Aspects of Random Graphs

Colourings, walkers and Hamiltonian cycles

Tesi doctoral presentada al
Programa de Doctorat de Matemàtica Aplicada
de la Universitat Politècnica de Catalunya

per a optar al grau de
Doctor en Matemàtiques

per
Xavier Pérez i Giménez

sota la direcció del Doctor
Josep Díaz i Cort

Barcelona, juny de 2007

This thesis was presented at the Facultat de Matemàtiques i Estadística of the Universitat Politècnica de Catalunya on June 22th, 2007.

The dissertation committee consisted of:

- *Marc Noy Serrano* (Chair) — Departament de Matemàtica Aplicada II, Universitat Politècnica de Catalunya.
- *María José Serna Iglesias* (Secretary) — Departament de Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya.
- *Lefteris M. Kirousis* — Computer Engineering and Informatics Department, University of Patras.
- *Nicholas C. Wormald* — Department of Combinatorics and Optimization, University of Waterloo.
- *Colin McDiarmid* — Departament of Statistics, University of Oxford.

Dedico aquest treball al Marrofí.

Abstract

This dissertation presents the author's work in some problems involving different models of random graphs. First it contains a technical contribution towards solving the open problem of deciding whether with high probability a random 5-regular graph can be coloured with three colours. Next, the author proposes a model for the establishment and maintenance of communication between agents in a *mobile ad-hoc network* (MANET), which is called the *walkers model*. We assume that the agents move through a fixed environment modelled by a motion graph, and are able to communicate only if they are at a distance of at most d. As the agents move randomly, we analyse how the connectivity between a set of w agents evolves over time, asymptotically for a large number N of vertices, when w also grows large. The particular topologies of the environment which are studied here are the cycle and the toroidal grid. For the latter, the results apply under any ℓ_p-normed distance, for $1 \leq p \leq \infty$. Then, the dissertation follows with a continuous counterpart of the walkers model. Namely, it presents a model for MANETs based on random geometric graphs over the 2-dimensional unit torus, where each node moves under the *random walk* mobility model. More precisely, our model starts from a random geometric graph over the torus $[0, 1)^2$, with n nodes and radius exactly at the connectivity threshold r_c. Then each node chooses independently a random angle in $[0, 2\pi)$ and moves for a number m of steps a fixed distance $s > 0$ in that direction. After these steps, each node again chooses a new angle and starts moving in that new direction, repeating the change of direction every m steps. We compute the expected number of steps for which the resulting graph stays connected or disconnected. In addition, for static random geometric graphs with radius at the connectivity threshold r_c, we provide asymptotic expressions on the probability of existence of components according to their sizes. Finally, in the last part of this work, we show in a constructive way that, for any arbitrary ℓ_p-normed distance, $1 \leq p \leq \infty$, the property that a random geometric graph under that distance contains a Hamiltonian cycle exhibits a sharp threshold at radius $r = \sqrt{\log n/(\alpha_p n)}$, where α_p is the area of the unit disk in the ℓ_p norm.

Resum

Aquesta tesi presenta l'aportació de l'autor en alguns problemes relacionats amb diferents models de grafs aleatoris. Primer conté la contribució tècnica envers la solució del problema obert de decidir si amb alta probabilitat un graf 5-regular aleatori pot ésser acolorit amb tres colors. A continuació, l'autor proposa un model per a l'establiment i manteniment de la comunicació entre agents mòbils en una xarxa mòbil ad-hoc (MANET), anomenat el model dels *walkers*. Suposem que els agents es mouen a través d'un medi modelitzat per un graf motriu, i que són capaços de comunicar-se entre ells si són a distància com a molt d. A mesura que els agents es belluguen a l'atzar, analitzem com evoluciona en el temps la connectivitat entre un conjunt de w agents, asimptòticament per a un gran nombre N de vèrtexs, quan w també creix. Les topologies particulars del medi que estudiem aquí són el cicle i la graella toroïdal. En aquesta darrera, els resultats fan referència a qualsevol distància normada ℓ_p, amb $1 \leq p \leq \infty$. Seguidament, la tesi continua amb una variant contínua del models dels *walkers*. Concretament, es presenta un model per a MANETs basat en grafs aleatoris geomètrics sobre el torus unitat 2-dimensional, on cada node es belluga segons el model de mobilitat *random walk*. Més detalladament, el nostre model parteix d'un graf aleatori geomètric en el torus $[0, 1)^2$, amb n nodes i radi exactament en el llindar r_c de la connectivitat. Aleshores, cada node escull a l'atzar i de manera independent un angle de $[0, 2\pi)$ i es mou durant m passes una distància fixada $s > 0$ en aquella direcció. Després d'aquestes passes, tots els nodes escullen de nou un altre angle i comencen a moure's cap allà, repetint el canvi de direcció cada m passes. Es calcula el nombre esperat de passes durant les quals el graf resultant es manté connex o inconnex. A més, per als grafs aleatoris geomètrics estàtics en el llindar de la connectivitat r_c, donem expressions asimptòtiques de la probabilitat d'existència de components segons les seves talles. Finalment, en la darrera part d'aquest treball, mostrem de manera constructiva que, per a qualsevol distància normada ℓ_p arbitrària, $1 \leq p \leq \infty$, la propietat que un graf aleatori geomètric contingui un cicle Hamiltonià exhibeix un llindar abrupte en radi $r = \sqrt{\log n/(\alpha_p n)}$, on α_p és l'àrea del disc unitat en la norma ℓ_p.

Resumen

Esta tesis presenta la aportación del autor en algunos problemas relacionados con distintos modelos de grafos aleatorios. Primero contiene la contribución técnica hacia la solución del problema abierto de decidir si con alta probabilidad un grafo 5-regular aleatorio puede ser coloreado con tres colores. A continuación, el autor propone un modelo para el establecimiento y mantenimiento de la comunicación entre agentes móviles en una red móvil ad-hoc (MANET), llamado el modelo de los *walkers*. Supongamos que los agentes se mueven a través de un medio modelizado por un grafo motriz, y que son capaces de comunicarse entre ellos si están a distancia como mucho d. A medida que los agentes se mueven al azar, analizamos cómo evoluciona en el tiempo la conectividad entre un conjunto de w agentes, asintóticamente para un número grande N de vértices, cuando w también crece. Las topologías particulares del medio que estudiamos aquí son el ciclo y la malla toroidal. En ésta última, los resultados se refieren a cualquier distancia normada ℓ_p, con $1 \leq p \leq \infty$. Seguidamente, la tesis continúa con una variante continua del modelo de los *walkers*. Concretamente, se presenta un modelo para MANETs basado en los grafos aleatorios geométricos sobre el toro unidad 2-dimensional, donde cada nodo se mueve según el modelo de movilidad *random walk*. Más en detalle, nuestro modelo parte de un grafo aleatorio geométrico en el toro $[0,1)^2$, con n nodos y radio exactamente en el umbral r_c de la conectividad. Entonces, cada nodo escoge al azar y de manera independiente un ángulo de $[0, 2\pi)$ y se mueve durante $m \in \mathbb{Z}$ pasos una distancia fijada $s > 0$ en aquella dirección. Después de esos pasos, todos los nodos escogen de nuevo otro ángulo y empiezan a moverse hacia allí, repitiendo el cambio de dirección cada m pasos. Se calcula el número esperado de pasos durante los cuales el grafo resultante se mantiene conectado o desconectado. Además, para los grafos aleatorios geométricos estáticos en el umbral de la conectividad r_c, damos expresiones asintóticas de la probabilidad de existencia de componentes según sus tallas. Finalmente, en la última parte de este trabajo, mostramos de manera constructiva que, para cualquier distancia normada arbitraria ℓ_p, $1 \leq p \leq \infty$, la propiedad de que un grafo aleatorio geométrico contenga un ciclo Hamiltoniano presenta un umbral abrupto en radio $r = \sqrt{\log n/(\alpha_p n)}$, donde α_p es el área del disco unidad en la norma ℓ_p.

Acknowledgements

This dissertation could not have been completed without the generous assistance and advice of many people. For this reason, I would like to express my appreciation and gratitude. In the first place I would like to thank my thesis advisor Josep Díaz for his infinite patience and continuous encouragement. He always believed in me (more than I do myself), and devoted a great amount of energy to transmit his enthusiasm for research to me. I can positively assert that he achieved the goal. During these years, I had the privilege to collaborate with Nicholas Wormald. I have benefited enormously from working with him and have gained much insight and perspective into many probabilistic issues. I am also indebted to him and to his wife Hania for their warm hospitality in Waterloo. I also had the opportunity to work very closely with Lefteris Kirousis. I am very grateful to him for the wonderful time we spent together. During those days of intensive work, I had indeed great fun besides learning a lot; sometimes research can be as thrilling as a suspense movie και το τέλος της ιστορίας δε γράφτηκε ακόμα... During the last stages of my thesis, I had the chance to work with Dieter Mitsche. I would like to thank him for all his help and encouragement, and for carefully proofreading some parts of this work. *Danke*, Dieter! I consider all of you much more than simply co-authors, and hope this is just the beginning of a fruitful collaboration and long-lasting friendship.

Additionally, I have had the pleasure of collaborating with María José Serna, Alexios Kaporis, Giannis Grammatikopoulos, Dimitris Sotiropoulos and Graeme Kemkes. They all have my deepest appreciation. Thanks to all of them for allowing me to include joint work in this dissertation. I am particularly indebted to María José Serna for her useful advice in the earliest stages of my research. The motivating discussions we had during those days guided my first steps in this "terrible" jungle. Moreover I would like to express my gratitude towards Dimitris Achlioptas and Christopher Moore for generously sharing their ideas, which proved very useful for this research. I am also very grateful to all the people in the LSI department for providing a friendly environment and optimal working conditions. I will try to avoid listing too many names, but I want to make particular mention of Maria Josep Blesa, Xavier Molinero and Mihai Surdeanu, from my old office, and also Jordi Petit.

However, the real beginning of the story goes back 10 years, when I first began to explore the beauty of the realm of Mathematics under the encouragement of Josep Grané. I owe much of my present success to him.

Finally, I would like to have some kind words for my family, my closest friends and Marta. They have offered me warmth and support, mainly during this last year, which has been particularly hard for several reasons. *Gràcies per la vostra paciència; el Xavier de sempre ja torna a ser aquí!!!*

This work was partially supported by the Future and Emerging Technologies programme of the EU under contract 001907 *"Dynamically Evolving, Large-Scale Information Systems (DELIS)"*.

Always keep Ithaca in your mind.
To arrive there is your ultimate goal.
But do not hurry the voyage at all.
It is better to let it last for many years;
and to anchor at the island when you are old,
rich with all you have gained on the way,
not expecting that Ithaca will offer you riches.

Konstantinos P. Kavafis [1911]

Contents

Introduction

Random graphs have transcended the pure mathematical framework, to become relevant models and benchmarks in scientific disciplines such as information technology, physics and biology. Chapter 1 of this dissertation deals with the chromatic number of random d-regular graphs. As it will be explained there, this problem has given rise to beautiful research in the physics community. The chapter begins with a brief introduction on the colourability of random graphs. It follows the author's technical contribution towards the solution of the open problem of deciding whether with high probability a random 5-regular graph can be coloured with three colours. An extended abstract of the result was presented at the *European Symposium on Algorithms* (ESA'05) [22] and at the *European Conference on Complex Systems* [24]. A long version has been sent for publication[1].

Nowadays, communication networks have become an ubiquitous component of society. In particular, an increasing interest has arisen in the study of *mobile ad-hoc networks* (MANETs) and also in the theoretical aspects of *random geometric graphs*, as models for such networks. A random geometric graph results from taking n uniformly distributed points in some metric space $\mathcal{S}$ (usually the unit cube $[0,1]^d$) and connecting two points if their distance is at most r, for some prescribed radius $r \in \mathbb{R}^+$. Chapter 2 of this dissertation starts by briefly introducing MANETs, and then it surveys the main known results on the static setting of random geometric graphs. For instance, it is known that there exists a value for r denoted by $r_c = r_c(n)$ below which a random geometric is disconnected with high probability and above which it is connected with high probability. In the core of Chapter 2, we propose a model for the establishment and maintenance of communication between mobile agents in a MANET, which is called the *walkers model*. We assume that the agents move through a fixed environment modelled by a motion graph, and are able to communicate if they are at a distance of at most d. The positions of the agents over the vertices of the

[1] A preprint of this paper is available at http://www.lsi.upc.es/~xperez

motion graph determine an ad-hoc random geometric graph. As the agents move randomly, we analyse how the connectivity between a set of w agents evolves in time, asymptotically for a large number N of vertices, when w also grows large. The particular topologies of the environment which are studied here are the cycle and the toroidal grid. For the latter, the results apply under any normed ℓ_q distance, for $1 \leq q \leq \infty$. Extended abstracts of the results in this chapter, were presented at the *Symposium on Theoretical Aspects of Computer Science* (STACS'05) [26] and at the *Workshop on Efficient and Experimental Algorithms* (WEA'05) [67]. A long journal version has been sent for publication[1] (see also the technical report [27]).

