Kurtz, T. G. *Markov processes.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

[22] Frieze, A. M., and McDiarmid, C. Algorithmic theory of random graphs. *Random Struct. Algorithms 10*, 1-2 (1997), 5–42.

[23] Frieze, A. M., and Pittel, B. Perfect matchings in random graphs with prescribed minimal degree. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms, January 12-14, 2003, Baltimore, Maryland, USA.* (2003), pp. 148–157.

[24] Frieze, A. M., Radcliffe, A. J., and Suen, S. Analysis of a simple greedy matching algorithm on random cubic graphs. *Combinatorics, Probability & Computing 4* (1995), 47–66.

[25] Gabow, H. N., and Tarjan, R. E. Faster scaling algorithms for general graph matching problems. *J. ACM 38*, 4 (Oct. 1991), 815–853.

[26] Janson, S. Poisson convergence and poisson processes with applications to random graphs. *Stochastic Processes and their Applications 26* (1987), 1–30.

[27] Janson, S. Monotonicity, asymptotic normality and vertex degrees in random graphs. *Bernoulli 13*, 4 (2007), 952–965.

[28] Janson, S., and Luczak, M. A simple solution to the k-core problem. *Random Struct. Algorithms 30*, 1-2 (2007), 50–62.

[29] Janson, S., and Luczak, M. Asymptotic normality of the k-core in random graphs. *The Annals of Applied Probability 18*, 3 (2009), 1085–1137.

[30] Janson, S., Luczak, T., and Rucinski, A. *Random graphs.* John Wiley & Sons, 2000.

[31] Joos, F., Perarnau, G., Rautenbach, R., and Reed, B. How to determine if a random graph with a fixed degree sequence has a giant component. *Probab. Theory Related Fields*, to appear. `arXiv:1601.03714`, 2017+.

[32] Joseph, A. The component sizes of a critical random graph with given degree sequence. *The Annals of Applied Probability 24*, 6 (2014), 2560–2594.

[33] Kang, W., and Williams, R. J. An invariance principle for semimartingale reflecting Brownian motions in domains with piecewise smooth boundaries. *Ann. Appl. Probab. 17*, 2 (2007), 741–779.

[34] Karp, R. M., and Sipser, M. Maximum matchings in sparse random graphs. In *Proceedings of the Twenty-second Annual Symposium on Foundations of Computer Science* (1981), IEEE, pp. 364–375.

[35] Kurtz, T. G. Solutions of ordinary differential equations as limits of pure jump markov processes. *Journal of Applied Probability 7* (1970), 49–58.

[36] Kurtz, T. G. Limit theorems for sequences of jump markov processes approximating ordinary differential processes. *Journal of Applied Probability 8*, 2 (1971), 344–356.

[37] Kurtz, T. G., and Protter, P. Weak limit theorems for stochastic integrals and stochastic differential equations. *The Annals of Probability 19*, 3 (1991), 1035–1070.

[38] Lovász, L., and Plummer, M. Matching theory. *Annals of Discrete Mathematics*, 29 (1986).

[39] Luczak, T., Pittel, B., and Wierman, J. C. The structure of a random graph at the point of the phase transition. *Transactions of the American Mathematical Society 341*, 2 (1994), 721–748.

[40] Mitzenmacher, M., and Upfal, E. *Probability and computing: randomized algorithms and probabilistic analysis.* Cambridge University Press, 2005.

[41] Molloy, M., and Reed, B. A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms 6*, 2-3 (1995), 161–180.

[42] Molloy, M., and Reed, B. The size of the giant component of a random graph with a given degree sequence. *Comb. Probab. Comput. 7*, 3 (Sept. 1998), 295–305.

[43] Mulder, H. *Julius Petersen's theory of regular graphs.* Discrete Mathematics 100 (1992). North Holland, 1991.

[44] Pilipenko, A. *An introduction to Stochastic Differential Equations with Reflection.* Lectures in Pure and Applied Mathematics. Potsdam University Press, 2014.

[45] Pittel, B. On tree census and the giant component in sparse random graphs. *Random Struct. Algorithms 1*, 3 (1990), 311–342.

[46] Pittel, B., Spencer, J., and Wormald, N. C. Sudden emergence of a giant k-core in a random graph. *J. Comb. Theory, Ser. B 67*, 1 (1996), 111–151.

[47] Pittel, B., and Wormald, N. C. Counting connected graphs inside-out. *J. Comb. Theory, Ser. B 93*, 2 (2005), 127–172.

[48] Revuz, D., and Yor, M. *Continuous martingales and Brownian motion*, third ed., vol. 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.

[49] Riordan, O. The phase transition in the configuration mode. *Combinatorics, Probability and Computing 21*, 1-2 (2012), 265–299.

[50] Stroock, D. W., and Varadhan, S. R. S. Diffusion processes with boundary conditions. *Comm. Pure Appl. Math. 24* (1971), 147–225.

[51] van der Hofstad, R. Random graphs and complex networks. vol 2. In preparation; version of July 11, 2014 from http://www.win.tue.nl/~rhofstad/NotesRGCN.html, 2014.

[52] van der Hofstad, R. *Random graphs and complex networks. Vol. 1.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2017.

[53] Vazirani, V. A theory of alternating paths and blossoms for proving correctness of the $o(\sqrt{V}e)$ general graph maximum matching algorithm. *Combinatorica 14*, 1 (1994), 71–109.

[54] Wormald, N. C. *Some problems in the enumeration of labelled graphs.* PhD Thesis. University of Newcastle, 1978.

[55] Wormald, N. C. The asymptotic connectivity of labelled regular graphs. *Journal of Combinatorial Theory, Series B 31*, 2 (1981), 156 – 167.

[56] Wormald, N. C. The asymptotic distribution of short cycles in random regular graphs. *J. Combin. Theory Ser. B 31*, 2 (1981), 168–182.

[57] Wormald, N. C. Differential equations for random processes and random graphs. *Ann. Appl. Probab. 5*, 4 (1995), 1217–1235.