Chapter 3 presents a continuous model for mobile ad-hoc networks based on random geometric graphs over the 2-dimensional unit torus, and where each node moves under the *random walk* model. More precisely, the model starts from a random geometric graph over the torus $[0,1)^2$, with n nodes and radius exactly at the connectivity threshold r_c. Then each node chooses independently a random angle in $[0,2\pi)$ and moves for a number m of steps a fixed distance $s > 0$ in that direction. After these steps, all the nodes choose again a new angle and start moving in that new direction, repeating the change of direction every m steps. We compute the expected number of steps for which the resulting graph stays connected or disconnected. Notice that this model can be regarded as a continuous counterpart of the walkers model described in Chapter 2, since the torus $[0,1)^2$ is approximated by the toroidal grid when the number of vertices in the grid is large. In fact, both cases are studied as an attempt to model ad-hoc networks. However, while in Chapter 2 the walkers move randomly along the edges of a prescribed graph, in Chapter 3 the nodes are allowed to move "freely" around the unit torus, under the random walk mobility model. The developments of both chapters have some analogies but many technical details in the arguments are quite different. In addition, in Chapter 3, for static random geometric graphs with radius at the connectivity threshold r_c, we provide asymptotic expressions on the probability of existence of components according to their sizes, which was not known before the present work. This contributes towards the understanding of the behaviour of mid-size components at the connectivity threshold. The results have been submitted for publication[1].

In chapter 4, for an arbitrary ℓ_p-normed distance, $1 \leq p \leq \infty$, we show in a constructive way that the property that a random geometric graph contains a Hamiltonian cycle exhibits a sharp threshold at radius $r = \sqrt{\log n/(\alpha_p n)}$, where α_p is the area of the unit disk in the ℓ_p norm. The result appeared in *SIAM Journal on Discrete Mathematics* [25].

Each chapter of the thesis ends with a section containing conclusions and open problems relevant to the material exposed there.

Notes on Notation and More

Given any two sequences f_n and g_n taking positive values, we say that $f_n = O(g_n)$ or equivalently that $g_n = \Omega(f_n)$ if there exists some $C > 0$ and some $n^* \in \mathbb{N}$ such that for all $n \geq n^*$ we have $f_n \leq C g_n$. If $f_n = O(g_n)$ and $f_n = \Omega(g_n)$ then we say that $f_n = \Theta(g_n)$. Moreover, we say that $f_n = o(g_n)$ or equivalently that $g_n = \omega(f_n)$ if $\lim_{n\to\infty} f_n/g_n = 0$.

We say that f_n is finite if $f_n < +\infty$, i.e. it is not identically infinity but possibly $f_n \to \infty$. We usually reserve the word bounded to describe f_n in the case that $f_n = O(1)$. This distinction is relevant in Chapters 2 and 3, where $f_n = \mathbf{E}(X_n)$ for some sequence X_n

of random variables.

Observation It is well known that a real sequence $f(n)$ converges to $f \in \mathbb{R} \cup \{-\infty, +\infty\}$ iff for any subsequence $f(n_k)$ we can find a subsubsequence $f(n_{k_m})$ which has limit f. In addition, given any $c \geq 0$ and any sequence $f(n)$ of non-negative real numbers, we can find a subsequence $f(n_k)$ with the following properties: $f(n_k)$ is either $o(1)$, $\Theta(1)$ or $\omega(1)$, and moreover either $f(n_k) \leq c$ or $f(n_k) > c$.

Hence, for each non-negative expression or parameter $f(n)$ considered hereinafter, we can assume w.l.o.g. that it is either $o(1)$, $\Theta(1)$ or $\omega(1)$, and also that given $c \geq 0$ either $f(n) \leq c$ or $f(n) > c$. Otherwise in view of the above observation, for each subsequence of $f(n)$ we can find a subsubsequence with the above-mentioned properties, use that subsubsequence in the argument and extend the results to $f(n)$. The assumption is made throughout the dissertation without explicitly mentioning it.

In all the topics covered in this work, there is involved a sequence of probability spaces indexed by n (or also N in Chapter 2), which usually denotes the size of some graph. We derive asymptotic results as n (or N) grows large. Given a sequence of events $\mathcal{E}_n$, we say that $\mathcal{E}$ holds *asymptotically almost surely* (a.a.s.) if $\mathbf{P}(\mathcal{E}) = 1 - o(1)$.

Throughout the dissertation, u.a.r will abbreviate *uniformly at random*, which means 'selected at random with uniform probability'. We denote by $1[\mathcal{E}]$ the indicator function of an event $\mathcal{E}$.

Unless otherwise specified, the base of all logarithms is assumed to be e. Moreover, we follow the convention that $0 \log 0 = 0$ and $0^0 = 1$. In view of this, functions such as $x \log x + (1-x)\log(1-x)$ and $x^x(1-x)^{1-x}$ are continuous in $[0, 1]$.

Given $k_1, \ldots, k_r \in \mathbb{N}$ with $k = k_1 + \cdots k_r$, the usual multinomial coefficient will be denoted by $\binom{k}{k_1,\ldots,k_r} = \frac{k!}{k_1! \cdots k_r!}$. Given $k, r \in \mathbb{N}$, the falling factorial is written $[k]_r = k(k-1)\cdots(k-r+1)$. In particular $\mathbf{E}[X]_r = \mathbf{E}\big(X(X-1)\cdots(X-r+1)\big)$ denotes the r^{th} factorial moment of the random variable X.

1

The Chromatic Number of Random 5-Regular Graphs

1.1 A Succinct History of Colourability of Random Graphs.

Two of the early models of random graphs are the $\mathcal{G}(n,m)$ and $\mathcal{G}(n,p)$ models due to Erdős and Rényi [30]. The $\mathcal{G}(n,p)$ model was also independently proposed by Gilbert [34]. Namely given $p \in [0,1]$, a random graph $\mathcal{G}(n,p)$ is obtained by considering a set of n labelled vertices and selecting each of the possible $\binom{n}{2}$ edges independently from each other with probability p. On the other hand given $m \in \mathbb{Z}$ $(0 \leq m \leq \binom{n}{2})$, $\mathcal{G}(n,m)$ is a random instance selected with uniform probability from the set of all graphs on n labelled vertices and m edges. Usually p and m are functions of n, and we are interested in asymptotic properties of the models for n growing large. Hereinafter, we say that an event holds *asymptotically almost surely* (a.a.s.) if it occurs with probability tending to 1 as n goes to infinity.

Two important parameters of a random graph are the *average degree* and the *edge density*. One can verify that the average degree of $\mathcal{G}(n,p=\frac{d}{n})$ (respectively $\mathcal{G}(n,m=\frac{dn}{2})$) is $p(n-1) \sim d$ (respectively $2m/n = d$). The edge density of a random graph is defined to be one half of the average degree. It is known [46] that $\mathcal{G}(n,p)$ and $\mathcal{G}(n,m)$ with the same average degree are somehow equivalent, in the sense that in many cases they have analogous properties. For sake of simplicity, in the following $\mathcal{G}(n,\frac{d}{n})$ will stand for $\mathcal{G}(n,p=\frac{d}{n})$ whenever there is no possibility of confusion.

Given a graph $G=(V,E)$ *a legal colouring* is an assignment of colours to the vertices in V in such a way that no adjacent vertices get the same colour. Given $G=(V,E)$ and $k \in \mathbb{Z}^+$, the k-colourability (k-Col) problem is to decide whether there is a legal colouring of G with at most k colours. Similarly, the *chromatic number* problem is to find which is the minimal number of colours $\chi(G)$ required to colour G legally. It is known that for general deterministic graphs, the chromatic number problem is NP-Complete [33]. Note that the two problems are intimately related, and often their study cannot be carried separately.

The first results on the k-colourability problem on random graphs were due to Erdős and Rényi.

Theorem 1.1.1 ([30]).

(i). If $d \le 1-\epsilon$, then a.a.s. all the connected components of $\mathcal{G}(n, \frac{d}{n})$ have at most one cycle and $O(\log n)$ vertices. Consequently, a.a.s. $\chi(\mathcal{G}(n, \frac{d}{n})) \le 3$.

(ii). If $d \ge 1+\epsilon$, then a.a.s. there exists a unique connected component of $\mathcal{G}(n, \frac{d}{n})$ with $\Omega(n)$ cycles (and in particular one odd cycle). Consequently, a.a.s. $\chi(\mathcal{G}(n, \frac{d}{n})) \ge 3$.

It was left open by them the question whether the chromatic number of $\mathcal{G}(n, \frac{1}{n})$ is a.a.s. 3. The answer would be given 29 years latter in the positive by Łuczak and Wierman [56].

Cheeseman, Kanefsky and Taylor [19] and also Culberson and Gent [21] observed that for each fixed k, there is a threshold average degree d_k such that if $d < d_k$, then $\mathcal{G}(n, \frac{d}{n})$ is a.a.s. k-colourable, while if $d > d_k$ then $\mathcal{G}(n, \frac{d}{n})$ is a.a.s. not k-colourable. Moreover, they experimentally obtained that $d_3 \approx 4.7$. Braunstein, Mulet, Pagnani, Weigt and Zecchina [16] gave an analytic (non-rigorous) verification that $d_3 \approx 4.69$, by using the replica method from statistical physics.

In [2] Achlioptas and Friedgut gave a formal proof that k-Col has a sharp threshold:

Theorem 1.1.2 (Achlioptas–Friedgut). *$\forall k \ge 3$ there is a sequence $d_k(n)$ such that $\forall \epsilon > 0$:*

- *A random graph with average degree $d_k(n) - \epsilon$ is a.a.s. k-colourable.*
- *A random graph with average degree $d_k(n) + \epsilon$ is a.a.s. not k-colourable.*

Notice that Theorem 1.1.2 does not imply the convergence of $d_k(n)$. Thus it remains as an open problem to prove if $d_k(n)$ converges and if so, to what value. As a corollary to Theorem 1.1.2 they obtained that for any given d, if we have that $\mathcal{G}(n, \frac{d}{n})$ is k-colourable with probability at least $\epsilon > 0$ for all n, then $\mathcal{G}(n, \frac{d}{n})$ is a.a.s. k-colourable.

The basic upper bound technique applied to the 3-colourability problem for random graphs is the following: Given a graph G, let $\mathcal{C}(G)$ the class of the legal 3-colourings of G. Then for any model of random graph $\mathcal{G}$,

$$\mathbf{P}(\mathcal{G} \in 3\text{-COL}) = \mathbf{P}(|\mathcal{C}(\mathcal{G})| \ge 1) \le \mathbf{E}(|\mathcal{C}(\mathcal{G})|).$$

Therefore, by computing $\mathbf{E}(|\mathcal{C}(\mathcal{G}(n, \frac{d}{n}))|)$, which is easy, and by finding the values of d for which $\mathbf{E}(|\mathcal{C}(\mathcal{G}(n, \frac{d}{n}))|) \to 0$, we get a trivial upper bound of 5.42 for d_3. But as we pointed out before, experiments seem to indicate the the phase transition for 3-COL occurs below at $d_3 \approx 4.69$. The reason for this discrepancy is the fact that, when $4.69 < d < 5.42$, a class of graphs with small probability but with many legal 3-colourings contributes *too much* to $\mathbf{E}(|\mathcal{C}(\mathcal{G}(n, \frac{d}{n}))|)$. In order to improve this upper bound, it is natural to consider some restricted type of legal 3-colourings with the property that whenever one graph G is 3-colourable it must admit some colouring of that restricted type. Thus the expected number of such colourings of $\mathcal{G}(n, \frac{d}{n})$ is smaller, and we obtain a more realistic bound. Following this idea, Kaporis, Kirousis and Stamatiou [49] considered *rigid* 3-colourings, i.e. colourings where flipping the colour of one vertex to a higher colour in the $0 < 1 < 2$ ordering destroys the legality of the colouring. They proved that $d_3 \le 4.99$. A different approach was introduced by Dubois, Boufkhad and Mandler [29], where they restricted the class of graphs to avoid the perturbation of having a few instances with many legal 3-colourings. In fact, it is known that $\mathcal{G}(n, \frac{d}{n})$ has a.a.s. a degree sequence close to a truncated Poisson (cf. [46]), and thus they considered the space of graphs restricted to *typical graphs*, i.e. graphs with exactly a truncated Poisson degree sequence. More precisely, for any k, the percentage of vertices with degree k is $\frac{e^{-d}d^k}{k!}$, where d is the average degree. They proved that, in $\mathcal{G}(n, \frac{d}{n})$, the few graphs which have not a truncated Poisson degree sequence have many 3-colourings, which explains the bad estimates obtained from the expectation. Dubois *et al.* proved that for the class of typical graphs, $d_3 \leq 4.854$, which gets closer to the analytical non-rigorous results of Zecchina *et al.* [16]. At the present, this is the best known upper bound for the rigorous evaluation of d_3.

A graph G has empty k-core if it does not contain any subgraph with minimum degree at least k. It is straightforward to check that any graph G which has empty k-core must be k-colourable. Using this fact, Pittel, Spencer and Wormald [69] studied the emergence of the k-core in $\mathcal{G}(n, \frac{d}{n})$ and proved that $d_3 \geq 3.35$. Then, Achlioptas and Moore [3] analysed via the *Differential Equations Method*, a non-backtracking version of the Brélaz heuristic (cf. [17]) to obtain the best lower bound for 3-COL on $G(n, d/n)$ at the present, which is $d_3 \geq 4.03$.

Let us turn back our attention to the general chromatic number problem on random graphs. I.e., given $p = p(n)$ with $0 < p < 1$, we want to find the chromatic number of $\mathcal{G}(n,p)$. Grimmett and McDiarmid [37] gave the first lower bound on $\chi(\mathcal{G}(n,p))$, by determining the range of p such that $\mathcal{G}(n,p)$ has not and independent set of size n/k (and thus $\chi(\mathcal{G}(n,p)) > k$). After a long series of papers by different authors, Bollobás [14] and independently Kučera and Matula [57] proved that if p is fixed, then

$$\chi(\mathcal{G}(n,p)) \sim \frac{n}{2\log n}.$$

If we consider the sparse case $p = d/n$ for fixed $d \in \mathbb{R}^+$, Łuczak [55] proved that for all $d \in \mathbb{R}^+$, there exists $k_d \in \mathbb{N}$ such that a.a.s. $\chi(\mathcal{G}(n, \frac{d}{n}))$ is either k_d or $k_d + 1$. However, Łuczak did not give information on these two values. Achlioptas and Naor [5], using the second moment method, proved that if k_d is defined as the smallest k such that $d < 2k \log k$, then a.a.s. $\chi(\mathcal{G}(n, \frac{d}{n})) \in \{k_d, k_d + 1\}$. Moreover, if $d \in [(2k-1)\log k, 2k \log k]$ then a.a.s. $\chi(\mathcal{G}(n, \frac{d}{n})) = k_d + 1$, which determines the exact value for about half of the d's. The second moment method will play a very important role in this chapter, therefore it is worth to look at it with more detail (See also [45] and Remark 3.1 in [46]). Let X be a non-negative random variable that depends on n. As n grows large $\mathbf{E}X$ may also grow large, but $\mathbf{P}(X > 0)$ may approach zero. However, if $\mathbf{E}(X^2)$ does not approach infinity too fast compared to the square of $\mathbf{E}X$, then it may turn out that $\mathbf{P}(X > 0)$ stays away from zero. In fact,

$$\mathbf{P}(X > 0) \geq \frac{(\mathbf{E}(X))^2}{\mathbf{E}(X^2)}. \tag{1.1}$$

So if $\mathbf{E}(X^2) = \Theta((\mathbf{E}X)^2)$ then $\Pr[X > 0]$ is bounded away from 0. Achlioptas and Naor [5] used the second moment method, with X counting the number of *balanced* k-colourings of $G(n, d/n)$ (i.e. colourings with the same number of vertices with any given colour). Then $\mathbf{E}(X^2)$ turns out to be a sum of exponential terms. They used a martingale-based concentration result to find the term with the largest base, which yields the value of the sum, and thus they proved the above-mentioned result.

An important result was obtained by using tools from physics: Krząkała, Pagnani and Weigt [54] studied the geometry of the space of solutions for k-colourings. For $G \in G(n, d/n)$, they considered the space of all k-colour assignments (legal or not) to the vertices of G. They defined a cluster as a set of colourings such that it is possible to go from one colouring to another by changing the colours of at most 2 vertices. Then, they showed that if $d < (1-\epsilon)k\log k$, all legal k-colourings form a unique cluster in this space. Therefore if we use a greedy strategy to change colours, sooner or later the algorithm is going to hit the cluster. So it is easy to solve the chromatic problem for $d < (1-\epsilon)k\log k$. Moreover, they also proved that as d increases, the clusters break down into exponentially many legal clusters and exponentially many *almost legal* clusters. Therefore, any change in the colours of a few vertices gives rise to more illegally coloured edges. As a consequence, local search algorithms are not expected to produce results for average degrees beyond the breaking down point of the unique cluster, the algorithm will get stuck on the local maximal (local colours clusters). The range $d \geq (1-\epsilon)k\log k$ is called the *hard-colourability region.*

1.1.1 Colouring Random Regular Graphs

So far we have been discussing about the Erdős and Rényi model of random graphs. Let us turn our attention to another type of random graphs, which are the centre of this chapter. A *random d-regular graph* $\mathcal{G}(n,d)$ is a random instance selected with uniform probability from the class $\mathbb{G}(n,d)$ of d-regular graphs on n labelled vertices. In order to study this model, it is usually used the well-known *pairing* or *configuration model* $\mathcal{P}(n,d)$, which was first introduced by Bollobás [13]. A *d-pairing* is a perfect matching on a set of dn points which are grouped into n cells of d points each. Let $\mathbb{P}(n,d)$ be the set of d-pairings of dn points, and denote by $\mathcal{P}(n,d)$ a random element of $\mathbb{P}(n,d)$ selected u.a.r. A random pairing $\mathcal{P}(n,d)$ corresponds in a natural way to a random d-regular multigraph (possibly containing loops or multiple edges), in which each cell becomes a vertex. The reader should refer to [78] for further aspects of the pairing model such as the following well know result.

Theorem 1.1.3. *If a property holds a.a.s. for $\mathcal{P}(n,d)$, then it also holds a.a.s. for $\mathcal{G}(n,d)$.*

From the early 90's the there has been a considerable effort in studying the chromatic number of a random regular graph. In the present subsection, we sketch the highlights of the developments in the field. Frieze and Łuczak [31], proved that $\chi\big(\mathcal{G}(n,d)\big) \sim d/(2\log d)$ a.a.s. (as $d \to \infty$). However, they did not give any result for fixed values of d. Molloy and Reed [61] proved that if $k(1-1/k)^{d/2} < 1$ then $\mathcal{G}(n,d)$ is not k-colourable a.a.s. Notice that this result implies that $\forall d \geq 6$, a.a.s. $\mathcal{G}(n,d)$ is not 3-colourable. The result was proved by a clever use of the first moment method to show that the expected number of k-colourings of a d-regular pairing is at most $k(1-1/k)^{d/2}$. Achlioptas and Moore [3] showed that $\chi\big(\mathcal{G}(n,4)\big) = 3$ with probability bounded away from 0. The proof was based on the analysis of a backtracking-free version of Brélaz heuristic, together with the fact that $\mathcal{G}(n,4)$ is a.a.s. not bipartite. Later, they showed in [4] that the chromatic number of a random regular graph of degree d ranges a.a.s. in $\{k_d, k_d+1, k_d+2\}$, where k_d is the smallest integer k such that $d < 2k\log k$. Shi and Wormald [74, 75] proved that: a.a.s. $\chi\big(\mathcal{G}(n,4)\big) = 3$; a.a.s. $\chi\big(\mathcal{G}(n,5)\big) \in \{3,4\}$; a.a.s. $\chi\big(\mathcal{G}(n,6)\big) = 4$; if $7 \leq d \leq 9$ then a.a.s. $\chi\big(\mathcal{G}(n,d)\big) \in \{4,5\}$; and a.a.s. $\chi\big(\mathcal{G}(n,10)\big) \in \{5,6\}$. The proof of the previous result was algorithmic: Properly colour all short cycles of $\mathcal{G}(n,d)$ up to some length l. Then, greedily colour the remaining vertices following some priority given by an ordering of the possible labels. At any moment, each vertex is labelled (i,j), where i is the number of non-coloured neighbours and j is the number of available colours.

Recall the above-mentioned work by Krząkała, Pagnani and Weigt [54], where they defined the hard colouring region. In that paper they also proved that the solution space of 3-colourings of 5-regular graphs has many clusters. Therefore, the 3-colouring of 5-regular graphs is in the hard colouring region, so it seems difficult that classical algorithmic techniques will work to find out if a 5-regular graph is 3-colourable.

In the same paper, they also observed by using *Survey propagation* techniques that almost all 5-regular graphs seem to have chromatic number 3. Therefore regarding the the 3-colourability of d-regular graphs, the situation was: 4-regular graphs are a.a.s. 3-colourable; 6-regular are a.a.s. not 3-colourable; the problem of deciding whether a 5-regular is 3-colourable falls in the hard colouring region, and the physicists gave non-rigorous evidence that most of the 5-regular are 3-colourable.

1.2 Preliminaries

In this chapter we study the locally rainbow balanced 3-colourings of a 5-regular graph, where a colouring is *balanced* if the number of vertices of each colour is equal, and *locally rainbow* if every vertex is adjacent to vertices of all the other colours. We show that a 5-regular graph admits such a colouring with probability bounded away from 0. In Section 1.8 we briefly describe how this probability can be raised to $1-o(1)$. The proof contains one non-rigorous step which is sustained by empirical evidence (see Maximum Hypothesis below). Our results are asymptotic with respect to the number of vertices n, which is restricted to the multiples of 6. (The number of vertices of a 5-regular graph must be even, and having a balanced colouring requires n to be also divisible by 3.)

The main technique in the argument is the second moment method (see (1.1)): Indeed, we compute the expectation $\mathbf{E}X$ and second moment $\mathbf{E}(X^2)$ of the number X of locally rainbow balanced colourings, asymptotically. Then, assuming that a certain four-variable function has a unique maximum at a given point in a bounded domain (see again Maximum Hypothesis), we prove that $\mathbf{E}(X^2)$ is asymptotically a constant times $(\mathbf{E}X)^2$.

For our calculations, we use the previously described *pairing model* $\mathcal{P}(n,d)$. A colouring of a paring is an assignment of colours to the cells such that it defines a colouring of the corresponding multigraph. Moreover, each point of the pairing inherits the colour of the cell containing it.

The estimation of the second moment amounts essentially to counting the number of *pairs* of locally rainbow balanced colourings on pairings in $\mathbb{P}(n,5)$. To give an exact expression for $\mathbf{E}(X^2)$ we have to sum over a large number of variables (9×36). These variables express the number of cells that have a given pair of colours (out of the nine possible pairs) and also have a given distribution of their five points with respect to the pair of colours on the cells of the points to which they are matched. As we will see there are 36 possible distributions. The computation of the asymptotic value of this expression (even within a polynomial factor) entails the computation of the global maximum of a function of

9×36 variables. In Section 1.5 we show how to reduce this computation to the computation of the maximum of a four-variable continuous function F defined over a closed and bounded convex domain. As the definitions of F and its domain are technically involved, we postpone presenting them until Section 1.5, at which point the motivation behind the technicalities becomes clearer. For the sake of easy reference, in Section 1.7 we repeat these definitions, and also give an equivalent definition of F.

Regarding the maximisation of F, we show that the boundary of its domain contains no local maximiser and that there is a local maximum at the interior point $(1/9, 1/9, 1/9, 1/9)$, by showing that the Hessian of $\log F$ is negative definite at this point. Furthermore, by numerically computing the values of F over a fine grid of its domain we obtain strong numerical evidence that the point $(1/9, 1/9, 1/9, 1/9)$ is actually the unique maximiser of F. We state the assumption corroborated by this evidence as:

Maximum Hypothesis. *The four-variable function $F(\boldsymbol{n})$ has a unique global maximum over its domain at the point* $(1/9, 1/9, 1/9, 1/9)$.

Under the Maximum Hypothesis, we can establish the chromatic number of a positive fraction of the random 5-regular graphs.

Theorem 1.2.1. *Under the Maximum Hypothesis, for n divisible by 6 the chromatic number of $\mathcal{G}(n, 5)$ is 3 with probability bounded away from 0.*

In the remaining of this chapter we prove Theorem 1.2.1. In Section 1.3, we develop an exact expression for the first and second moments of X. The asymptotic value of $\mathbf{E}(X)$ is determined in Section 1.4. In Section 1.5, we compute the asymptotic value of $\mathbf{E}(X^2)$, under the Maximum Hypothesis. The proof of Theorem 1.2.1 is completed in Section 1.6, where the previously obtained results on pairings are transferred to simple graphs. In Section 1.7 we present the empirical validation of the Maximum Hypothesis. Finally, Section 1.8 contains a brief description on how to override the restriction of n to multiples of 6 and also extend our result to a.a.s.

1.3 Exact Expression for the Moments

Given a pairing $P \in \mathbb{P}(n, 5)$, let $\mathcal{R}_P$ be the class of locally rainbow balanced 3-colourings of P. Let $X = \left|\mathcal{R}_{\mathcal{P}(n,5)}\right|$ be the random variable that counts the number of locally rainbow balanced 3-colourings in $\mathcal{P}(n, 5)$. Then,

$$\mathbf{E}X = \frac{\left|\{(P, C) \ : \ P \in \mathbb{P}(n, 5), \ C \in \mathcal{R}_P\}\right|}{\left|\mathbb{P}(n, 5)\right|}, \tag{1.2}$$

$$\mathbf{E}(X^2) = \frac{\left|\{(P, C_1, C_2) \ : \ P \in \mathbb{P}(n, 5), \ C_1, C_2 \in \mathcal{R}_P\}\right|}{\left|\mathbb{P}(n, 5)\right|}, \tag{1.3}$$

where $|\mathbb{P}(n, 5)| = (5n)!/(2^{5n/2}(5n/2)!)$.

1.3.1 First Moment

Below we assume that we are given a pairing P and a locally rainbow balanced 3-colouring C on P. Recall that a pairing is a perfect matching on $5n$ points which are organised into

n cells of 5 points each. Here and throughout the chapter, the cells get colours 0, 1 and 2, and the arithmetic in the colours is modulo 3, so we can regard colours as elements in $\mathbb{Z}_3$.

Let v be a cell of colour $i \in \mathbb{Z}_3$. The 1-*spectrum* of cell v is an ordered pair of non-negative integers. Cell v is said to have 1-spectrum $s = (s_{-1}, s_1)$ if s_r out of its five points, $r \in \{-1, 1\}$, are matched to points in cells of colour $i + r$. The entries of s are non-negative because C is locally rainbow, and their sum is 5 because of the 5-regularity of the pairing. One can check that there are four possible 1-spectra, namely $(1, 4)$, $(2, 3)$, $(3, 2)$ and $(4, 1)$. We let $\mathcal{S}_1$ denote the set of all 1-spectra.

For each $i \in \mathbb{Z}_3$ and 1-spectrum $s \in \mathcal{S}_1$, we denote by d_s^i the scaled with respect to n number of cells of P which have colour i and 1-spectrum s. Then,

$$\sum_{s \in \mathcal{S}_1} d_s^i = \frac{1}{3}, \quad \forall i \in \mathbb{Z}_3 \tag{1.4}$$

and therefore $\sum_{i,s} d_s^i = 1$.

Given any two colours i and j in $\mathbb{Z}_3$, observe that there are exactly $5n/6$ pairs of points in cells of colours i and j respectively. Hence,

$$\sum_{s \in \mathcal{S}_1} s_r d_s^i = \frac{5}{6}, \quad \forall i \in \mathbb{Z}_3,\ \forall r \in \{-1, 1\}, \tag{1.5}$$

which also implies (1.4).

We consider the 6-dimensional polytope

$$\mathcal{D}_1 = \left\{ (d_s^i)_{i \in \mathbb{Z}_3, s \in \mathcal{S}_1} \in \mathbb{R}^{12} \ :\ d_s^i \geq 0\ \forall i, s,\ \sum_s s_r d_s^i = \frac{5}{6}\ \forall i, r \right\},$$

and the discrete subset

$$\mathcal{I}_1 = \mathcal{D}_1 \cap \left(\frac{1}{n} \mathbb{Z}^{12} \right).$$

We observe that $\mathcal{I}_1$ contains all the sequences $(d_s^i)_{i \in \mathbb{Z}_3, s \in \mathcal{S}_1}$ that correspond to some locally rainbow balanced 3-colouring, since they all satisfy (1.5). Given a fixed sequence $(d_s^i) \in \mathcal{I}_1$, let us denote by $\binom{n}{(d_s^i n)}$ the multinomial coefficient that counts the number of ways to distribute the n vertices into classes of cardinality $d_s^i n$ for all possible values of i and s. Let $\binom{5}{s}$ stand for $5!/(s_{-1}! s_1!)$.

By counting the ways to assign 1-spectra to cells, and then colours to points in cells given their 1-spectra, and finally the number of matchings between colour classes, we have

$$|\{(P, C) \ :\ P \in \mathbb{P}(n, 5),\ C \in \mathcal{R}_P\}| = \sum_{(d_s^i) \in \mathcal{I}_1} \left\{ \binom{n}{(d_s^i n)} \left(\prod_{i,s} \binom{5}{s}^{d_s^i n} \right) \left(\frac{5n}{6}! \right)^3 \right\}.$$

In view of (1.2), we divide this by $|\mathbb{P}(n, 5)|$ and obtain

$$\mathbf{E}(X) = \frac{2^{5n/2}(5n/2)!}{(5n)!} \sum_{(d_s^i) \in \mathcal{I}_1} \left\{ \binom{n}{(d_s^i n)} \left(\prod_{i,s} \binom{5}{s}^{d_s^i n} \right) \left(\frac{5n}{6}! \right)^3 \right\}. \tag{1.6}$$

1.3.2 Second Moment

Below we assume we are given a pairing P and two locally rainbow balanced 3-colourings C_1 and C_2 on P. For $i,j \in \mathbb{Z}_3$, let $V^{i,j}$ be the set of cells coloured with i and j with respect to colourings C_1 and C_2, respectively. Let $n^{i,j} = |V^{i,j}|/n$ and let $E^{i,j}$ be the set of points in cells of $V^{i,j}$. Since C_1 and C_2 are balanced, we have

$$\sum_{i\in\mathbb{Z}_3} n^{i,j} = 1/3, \quad \forall j \in \mathbb{Z}_3, \qquad \sum_{j\in\mathbb{Z}_3} n^{i,j} = 1/3, \quad \forall i \in \mathbb{Z}_3, \tag{1.7}$$

and therefore $\sum_{i,j} n^{i,j} = 1$.

Also, for $r,t \in \{-1,1\}$, let $E^{i,j}_{r,t}$ be the set of points in $E^{i,j}$ which are matched with points in $E^{i+r,j+t}$. Recall that the arithmetic in the indices is modulo 3. Let $m^{i,j}_{r,t} = |E^{i,j}_{r,t}|/n$. For fixed i and j in $\mathbb{Z}_3$, it is convenient to think of the four variables $(m^{i,j}_{r,t})_{r,t\in\{-1,1\}}$ as the entries of a 2×2 matrix $m^{i,j}$. The rows and columns are indexed by -1 and 1, with -1 for the first row or column.

$$m^{i,j} = \begin{bmatrix} m^{i,j}_{-1,-1} & m^{i,j}_{-1,1} \\ m^{i,j}_{1,-1} & m^{i,j}_{1,1} \end{bmatrix}.$$

We have that $\sum_{r,t} m^{i,j}_{r,t} = 5n^{i,j}$, and therefore $\sum_{i,j,r,t} m^{i,j}_{r,t} = 5$. And, since matching sets of points should have equal cardinalities, we also have that

$$m^{i,j}_{r,t} = m^{i+r,j+t}_{-r,-t}. \tag{1.8}$$

Let v be a cell in $V^{i,j}$. The 2-*spectrum* s of cell v is a 2×2 non-negative integer matrix. The rows and columns are indexed by -1 and 1, with -1 for the first row or column. Cell v is said to have 2-spectrum s if $s_{r,t}$ out of its five points, $r,t \in \{-1,1\}$, are matched to points in cells of $V^{i+r,j+t}$. The sum of the entries of s is 5 because of the 5-regularity of the pairing. Each row and column sum is at least 1 because both C_1 and C_2 are locally rainbow. We let $\mathcal{S}_2$ denote the set of possible 2-spectra. See that $|\mathcal{S}_2|$ has 36 elements, namely

$$\begin{aligned}\mathcal{S}_2 = \Bigg\{ &\begin{bmatrix}1&0\\0&4\end{bmatrix}, \begin{bmatrix}0&1\\4&0\end{bmatrix}, \begin{bmatrix}4&0\\0&1\end{bmatrix}, \begin{bmatrix}0&4\\1&0\end{bmatrix}, \begin{bmatrix}2&0\\0&3\end{bmatrix}, \begin{bmatrix}0&2\\3&0\end{bmatrix},\\
&\begin{bmatrix}3&0\\0&2\end{bmatrix}, \begin{bmatrix}0&3\\2&0\end{bmatrix}, \begin{bmatrix}0&1\\1&3\end{bmatrix}, \begin{bmatrix}1&0\\3&1\end{bmatrix}, \begin{bmatrix}3&1\\1&0\end{bmatrix}, \begin{bmatrix}1&3\\0&1\end{bmatrix},\\
&\begin{bmatrix}0&1\\3&1\end{bmatrix}, \begin{bmatrix}1&0\\1&3\end{bmatrix}, \begin{bmatrix}3&1\\0&1\end{bmatrix}, \begin{bmatrix}1&3\\1&0\end{bmatrix}, \begin{bmatrix}1&1\\0&3\end{bmatrix}, \begin{bmatrix}3&0\\1&1\end{bmatrix},\\
&\begin{bmatrix}1&1\\3&0\end{bmatrix}, \begin{bmatrix}0&3\\1&1\end{bmatrix}, \begin{bmatrix}1&1\\1&2\end{bmatrix}, \begin{bmatrix}1&1\\2&1\end{bmatrix}, \begin{bmatrix}2&1\\1&1\end{bmatrix}, \begin{bmatrix}1&2\\1&1\end{bmatrix},\\
&\begin{bmatrix}0&2\\2&1\end{bmatrix}, \begin{bmatrix}2&0\\1&2\end{bmatrix}, \begin{bmatrix}1&2\\2&0\end{bmatrix}, \begin{bmatrix}2&1\\0&2\end{bmatrix}, \begin{bmatrix}0&2\\1&2\end{bmatrix}, \begin{bmatrix}2&0\\2&1\end{bmatrix},\\
&\begin{bmatrix}1&2\\0&2\end{bmatrix}, \begin{bmatrix}2&1\\2&0\end{bmatrix}, \begin{bmatrix}2&2\\0&1\end{bmatrix}, \begin{bmatrix}1&0\\2&2\end{bmatrix}, \begin{bmatrix}2&2\\1&0\end{bmatrix}, \begin{bmatrix}0&1\\2&2\end{bmatrix}\Bigg\}.\end{aligned} \tag{1.9}$$

For each $i, j \in \mathbb{Z}_3$ and 2-spectrum $s \in \mathcal{S}$, we denote by $d_s^{i,j}$ the scaled, with respect to n, number of cells which belong to $V^{i,j}$ and have spectrum s. We have

$$m^{i,j} = \sum_{s \in \mathcal{S}_2} d_s^{i,j}\, s, \tag{1.10}$$

$$n^{i,j} = \sum_{s \in \mathcal{S}_2} d_s^{i,j}, \tag{1.11}$$

and therefore $\sum_{i,j,s} d_s^{i,j} = 1$.

We consider the 301-dimensional polytope

$$\mathcal{D}_2 = \left\{ (d_s^{i,j})_{i,j \in \mathbb{Z}_3, s \in \mathcal{S}} \in \mathbb{R}^{324} \;:\; d_s^{i,j} \geq 0\ \forall i,j,s,\ \sum_{j,s} d_s^{i,j} = \frac{1}{3}\ \forall i, \right.$$
$$\left. \sum_{i,s} d_s^{i,j} = \frac{1}{3}\ \forall j,\ \sum_s s_{r,t} d_s^{i,j} = \sum_s s_{-r,-t} d_s^{i+r,j+t}\ \forall i,j,r,t \right\},$$

and the discrete subset

$$\mathcal{I}_2 = \mathcal{D}_2 \cap \left(\frac{1}{n} \mathbb{Z}^{324} \right).$$

In view of (1.7)–(1.11), note that $\mathcal{I}_2$ contains the set of sequences $(d_s^{i,j})_{i,j \in \mathbb{Z}_3, s \in \mathcal{S}_2}$ that correspond to some pair of locally rainbow balanced 3-colourings. Given a fixed sequence $(d_s^{i,j}) \in \mathcal{I}_2$, let us denote by $\binom{n}{(d_s^{i,j} n)}$ the multinomial coefficient that counts the number of ways to distribute the n vertices into classes of cardinality $d_s^{i,j} n$ for all possible values of i, j and s. Define $m^{i,j}$ by (1.10). Also let $\binom{5}{s}$ stand for $5!/\prod_{r,t} s_{r,t}!$.

By counting the ways to assign 2-spectra to cells, the ways to assign colours to points in cells given their spectra, and finally the number of matchings between colour classes, we have

$$|\{(P, C_1, C_2) \;:\; P \in \mathbb{P}(n,5),\ C_1, C_2 \in \mathcal{R}_P\}| =$$
$$\sum_{(d_s^{i,j}) \in \mathcal{I}_2} \left\{ \binom{n}{(d_s^{i,j} n)} \left(\prod_{i,j,s} \binom{5}{s}^{d_s^{i,j} n} \right) \left(\prod_{i,j,r,t} \left((m_{r,t}^{i,j} n)! \right)^{1/2} \right) \right\}.$$

In view of (1.3), we divide this by $|\mathbb{P}(n,5)|$ and obtain

$$\mathbf{E}(X^2) =$$
$$\frac{2^{5n/2}(5n/2)!}{(5n)!} \sum_{(d_s^{i,j}) \in \mathcal{I}_2} \left\{ \binom{n}{(d_s^{i,j} n)} \left(\prod_{i,j,s} \binom{5}{s}^{d_s^{i,j} n} \right) \left(\prod_{i,j,r,t} \left((m_{r,t}^{i,j} n)! \right)^{1/2} \right) \right\}. \tag{1.12}$$

1.4 Asymptotic Value of the First Moment

For sake of simplicity, here and throughout the section we will often write $\boldsymbol{d}$ to denote the tuple $(d_s^i)_{i \in \mathbb{Z}_3, s \in \mathcal{S}_1}$. Let us consider the functions

$$f_1(\boldsymbol{d}) = \sum_{i,s} d_s^i (\log \binom{5}{s} - \log d_s^i), \qquad g_1(\boldsymbol{d}) = \sqrt{\frac{1}{\prod_{i,s} d_s^i}},$$

$$h_1(n) = \frac{5^{3/2}}{2^2 3^{3/2}} \; (2\pi n)^{-4} \; 6^{-5n/2}, \tag{1.13}$$

where f_1 is defined over $\mathcal{D}_1$ and g_1 is defined over the interior of $\mathcal{D}_1$. Recall that we follow the convention that $0 \log 0 = 0$.

Lemma 1.4.1. *The first moment satisfies*

$$\mathbf{E}(X) = h_1(n) \sum_{\boldsymbol{d} \in \mathcal{I}_1} q_1(n, \boldsymbol{d}) e^{f_1(\boldsymbol{d})n} \tag{1.14}$$

where, as $n \to \infty$ and uniformly over all $\boldsymbol{d}$, $q_1(n, \boldsymbol{d}) = O(n^6)$ and $q_1(n, \boldsymbol{d}) \sim g_1(\boldsymbol{d})$ provided all d_s^i are bounded away from 0.

Proof. We apply Stirling's formula and perform simple manipulations to (1.6) to obtain:

$$\begin{aligned}
\mathbf{E}(X) &= \frac{2^{5n/2}(5n/2)!n!(5n/6)!^3}{(5n)!} \sum_{\boldsymbol{d} \in \mathcal{I}_1} \left(\prod_{i,s} \frac{\binom{5}{s}^{d_s^i n}}{(d_s^i n)!} \right) \\
&\sim \sqrt{\pi n \; (5\pi n/3)^3} \; 6^{-5n/2} (n/e)^n \sum_{\boldsymbol{d} \in \mathcal{I}_1} \left(\prod_{i,s} \frac{\binom{5}{s}^{d_s^i n}}{(d_s^i n)!} \right) \\
&= h_1(n) \sum_{\boldsymbol{d} \in \mathcal{I}_1} \left((2\pi n)^6 (n/e)^n \prod_{i,s} \frac{\binom{5}{s}^{d_s^i n}}{(d_s^i n)!} \right).
\end{aligned} \tag{1.15}$$

We need to uniformly approximate the factorial of several numbers not necessarily growing large with n. Stirling's formula also implies $k! = \sqrt{2\pi \eta(k)} (k/e)^k$ for all $k \geq 0$, where $\eta(k) \sim k$ if $k \to \infty$, and $\eta(k) = \Theta(k+1)$ for all $k \geq 0$. In particular, η is non-zero. So we have

$$\begin{aligned}
\prod_{i,s} \frac{\binom{5}{s}^{d_s^i n}}{(d_s^i n)!} &= \prod_{i,s} \frac{\binom{5}{s}^{d_s^i n}}{\sqrt{2\pi \eta(d_s^i n)} (d_s^i n/e)^{d_s^i n}} \\
&= \frac{1}{(2\pi n)^6 (n/e)^n} \prod_{i,s} \frac{\left(\binom{5}{s}/d_s^i\right)^{d_s^i n}}{\sqrt{\eta(d_s^i n)/n}} \\
&= \frac{1}{(2\pi n)^6 (n/e)^n} q_1(n, \boldsymbol{d}) e^{f_1(\boldsymbol{d})n},
\end{aligned}$$

for a function q_1 of the type described in the statement of the lemma. Combining this with (1.15) yields the lemma. □

We consider the maximum base of the exponential part of the terms in (1.14), taken over all points in the polytope $\mathcal{D}_1$:

$$M_1 = \max_{\boldsymbol{d} \in \mathcal{D}_1} \left\{ 6^{-5/2} e^{f_1(\boldsymbol{d})} \right\}.$$

This is well defined, due to the compactness of the domain and the continuity of the expression. Note that the exponential behaviour of the first moment is governed by M_1 since the number of terms in the sum in (1.14) is polynomial with respect to n.

In the next subsection we determine the value of M_1. In the following subsection, based on that result and using a Laplace-type integration argument, we compute the subexponential factors in the asymptotic expression of the first moment.

1.4.1 Computing M_1

Let $\boldsymbol{b} = (b_s^i)_{i\in\mathbb{Z}_3, s\in\mathcal{S}_1}$ be the point in $\mathcal{D}_1$ where $b_s^i = \frac{\binom{5}{s}}{90}$ for all i, s. The following lemma follows from the application of elementary analysis techniques and the computation of Lagrange multipliers.

Lemma 1.4.2. *The function f_1 is strictly concave and has a unique maximum in $\mathcal{D}_1$ at $\boldsymbol{b}$.*

Proof. We will maximise f_1 over the larger domain $\mathcal{R} \supset \mathcal{D}_1$ of all non-negative tuples $\boldsymbol{d}$ such that

$$\sum_{i\in\mathbb{Z}_3, s\in\mathcal{S}_1} d_s^i = 1. \tag{1.16}$$

We temporarily relax the constraint (1.16) and observe that the Hessian of f_1 is negative definite for any tuple of positive d_s^i. Therefore f_1 is concave in this domain. Then f_1 is also concave in $\mathcal{R}$ and in $\mathcal{D}_1$, since linear constraints do not affect concavity.

We use the Lagrange multipliers method to find stationary points in the interior of R. We obtain the following equations

$$\frac{\partial f_1}{\partial d_s^i} = \log \binom{5}{s} - 1 - \log d_s^i = \lambda, \quad \forall i \in \mathbb{Z}_3,\ \forall s \in \mathcal{S}_1,$$

where λ is the Lagrange multiplier introduced by the constraint (1.16). There is a unique solution at the point

$$d_s^i = \frac{\binom{5}{s}}{e^{\lambda+1}}, \quad \forall i \in \mathbb{Z}_3,\ \forall s \in \mathcal{S}_1,$$

which must be a maximum by concavity of the function. We observe that this maximiser also belongs to $\mathcal{D}_1$ and the statement follows. $\square$

By direct substitution, we obtain:

Lemma 1.4.3. $M_1 = 6^{-5/2}\left(\prod_{i,s} 90^{d(i,s)}\right) = \left(\frac{1}{6}\right)^{5/2} 90 = \sqrt{\frac{25}{24}}$.

From the above and from (1.14), we get:

Theorem 1.4.4. *The expected number of locally rainbow balanced 3-colourings of a 5-regular pairing approaches infinity as n grows large.*

1.4.2 Subexponential Factors of $\mathbf{E}(X)$

Here we complete the computation of the asymptotic expression of $\mathbf{E}(X)$ by using a Laplace-type integration technique described in the following result:

Lemma 1.4.5. *Let $\mathcal{A} \subseteq \mathbb{R}^d$ be a compact set with non-empty interior. Let $g \in \mathcal{C}^1(\mathcal{A})$ and $f \in \mathcal{C}^3(\mathcal{A})$ be real functions. Suppose that f has a unique maximum in $\mathcal{A}$, at the interior point $\boldsymbol{x}_0$, and that the Hessian H of f at $\boldsymbol{x}_0$ is definite negative. Suppose furthermore that $g(\boldsymbol{x}_0) \neq 0$. Then, we have as n grows large*

$$I := \int_{\mathcal{A}} g(\boldsymbol{x}) e^{f(\boldsymbol{x})n} d\boldsymbol{x} \sim \frac{1}{\sqrt{|\det H|}} \left(\frac{2\pi}{n}\right)^{d/2} g(\boldsymbol{x}_0) e^{f(\boldsymbol{x}_0)n}.$$

Proof. Since H is definite negative, there exist an invertible matrix Q such that

$$H = Q^{\perp}(-Id)Q,$$

and hence

$$\det H = (-1)^d (\det Q)^2.$$

Define the following change of variables:

$$\boldsymbol{y} = \phi(\boldsymbol{x}) = Q(\boldsymbol{x} - \boldsymbol{x}_0).$$

Let us consider the functions $\hat{g} = g \circ \phi^{-1}$ and $\hat{f} = f \circ \phi^{-1} - f(\boldsymbol{x}_0)$ defined in $\hat{\mathcal{A}} = \phi(\mathcal{A})$. Notice that $\hat{\mathcal{A}}$ is compact with non-empty interior, that $\hat{f}$ has a unique global maximum at $\bar{0}$, which is an interior point of $\hat{\mathcal{A}}$, that $\hat{f}(\bar{0}) = 0$ but $\hat{g}(\bar{0}) \neq 0$, and that $\left(\frac{\partial x_i}{\partial y_j}\right) = Q^{-1}$. Then, after performing a change of variables, we obtain

$$I = \frac{1}{|\det Q|} \int_{\hat{\mathcal{A}}} \hat{g}(\boldsymbol{y}) e^{\left(\hat{f}(\bar{y}) + f(\boldsymbol{x}_0)\right)n} d\boldsymbol{y} = \frac{e^{f(\boldsymbol{x}_0)n}}{\sqrt{|\det H|}} \int_{\hat{\mathcal{A}}} \hat{g}(\boldsymbol{y}) e^{\hat{f}(\boldsymbol{y})n} d\boldsymbol{y}.$$

Let us call

$$I_0 = \int_{\hat{\mathcal{A}}} \hat{g}(\boldsymbol{y}) e^{\hat{f}(\boldsymbol{y})n} d\boldsymbol{y}.$$

Taking into account that $\hat{\mathcal{A}}$ is compact, that $\hat{f}$ is continuous and that the unique maximum occurs at the interior point $\bar{0}$, we can claim that there exists some $\alpha > 0$ such that, for small enough $\epsilon > 0$, we can assure that $[-\epsilon, \epsilon]^d \subseteq \hat{\mathcal{A}}$ and $\hat{f}(\boldsymbol{y}) \leq -\alpha$ in $\hat{\mathcal{A}} \backslash [-\epsilon, \epsilon]^d$. Then, we can write

$$I_0 = \int_{[-\epsilon,\epsilon]^d} \hat{g}(\boldsymbol{y}) e^{\hat{f}(\boldsymbol{y})n} d\boldsymbol{y} + \int_{\hat{\mathcal{A}} \backslash [-\epsilon,\epsilon]^d} \hat{g}(\boldsymbol{y}) e^{\hat{f}(\boldsymbol{y})n} d\boldsymbol{y},$$

and call

$$I_1 = \int_{\hat{\mathcal{A}} \backslash [-\epsilon,\epsilon]^d} \hat{g}(\boldsymbol{y}) e^{\hat{f}(\boldsymbol{y})n} d\boldsymbol{y}, \qquad I_2 = \int_{[-\epsilon,\epsilon]^d} \hat{g}(\boldsymbol{y}) e^{\hat{f}(\boldsymbol{y})n} d\boldsymbol{y}.$$

Since $\hat{\mathcal{A}}$ is compact, we can write $\int_{\hat{\mathcal{A}}} |\hat{g}(\bar{y})| \, d\boldsymbol{y} \leq \max_{\hat{\mathcal{A}}} \{|\hat{g}(\boldsymbol{y})|\} \int_{\hat{\mathcal{A}}} d\boldsymbol{y} \leq K$, and then

$$|I_1| \leq \int_{\hat{\mathcal{A}} \backslash [-\epsilon,\epsilon]^d} \left|\hat{g}(\boldsymbol{y}) e^{-\alpha n}\right| d\boldsymbol{y} \leq e^{-\alpha n} \int_{\hat{\mathcal{A}}} |\hat{g}(\boldsymbol{y})| \, d\boldsymbol{y} \leq e^{-\alpha n} K. \tag{1.17}$$

Observe that $\hat{g} \in \mathcal{C}^1(\hat{\mathcal{A}})$ and $\hat{f} \in \mathcal{C}^3(\hat{\mathcal{A}})$. Then we can find positive constants C and M such that, for all $\boldsymbol{y}$ in a neighbourhood of $\bar{0}$, we have

$$\left|\hat{g}(\boldsymbol{y}) - \hat{g}(\bar{0})\right| \leq \frac{C}{d} \Big| \sum_i y_i \Big|, \quad \text{and} \quad \Big|\hat{f}(\boldsymbol{y}) + \frac{1}{2} \sum_i y_i^2 \Big| \leq \frac{M}{2d^3} \Big| \sum_{i,j,k} y_i y_j y_k \Big|.$$

Then, for any $\boldsymbol{y} \in [-\epsilon, \epsilon]^d$ and assuming that $|y_m| \geq |y_i| \ \forall i$, we can write

$$\left|\hat{g}(\boldsymbol{y}) - \hat{g}(\bar{0})\right| \leq \frac{C}{d}\Big|\sum_i y_i\Big| \leq \frac{C}{d}\sum_i \Big|y_i\Big| \leq C\epsilon,$$

and

$$\left|\hat{f}(\boldsymbol{y}) + \frac{1}{2}\sum_i y_i^2\right| \leq \frac{M}{2d^3}\Big|\sum_{i,j,k} y_i y_j y_k\Big| \leq \frac{M}{2d^3}\sum_{i,j,k}\Big|y_i y_j y_k\Big| \leq \frac{M}{2}\epsilon y_m^2 \leq M\epsilon\frac{1}{2}\sum_i y_i^2.$$

Now we obtain upper and lower bounds for I_2. We use the following well known identity:

$$\int_{\mathbb{R}} e^{-t^2}\, dt = \sqrt{\pi}.$$

$$\begin{aligned}
I_2 &\leq \int_{[-\epsilon,\epsilon]^d} \left(\hat{g}(\bar{0}) + C\epsilon\right) e^{-\frac{1}{2}(1-M\epsilon)n\sum_i y_i^2} d\boldsymbol{y} \\
&= \left(\frac{1}{2}(1-M\epsilon)n\right)^{-d/2} \left(\hat{g}(\bar{0}) + C\epsilon\right) \int_{\left[-\sqrt{\frac{1}{2}(1-M\epsilon)n}\ \epsilon, \sqrt{\frac{1}{2}(1-M\epsilon)n}\ \epsilon\right]^d} e^{-\sum_j z_j^2} d\bar{z} \\
&\sim \left(\frac{1}{2}(1-M\epsilon)n\right)^{-d/2} \left(\hat{g}(\bar{0}) + C\epsilon\right) \pi^{d/2} \\
&= \left(\frac{2\pi}{(1-M\epsilon)n}\right)^{d/2} \left(\hat{g}(\bar{0}) + C\epsilon\right), \qquad (1.18)
\end{aligned}$$

where we did the change of variables $z_i = \sqrt{\frac{1}{2}(1-M\epsilon)n}\ y_i$.

Similarly,

$$\begin{aligned}
I_2 &\geq \int_{[-\epsilon,\epsilon]^d} \left(\hat{g}(\bar{0}) - C\epsilon\right) e^{-\frac{1}{2}(1+M\epsilon)n\sum_i y_i^2} d\boldsymbol{y} \\
&= \left(\frac{1}{2}(1+M\epsilon)n\right)^{-d/2} \left(\hat{g}(\bar{0}) - C\epsilon\right) \int_{\left[-\sqrt{\frac{1}{2}(1+M\epsilon)n}\ \epsilon, \sqrt{\frac{1}{2}(1+M\epsilon)n}\ \epsilon\right]^d} e^{-\sum_j z_j^2} d\bar{z} \\
&\sim \left(\frac{1}{2}(1+M\epsilon)n\right)^{-d/2} \left(\hat{g}(\bar{0}) - C\epsilon\right) \pi^{d/2} \\
&= \left(\frac{2\pi}{(1+M\epsilon)n}\right)^{d/2} \left(\hat{g}(\bar{0}) - C\epsilon\right), \qquad (1.19)
\end{aligned}$$

where we did the change of variables $z_i = \sqrt{\frac{1}{2}(1+M\epsilon)n}\ y_i$.

By putting together (1.17), (1.18) and (1.19), we get

$$I_0 \leq e^{-\alpha n}K + (1+o(1))\left(\frac{2\pi}{(1-M\epsilon)n}\right)^{d/2} \left(\hat{g}(\bar{0}) + C\epsilon\right),$$

and

$$I_0 \geq -e^{-\alpha n}K + (1+o(1))\left(\frac{2\pi}{(1+M\epsilon)n}\right)^{d/2} \left(\hat{g}(\bar{0}) - C\epsilon\right).$$

Thus, since ϵ was small enough but arbitrarily chosen, we can write

$$I_0 \sim \left(\frac{2\pi}{n}\right)^{d/2} \hat{g}(\bar{0}),$$

and then

$$I \sim \frac{1}{\sqrt{|\det H|}} \left(\frac{2\pi}{n}\right)^{d/2} g(\boldsymbol{x}_0) e^{f(\boldsymbol{x}_0)n}. \qquad \square$$

We need the following technical result.

Lemma 1.4.6. *The following system of* 6 *equations in the variables* $d_{i;s}$ *has rank* 6*:*

$$\sum_s s_r d_{i;s} = \frac{5}{6}, \qquad \forall i, r.$$

Moreover, after relabelling the variables as $d_1, \dots, d_{12}$*, the solutions can be expressed by*

$$\begin{aligned} &d_1, \dots, d_6 \quad \text{are free,} \\ &d_k = L_k(d_1, \dots, d_6, 1/6), \quad k = 7, \dots, 12, \end{aligned}$$

where L_k *are linear functions with coefficients in* $\mathbb{Z}$.

Proof. Let us call $s_1 = (1,4)$, $s_2 = (2,3)$, $s_3 = (3,2)$, $s_4 = (4,1)$ the four possible 1-spectra. We can easily found by hand the following solution: For each $i \in \mathbb{Z}_3$,

$$\begin{aligned} d^i_{s_3} &= -2d^i_{s_4} + d^i_{s_1} + 1/6, \\ d^i_{s_2} &= -2d^i_{s_1} + d^i_{s_4} + 1/6, \end{aligned}$$

and the remaining variables are free. $\square$

Hereinafter, we relabel d^i_s as $d_1, \dots, d_{12}$ in the sense of Lemma 1.4.6. The b^i_s are also relabelled as $b_1, \dots, b_{12}$ accordingly. (Recall that b^i_s was defined as $\binom{5}{s}/90$.) For a point $\boldsymbol{d} = (d_1, \dots, d_{12}) \in \mathcal{D}_1$, the first six coordinates will be often denoted by $\tilde{\boldsymbol{d}} = (d_1, \dots, d_6)$ for simplicity.

Let $\epsilon > 0$ be fixed but small enough. We consider the cube of side 2ϵ centred on $\tilde{\boldsymbol{b}}$

$$\tilde{\mathcal{Q}}_1 = \{(d_1, \dots, d_6) \in \mathbb{R}^6 \ : \ d_k \in [b_k - \epsilon, b_k + \epsilon], \ \forall k\},$$

and the discrete set in $\mathbb{Z}^6$

$$\tilde{\mathcal{J}}_1 = \tilde{\mathcal{Q}}_1 \cap \left(\frac{1}{n}\mathbb{Z}^6\right).$$

Let us define their extension to higher dimension

$$\begin{aligned} \mathcal{Q}_1 = \{(d_1, \dots, d_{12}) \in \mathbb{R}^{12} \ &: \ (d_1, \dots, d_6) \in \tilde{\mathcal{Q}}_1, \\ &d_k = L_k(d_1, \dots, d_6, 1/6), \ \forall k = 7, \dots, 12\}, \end{aligned}$$

where the L_k's are as in Lemma 1.4.6, and

$$\mathcal{J}_1 = \mathcal{Q}_1 \cap \left(\frac{1}{n}\mathbb{Z}^{12}\right).$$

Note that $\boldsymbol{b}$ is an interior point of $\mathcal{D}_1$, and that for each k the function $L_k(\cdot, 1/6)$ is continuous. Then, if ϵ is chosen small enough, we can ensure that for some $\delta > 0$

$$\forall \boldsymbol{d} \in \mathcal{Q}_1, \quad d_k > \delta \text{ and } |d_k - b_k| < \delta, \quad k = 1, \ldots, 12, \tag{1.20}$$

and hence $\mathcal{Q}_1 \subset \mathcal{D}_1$. Moreover, since n is always divisible by 6, for each k the function $L_k(\cdot, 1/6)$ maps points from $\frac{1}{n}\mathbb{Z}^{301}$ into $\frac{1}{n}\mathbb{Z}$, and so $\mathcal{J}_1 \subset \mathcal{I}_1$.

Recalling the descriptions of f_1, g_1 and h_1 from (1.4), define for any $(d_1, \ldots, d_6) \in \tilde{\mathcal{Q}}_1$

$$\begin{array}{l} \tilde{f}_1(d_1, \ldots, d_6) = f_1(d_1, \ldots, d_{12}) \\ \tilde{g}_1(d_1, \ldots, d_6) = g_1(d_1, \ldots, d_{12}) \end{array}, \quad \text{where } d_k = L_k(d_1, \ldots, d_6, 1/6), \ \forall k = 7, \ldots, 12.$$

Lemma 1.4.7. *The following statements hold:*

- *f_1 has a unique maximum in $\mathcal{D}_1$ at $\boldsymbol{b}$.*
- *$\tilde{f}_1$ has a unique maximum in $\tilde{\mathcal{Q}}_1$ at $\tilde{\boldsymbol{b}}$, with $e^{f_1(\boldsymbol{b})} = e^{\tilde{f}_1(\tilde{\boldsymbol{b}})} = 90$.*
- *The Hessian $\tilde{H}_1$ of $\tilde{f}_1$ at $\tilde{\boldsymbol{b}}$ is definite negative, and $\det \tilde{H}_1 = 3^{15}11^3$.*
- $\tilde{g}_1(\tilde{\boldsymbol{b}}) = 2^3 3^{12} \neq 0$.
- *Both $\tilde{f}_1$ and $\tilde{g}_1$ are of class $\mathcal{C}^\infty$ in $\tilde{\mathcal{Q}}_1$.*

Proof. The proof follows from Lemma 1.4.3 and direct calculations. □

We compute the contribution to $\mathbf{E}(X)$ of the terms around $\boldsymbol{b}$ and get the following.

Lemma 1.4.8.

$$\sum_{\boldsymbol{d} \in \mathcal{J}_1} q_1(n, \boldsymbol{d}) e^{f_1(\boldsymbol{d})n} \sim \frac{(2\pi n)^{6/2}}{\sqrt{|\det \tilde{H}_1|}} \tilde{g}_1(\tilde{\boldsymbol{b}}) e^{n\tilde{f}_1(\tilde{\boldsymbol{b}})} = \sqrt{\frac{2^6 3^9}{11^3}} (2\pi n)^3 90^n$$

Proof. From (1.20), we see that for all $\boldsymbol{d} \in \mathcal{J}_1 \subset \mathcal{Q}_1$ we must have $d_k > \delta \ \forall k$. Thus, from their definition, all the $m_{r,t}^{i,j}$ are bounded away from 0, $q_1(n, \boldsymbol{d}) \sim g_1(\boldsymbol{d})$ and we can write

$$\sum_{\boldsymbol{d} \in \mathcal{J}_1} q_1(n, \boldsymbol{d}) e^{f_1(\boldsymbol{d})n} \sim \sum_{\mathcal{J}_1} g_1(\boldsymbol{d}) e^{n f_1(\boldsymbol{d})} = \sum_{\tilde{\mathcal{J}}_1} \tilde{g}_1(\tilde{\boldsymbol{d}}) e^{n\tilde{f}_1(\tilde{\boldsymbol{d}})}. \tag{1.21}$$

We note that both $\tilde{f}_1$, $\tilde{g}_1$ and its partial derivatives up to any fixed order are uniformly bounded in the compact set $\tilde{\mathcal{Q}}_1$. By repeated application of the Euler-Maclaurin summation formula (see [1], p. 806), we have that asymptotically as n grows large

$$\sum_{\tilde{\mathcal{J}}_1} \tilde{g}_1(\tilde{\boldsymbol{d}}) e^{n\tilde{f}_1(\tilde{\boldsymbol{d}})} \sim n^6 \int_{\tilde{\mathcal{Q}}_1} \tilde{g}_1(\tilde{\boldsymbol{x}}) e^{n\tilde{f}_1(\tilde{\boldsymbol{x}})} d\tilde{\boldsymbol{x}}. \tag{1.22}$$

Observe from Lemma 1.4.7 that all the conditions required in Lemma 1.4.5 are satisfied. In view of this, we obtain

$$\int_{\tilde{\mathcal{Q}}_1} \tilde{g}_1(\tilde{\boldsymbol{x}}) e^{n\tilde{f}_1(\tilde{\boldsymbol{x}})} d\tilde{\boldsymbol{x}} \sim \frac{1}{\sqrt{|\det \tilde{H}_1|}} \left(\frac{2\pi}{n}\right)^{6/2} \tilde{g}_1(\tilde{\boldsymbol{b}}) e^{n\tilde{f}_1(\tilde{\boldsymbol{b}})}. \tag{1.23}$$

The result follows from (1.21), (1.22), (1.23) and Lemma 1.4.7. □

Now we deal with the remaining terms of the sum.

Lemma 1.4.9. *There exists some positive real* $\alpha < e^{f_1(\boldsymbol{b})}$ *such that*

$$\sum_{\mathcal{I}_1 \setminus \mathcal{J}_1} q_1(n, \boldsymbol{d}) e^{f_1(\boldsymbol{d})n} = o\left(\alpha^n\right).$$

Proof. Let B be the topological closure of $\mathcal{D}_1 \setminus \mathcal{Q}_1$. We recall from Lemma 1.4.7 that f_1 has a unique maximum in $\mathcal{D}_1$ at point $\boldsymbol{b} \notin B$. Then, since B is a compact set and f_1 is continuous, there must be some real $\beta < f_1(\boldsymbol{b})$ such that $f_1(\boldsymbol{x}) \leq \beta \; \forall \boldsymbol{x} \in B$. Now we observe that all terms in the sum $\sum_{\mathcal{I}_1 \setminus \mathcal{J}_1} q_1(n, \boldsymbol{d}) e^{f_1(\boldsymbol{d})n}$ can be uniformly bounded by $Cn^6 e^{\beta n}$, for some fixed constant C. Note furthermore that there are at most $(n+1)^{12}$ terms in the sum. Hence, the result holds by taking $\alpha = (e^{\beta} + e^{f_1(\boldsymbol{b})})/2$. ☐

From Lemmata 1.4.8 and 1.4.9, we obtain

$$\sum_{\mathcal{I}_1} q_1(\boldsymbol{d}) e^{f_1(\boldsymbol{d})n} \sim \sqrt{\frac{2^6 3^9}{11^3}} (2\pi n)^3 90^n.$$

From this and Lemma 1.4.1, we conclude the following:

Theorem 1.4.10.

$$\mathbf{E}(X) \sim \sqrt{\frac{2^2 3^6 5^3}{11^3} \frac{1}{(2\pi n)^2} \left(\frac{25}{24}\right)^n}.$$

1.5 Asymptotic Value of the Second Moment

By analogy to Section 1.4, we denote the tuple $(d_s^{i,j})_{i,j \in \mathbb{Z}_3, s \in \mathcal{S}_2}$ by $\boldsymbol{d}$. Let us consider the function

$$\hat{F}(\boldsymbol{d}) = \left(\prod_{i,j,s} \left(\frac{\binom{5}{s}}{d_s^{i,j}} \right)^{d_s^{i,j}} \right) \left(\prod_{i,j,r,t} \left(m_{r,t}^{i,j} \right)^{\frac{1}{2} m_{r,t}^{i,j}} \right),$$

defined in $\mathcal{D}_2$, where $m_{r,t}^{i,j}$ denotes $\sum_s s_{r,t} d_s^{i,j}$ as before. Throughout this chapter we observe the conventions that $0^0 = 1$ and $0 \log 0 = 0$.

We define

$$f_2(\boldsymbol{d}) = \log \hat{F}(\boldsymbol{d}) = \sum_{i,j,s} d_s^{i,j} \left(\log \binom{5}{s} - \log d_s^{i,j} \right) + \sum_{i,j,r,t} \frac{1}{2} m_{r,t}^{i,j} \log m_{r,t}^{i,j},$$

$$g_2(\boldsymbol{d}) = \frac{\prod_{i,j,r,t} \left(m_{r,t}^{i,j} \right)^{1/4}}{\prod_{i,j,s} (d_s^{i,j})^{1/2}}, \qquad h_2(n) = 2^{-1/2} (2\pi n)^{-305/2} \, 5^{-5n/2}. \tag{1.24}$$

Lemma 1.5.1. *The second moment satisfies*

$$\mathbf{E}(X^2) = h_2(n) \sum_{\boldsymbol{d} \in \mathcal{I}_2} q_2(n, \boldsymbol{d}) e^{f_2(\boldsymbol{d})n}, \tag{1.25}$$

where, as $n \to \infty$ *and uniformly over all* $\boldsymbol{d}$, $q_2(n, \boldsymbol{d}) = O(n^{162})$ *and* $q_2(n, \boldsymbol{d}) \sim g_2(\boldsymbol{d})$ *provided all* $d_s^{i,j}$ *and* $m_{r,t}^{i,j}$ *are bounded away from 0.*

Proof. Apply Stirling's formula to (1.12), we obtain

$$\begin{aligned}
\mathbf{E}(X^2) &= \frac{2^{5n/2}(5n/2)!n!}{(5n)!} \sum_{\boldsymbol{d}\in\mathcal{I}_2} \left\{ \left(\prod_{i,j,s} \frac{\binom{5}{s}^{d_s^{i,j}n}}{(d_s^{i,j}n)!} \right) \left(\prod_{i,j,r,t} \left((m_{r,t}^{i,j}n)! \right)^{1/2} \right) \right\} \\
&\sim \sqrt{\pi n}\; 5^{-5n/2} \sum_{\boldsymbol{d}\in\mathcal{I}_2} \left\{ \frac{(n/e)^n}{(n/e)^{5n/2}} \left(\prod_{i,j,s} \frac{\binom{5}{s}^{d_s^{i,j}n}}{(d_s^{i,j}n)!} \right) \left(\prod_{i,j,r,t} \left((m_{r,t}^{i,j}n)! \right)^{1/2} \right) \right\} \\
&= h_2(n) \sum_{\boldsymbol{d}\in\mathcal{I}_2} \left\{ \frac{(2\pi n)^{153}(n/e)^n}{(n/e)^{5n/2}} \left(\prod_{i,j,s} \frac{\binom{5}{s}^{d_s^{i,j}n}}{(d_s^{i,j}n)!} \right) \left(\prod_{i,j,r,t} \left((m_{r,t}^{i,j}n)! \right)^{1/2} \right) \right\}. \qquad (1.26)
\end{aligned}$$

Recall from the proof of Lemma 1.4.1 that Stirling's formula implies $k! = \sqrt{2\pi\eta(k)}(k/e)^k$ for all $k \geq 0$, where $\eta(k) > 0$ and $\eta(k) \sim k$ as $k \to \infty$. So we have

$$\begin{aligned}
&\left(\prod_{i,j,s} \frac{\binom{5}{s}^{d_s^{i,j}n}}{(d_s^{i,j}n)!} \right) \left(\prod_{i,j,r,t} \left((m_{r,t}^{i,j}n)! \right)^{1/2} \right) \\
&= \left(\prod_{i,j,s} \frac{\binom{5}{s}^{d_s^{i,j}n}}{\sqrt{2\pi\eta(d_s^{i,j}n)} \left(\frac{d_s^{i,j}n}{e} \right)^{d_s^{i,j}n}} \right) \left(\prod_{i,j,r,t} \left(\sqrt{2\pi\eta(m_{r,t}^{i,j}n)} \left(\frac{m_{r,t}^{i,j}n}{e} \right)^{m_{r,t}^{i,j}n} \right) \right)^{1/2} \\
&= \frac{(n/e)^{5n/2}}{(2\pi n)^{153}(n/e)^n} \frac{\prod_{i,j,r,t}(\eta(nm_{r,t}^{i,j})^{1/4}n^{-1/4})}{\prod_{i,j,s}(\eta(nd_s^{i,j})^{1/2}n^{-1/2})} \left(\prod_{i,j,s} \left(\frac{\binom{5}{s}}{d_s^{i,j}} \right)^{d_s^{i,j}n} \right) \left(\prod_{i,j,r,t} \left(m_{r,t}^{i,j} \right)^{m_{r,t}^{i,j}n/2} \right) \\
&= \frac{(n/e)^{5n/2}}{(2\pi n)^{153}(n/e)^n} q_2(n, \boldsymbol{d}) e^{f_2(\boldsymbol{d})n},
\end{aligned}$$

for a function q_2 of the type described in the statement of the lemma. Combining this with (1.26) yields the lemma. □

We consider the maximum base of the exponential part of the terms in (1.25), taken over all points in the polytope $\mathcal{D}_2$:

$$M_2 = \max_{\boldsymbol{d}\in\mathcal{D}_2} \left\{ 5^{-5/2} e^{f_2(\boldsymbol{d})} \right\}.$$

This is well defined, due to the compactness of the domain and the continuity of the expression. Note that the exponential behaviour of the second moment is governed by M_2, since the number of terms in the sum in (1.25) is polynomial with respect to n.

In the next subsection we determine the value of M_2 under the Maximum Hypothesis. In Subsection 1.5.2, based on that fact and using a Laplace-type integration argument, we compute the sub-exponential factors in the asymptotic expression of the second moment.

1.5.1 Computing M_2

We will maximise $\hat{F}$ in two phases. In the first one, we maximise $\hat{F}$ assuming the matching variables $m_{r,t}^{i,j}$ are fixed constants. These constants must be compatible with the polytope

$\mathcal{D}_2$ over which $\hat{F}$ is defined, so we define $\mathcal{M}$ to be the set of vectors $\boldsymbol{m}$ of 2×2 matrices $(m^{i,j})_{i,j\in\mathbb{Z}_3}$ such that (1.10) holds for some $\boldsymbol{d}\in\mathcal{D}_2$.

We often consider variables $d_s^{i,j}$ and $m_{r,t}^{i,j}$ for fixed $i,j\in\mathbb{Z}_3$. To simplify notation, we delete the indices i and j when they are fixed throughout the formula. For any $0<c\in\mathbb{R}$, define

$$\mathcal{D}'(c)=\{(d_s)_{s\in\mathcal{S}}\in\mathbb{R}^{36}\ :\ d_s\geq 0\ \forall s,\ \sum_s d_s=c\},$$

and let $\mathcal{M}'(c)$ be the set of 2×2 matrices m such that (1.10) holds for some $(d_s)_{s\in\mathcal{S}}\in\mathcal{D}'(c)$ (after deleting superscripts i and j). We will use $\boldsymbol{d}$ to denote points in $\mathcal{D}_2$ and also points in $\mathcal{D}'(c)$. The meaning will be clear from the context.

In order to give an alternative characterisation of the matching variables $m_{r,t}^{i,j}$, we consider the following equations for all ordered pairs (i,j), $i,j\in\mathbb{Z}_3$, and all $r,t\in\{-1,1\}$:

$$\begin{array}{rcl} m_{r,t}^{i,j} & \geq & 0,\\ m_{r,t}^{i,j}+m_{r,-t}^{i,j} & \leq & 4(m_{-r,t}^{i,j}+m_{-r,-t}^{i,j}),\\ m_{r,t}^{i,j}+m_{-r,t}^{i,j} & \leq & 4(m_{r,-t}^{i,j}+m_{-r,-t}^{i,j}). \end{array} \tag{1.27}$$

That is, for all such i and j, the entries of $m^{i,j}$ are non-negative, neither row sum is greater than 4 times the other, and neither column sum is greater than 4 times the other.

Lemma 1.5.2. *Let $c>0\in\mathbb{R}$. The set $\mathcal{M}'(c)$ can be alternatively described as the polytope containing all matrices m such that*

$$\sum_{r,t} m_{r,t}=5c, \tag{1.28}$$

and the constraints in (1.27) *hold. Similarly, $\mathcal{M}$ is the polytope containing all vectors $\boldsymbol{m}$ of matrices $m^{i,j}$ such that*

$$\sum_{i,r,t} m_{r,t}^{i,j}=5/3,\quad \sum_{j,r,t} m_{r,t}^{i,j}=5/3, \tag{1.29}$$

and the constraints in (1.8), (1.27) *hold.*

Proof. Let $\mathcal{A}$ be the set of matrices m satisfying (1.27) and (1.28). We first prove that $\mathcal{M}'(c)\subseteq\mathcal{A}$. Let m be a matrix in $\mathcal{M}'(c)$. Then, for some $\boldsymbol{d}\in\mathcal{D}'(c)$, we have

$$\sum_{r,t} m_{r,t}=\sum_{r,t}\sum_s s_{r,t}d_s=\sum_s\sum_{r,t} s_{r,t}d_s=\sum_s 5d_s=5c,$$

and (1.28) is satisfied. Moreover, we observe that for any spectrum s, we have

$$s_{r,t}\geq 0,\ \ s_{r,t}+s_{r,-t}\leq 4(s_{-r,t}+s_{-r,-t})\ \ \text{and}\ \ s_{r,t}+s_{-r,t}\leq 4(s_{r,-t}+s_{-r,-t}).$$

Then m must satisfy the constraints in (1.27), since it is a positive linear combination of spectra, and $m\in\mathcal{A}$.

Now we prove that $\mathcal{A}\subseteq\mathcal{M}'(c)$. We have that $\mathcal{A}$ is a polytope and so it is the convex hull of its vertices:

$$\begin{bmatrix} c & 0\\ 0 & 4c\end{bmatrix},\begin{bmatrix} 0 & c\\ c & 3c\end{bmatrix},\begin{bmatrix} 0 & c\\ 4c & 0\end{bmatrix},\begin{bmatrix} c & 0\\ 3c & c\end{bmatrix},$$

$$\begin{bmatrix} 0 & 4c \\ c & 0 \end{bmatrix}, \begin{bmatrix} c & 3c \\ 0 & c \end{bmatrix}, \begin{bmatrix} 4c & 0 \\ 0 & c \end{bmatrix}, \begin{bmatrix} 3c & c \\ c & 0 \end{bmatrix}.$$

Each of these vertices v has the shape of some spectrum s times c. By making $d_s = c$ and $d_{s'} = 0$ for $s' \neq s$, we show that $v \in \mathcal{M}'(c)$.

Moreover, we observe that $\mathcal{M}'(c)$ is a convex set, since it is the image of $\mathcal{D}'(c)$ under a linear mapping. Then $\mathcal{M}'(c)$ must contain the convex hull of the vertices of $\mathcal{A}$, and thus contains $\mathcal{A}$.

The second statement in the lemma follows from the previous fact and from the definition of $\mathcal{M}$. □

For any fixed $\boldsymbol{m} \in \mathcal{M}$, let $\tilde{F}(\boldsymbol{m})$ be the maximum of $\hat{F}$ restricted to $\boldsymbol{d} \in \mathcal{D}_2$ such that (1.10) holds. To express $\tilde{F}(\boldsymbol{m})$ in terms of $\boldsymbol{m}$, we will use the matrix function

$$\Phi \begin{bmatrix} x & y \\ z & w \end{bmatrix} = (x+y+z+w)^5 - (x+y)^5 - (x+z)^5 - (y+w)^5 - (z+w)^5 + x^5 + y^5 + z^5 + w^5 \quad (1.30)$$

and, for each of the nine possible pairs (i, j), $i, j \in \mathbb{Z}_3$, consider the 4×4 system

$$\frac{\partial\ \Phi\mu^{i,j}}{\partial\mu^{i,j}_{r,t}}\ \mu^{i,j}_{r,t} = m^{i,j}_{r,t}, \quad r, t = -1, 1, \quad (1.31)$$

in the matrix variables $\mu^{i,j}$.

Lemma 1.5.3. *For any $\boldsymbol{m}$ in the interior of $\mathcal{M}$, each of the nine systems in (1.31) has a unique positive solution. Moreover, in terms of the solutions of these systems,*

$$\tilde{F}(\boldsymbol{m}) = \prod_{i,j,r,t} \left(\frac{(m^{i,j}_{r,t})^{\frac{1}{2}}}{\mu^{i,j}_{r,t}} \right)^{m^{i,j}_{r,t}},$$

and the equation remains valid for $\boldsymbol{m}$ on the boundary of $\mathcal{M}$ if the expression on the right is extended by continuity.

Proof. We assume that $\boldsymbol{m}$ is a fixed vector in the interior of $\mathcal{M}$. In order to compute $\tilde{F}(\boldsymbol{m})$, it is sufficient to maximise the function $\hat{F}(\boldsymbol{d})$ for non-negative $d^{i,j}_s$ subject to (1.10), since the other constraints are trivially satisfied. We observe that the factor $\prod \left(m^{i,j}_{r,t}\right)^{\frac{1}{2}m^{i,j}_{r,t}}$ is constant, and that variables $d^{i,j}_s$ with different pairs of indices (i, j) appear in different factors of $\hat{F}$ and also in different constraints. Thus, it is sufficient to maximise, separately for each $i, j \in \mathbb{Z}_3$, the function

$$G^{i,j} = \prod_{s \in \mathcal{S}} \left(\frac{\binom{5}{s}}{d^{i,j}_s} \right)^{d^{i,j}_s}, \quad (1.32)$$

over non-negative $d^{i,j}_s$ subject to the matrix constraint (1.10). For the remaining of this proof, we fix i and j and thus omit superscripts as discussed above.

Let $\mathcal{R}$ be the polytope containing all $\boldsymbol{d} = (d_s)_{s \in \mathcal{S}}$ such that d_s is non-negative for all $s \in \mathcal{S}$, and satisfying (1.10). The fact that $\boldsymbol{m}$ is in the interior of $\mathcal{M}$ implies that $\mathcal{R}$ contains points with all the d_s strictly positive. In fact, the interior of $\mathcal{R}$ consists of all those points in $\mathcal{R}$ with this property.

For any point $\boldsymbol{d}_0$ on the boundary of $\mathcal{R}$ we select a segment joining $\boldsymbol{d}_0$ with some interior point. We observe that, in moving along the segment from the interior of $\mathcal{R}$ towards $\boldsymbol{d}_0$, the directional derivative of $\log G$ contains the sum of some bounded terms plus some terms of the type $\log d_s$ with positive coefficient, which become large as we approach $\boldsymbol{d}_0$. Hence, G does not maximise at the boundary of $\mathcal{R}$.

We temporarily relax the constraint (1.10) and observe that the Hessian of $\log G$ is negative definite for any tuple of positive d_s. Hence $\log G$ is strictly concave in that domain and also in the interior of $\mathcal{R}$, since linear constraints do not affect concavity. Thus, the maximum of G is unique and occurs in the only stationary point of $\log G$ in the interior of $\mathcal{R}$.

We are now in a good position to apply the Lagrange multipliers method to look for stationary points of $\log G$. We consider

$$\log G = \sum_s d_s \left(\log \binom{5}{s} - \log d_s \right), \tag{1.33}$$

for positive d_s subject to the four constraints:

$$L_{r,t} = \sum_s s_{r,t} d_s - m_{r,t} = 0, \quad r, t \in \{-1, 1\}. \tag{1.34}$$

For each one of the four constraints $L_{r,t}$ in (1.34) a Lagrange multiplier $\lambda_{r,t}$ is introduced. Then we obtain the following equations:

$$\log \binom{5}{s} - 1 - \log d_s = \sum_{r,t} \lambda_{r,t} s_{r,t}, \quad \forall s \in \mathcal{S} \tag{1.35}$$

which, together with the constraints (1.34) have a unique solution when $\boldsymbol{d}$ is the only stationary point of $\log G$. Let us define $\mu_{r,t} = \exp(-\lambda_{r,t} - 1/5)$. After exponentiating (1.35), and noting that the sum of the $s_{r,t}$ is 5, we have

$$d_s = \binom{5}{s} \prod_{r,t} (\mu_{r,t})^{s_{r,t}}, \quad \forall s \in \mathcal{S}, \tag{1.36}$$

and combining this with (1.34) gives

$$m_{r,t} = \sum_s \left(s_{r,t} \binom{5}{s} \prod_{r,t} (\mu_{r,t})^{s_{r,t}} \right), \quad r, t \in \{-1, 1\}. \tag{1.37}$$

By construction, this system has a unique positive solution, and (1.36) gives the maximiser of G in terms of this solution. We observe that (1.37) is exactly the same system as the one in (1.31).

Now the maximum of G can be obtained by plugging (1.36) into (1.32), resulting in

$$\max_{\boldsymbol{d} \in \mathcal{R}} G(\boldsymbol{d}) = \prod_{r,t} \left(\frac{1}{\mu_{r,t}} \right)^{m_{r,t}}, \tag{1.38}$$

and the required expression for $\tilde{F}(\boldsymbol{m})$ follows by elementary computations. □

Let us now define for any $\boldsymbol{d} \in \mathcal{D}'(1/9)$ the auxiliary function

$$\hat{G}(\boldsymbol{d}) = \left(\prod_s \left(\frac{\binom{5}{s}}{d_s} \right)^{d_s} \right) \left(\prod_{r,t} (m_{r,t})^{\frac{1}{2} m_{r,t}} \right), \tag{1.39}$$

where $m_{r,t} = \sum_s s_{r,t}\, d_s$. (Recall that $0^0 = 1$.)

Lemma 1.5.4. *The function $\hat{G}$ takes its maximum on $\mathcal{D}'(1/9)$ in the interior of $\mathcal{D}'(1/9)$.*

Proof. Notice that the boundary of $\mathcal{D}'(1/9)$ comprises the points where for at least one s, $d_s = 0$ and $\sum_s d_s = 1/9$. Observe that it is sufficient to prove the statement for $\log \hat{G}$. The continuity of $\log \hat{G}$ at the boundary points of $\mathcal{D}'(1/9)$ follows from the fact that $\lim_{x \to 0} x^x = 1$. After proving $\log \hat{G}$ is continuous at the boundary of $\mathcal{D}'(1/9)$, take any $\boldsymbol{d}$ on the boundary. Here $d_{s_0} = 0$ for some s_0. Then $d_{s_1} > 0$ for some s_1 since the sum of entries of $\boldsymbol{d}$ is 1/9. At any point $\boldsymbol{d}$ such that $d_s > 0$,

$$\frac{\partial \log \hat{G}}{\partial d_s} = \log \binom{5}{s} - 1 - \log d_s + \frac{5}{2} + \sum_{r,t} \frac{1}{2} s_{r,t} \log m_{r,t}. \tag{1.40}$$

Note that if $d_s > 0$ then all the $m_{r,t}$ corresponding to a non-zero $s_{r,t}$ are also necessarily non-zero. As a first case, suppose none of the $m_{r,t}$ is zero at $\boldsymbol{d}$. Then at a point $\boldsymbol{d} + \epsilon E_{s_0} - \epsilon E_{s_1}$ we have $\frac{\partial \log \hat{G}}{\partial d_{s_0}} - \frac{\partial \log \hat{G}}{\partial d_{s_1}} \to \infty$ as $\epsilon \to 0$, where E_s denotes the vector with 1 in its s coordinate and zero elsewhere. (Since the first partial goes to ∞ and the second is bounded.) Hence there is no maximum at $\boldsymbol{d}$.

Next suppose precisely one $m_{r,t}$ is zero at $\boldsymbol{d}$. Fix such values of r and t. Pick an s such that $s_{r,t} = 1$. Then $d_s = 0$ at $\boldsymbol{d}$. So rename s as s_0 and use the above argument, choosing again any s_1 with $d_{s_1} > 0$. Now the unbounded terms in $\frac{\partial \log \hat{G}}{\partial d_{s_0}}$ are $-\log d_{s_0} + \frac{1}{2}(s_0)_{r,t} \log m_{r,t}$, and we have $m_{r,t} \geq d_{s_0}$ because $(s_0)_{r,t} = 1$. It follows that there is no maximum at $\boldsymbol{d}$.

For two different $m_{r,t}$ equal to zero at $\boldsymbol{d}$, pick the spectrum s_0 to have 1 in one of the corresponding positions, and zero in the other. Then the same argument as above gives the result.

So no local maximum occurs on the boundary. □

Lemma 1.5.5. *The function $\hat{G}$ has a unique maximum in $\mathcal{D}'(1/9)$ at the point where all the d_s are equal to $\binom{5}{s}/8100$. The function value at the maximum is $(5^{5/2}25/24)^{1/9}$.*

Proof. We note that (1.10) maps the interior of $\mathcal{D}'(1/9)$ into the interior of $\mathcal{M}'(1/9)$. As a result and in view of Lemma 1.5.4, the maximum of $\hat{G}$, under mapping (1.10), does not occur on the boundary of $\mathcal{M}'(1/9)$.

Assume that m is a fixed matrix in the interior of $\mathcal{M}'(1/9)$. We first maximise $\hat{G}$ in $\mathcal{D}'(1/9)$ subject to the matrix constraint (1.10). Denote this maximum by $\tilde{G}(m)$. By arguing as in the proof of Lemma 1.5.3, we have

$$\tilde{G}(m) = \prod_{r,t} \left(\frac{(m_{r,t})^{\frac{1}{2}}}{\mu_{r,t}} \right)^{m_{r,t}},$$

where the $\mu_{r,t}$ are the unique positive solution of the system in (1.31) after deleting superscripts i and j. Moreover, the maximiser is given in terms of this solution by (1.36).

We now maximise $\tilde{G}$ in the interior of $\mathcal{M}'(1/9)$, by applying the Lagrange multiplier method to

$$\log \tilde{G}(m) = \sum_{r,t} m_{r,t} \left(\frac{1}{2} \log m_{r,t} - \log \mu_{r,t} \right),$$

subject to

$$\sum_{r,t} m_{r,t} = 5/9.$$

We need some preliminary computations. By adding the four equations in (1.31) and taking into account (1.30), we have

$$5\Phi(\mu) = \sum_{r,t} m_{r,t}.$$

In view of this, we have that for all $r, t \in \{-1, 1\}$

$$\begin{aligned} \sum_{r',t' \in \{-1,1\}} m_{r',t'} \frac{\partial \log \mu_{r',t'}}{\partial m_{r,t}} &= \sum_{r',t' \in \{-1,1\}} \frac{m_{r',t'}}{\mu_{r',t'}} \frac{\partial \mu_{r',t'}}{\partial m_{r,t}} \\ &= \sum_{r',t' \in \{-1,1\}} \frac{\partial \Phi(\mu)}{\partial \mu_{r',t'}} \frac{\partial \mu_{r',t'}}{\partial m_{r,t}} = \frac{\partial \Phi(\mu)}{\partial m_{r,t}} = \frac{1}{5}. \end{aligned} \tag{1.41}$$

This allows us to compute

$$\begin{aligned} \frac{\partial \log \tilde{G}(m)}{\partial m_{r,t}} &= \frac{1}{2} \log m_{r,t} + \frac{1}{2} - \log \mu_{r,t} - \sum_{r',t'} m_{r',t'} \frac{\partial \log \mu_{r',t'}}{\partial m_{r,t}} \\ &= \frac{1}{2} \log m_{r,t} - \log \mu_{r,t} + \frac{3}{10}, \end{aligned} \tag{1.42}$$

and obtain the equations

$$\frac{1}{2} \log m_{r,t} - \log \mu_{r,t} + \frac{3}{10} = \lambda, \quad \forall r, t \in \{-1, 1\}, \tag{1.43}$$

where λ is the Lagrange multiplier introduced by the single constraint. After exponentiating (1.43), and defining $\lambda' = \exp(\lambda - 3/10)$, we can write

$$\frac{\sqrt{m_{r,t}}}{\mu_{r,t}} = \lambda', \quad \forall r, t \in \{-1, 1\}. \tag{1.44}$$

We relabel the entries of the matrices m and μ as

$$\begin{bmatrix} m_1 & m_2 \\ m_3 & m_4 \end{bmatrix}, \qquad \begin{bmatrix} \mu_1 & \mu_2 \\ \mu_3 & \mu_4 \end{bmatrix}.$$

Combining (1.44) and (1.31), we get

$$\mu_i \frac{\partial \Phi}{\partial \mu_j} - \mu_j \frac{\partial \Phi}{\partial \mu_i} = 0, \quad \forall i, j \in \{1, \ldots, 4\}.$$

We can factorise the following equation:

$$\mu_1 \frac{\partial \Phi}{\partial \mu_4} - \mu_4 \frac{\partial \Phi}{\partial \mu_1} = 0,$$

and get

$$(\mu_1 - \mu_4)\, P = 0,$$

where

$$\begin{aligned} P = \quad & 120\,\mu_1\,\mu_2\,\mu_3\,\mu_4 + 20\,{\mu_1}^3\,\mu_2 + 20\,{\mu_1}^3\,\mu_3 + 25\,{\mu_1}^3\,\mu_4 + 30\,{\mu_1}^2\,{\mu_2}^2 \\ & +30\,{\mu_1}^2\,{\mu_3}^2 + 35\,{\mu_1}^2\,{\mu_4}^2 + 5\,{\mu_4}^4 + 20\,\mu_1\,{\mu_2}^3 + 20\,\mu_1\,{\mu_3}^3 \\ & +25\,\mu_1\,{\mu_4}^3 + 5\,{\mu_1}^4 + 60\,{\mu_1}^2\,\mu_2\,\mu_3 + 80\,{\mu_1}^2\,\mu_2\,\mu_4 + 80\,{\mu_1}^2\,\mu_3\,\mu_4 \\ & +60\,{\mu_2}^2\,\mu_3\,\mu_4 + 60\,\mu_1\,{\mu_2}^2\,\mu_3 + 90\,\mu_1\,{\mu_2}^2\,\mu_4 + 60\,\mu_1\,\mu_2\,{\mu_3}^2 \\ & +80\,\mu_1\,\mu_2\,{\mu_4}^2 + 90\,\mu_1\,{\mu_3}^2\,\mu_4 + 80\,\mu_1\,\mu_3\,{\mu_4}^2 + 60\,\mu_2\,{\mu_3}^2\,\mu_4 \\ & +60\,\mu_2\,\mu_3\,{\mu_4}^2 + 20\,{\mu_2}^3\,\mu_3 + 20\,{\mu_2}^3\,\mu_4 + 30\,{\mu_2}^2\,{\mu_3}^2 + 30\,{\mu_2}^2\,{\mu_4}^2 \\ & +20\,\mu_2\,{\mu_3}^3 + 20\,\mu_2\,{\mu_4}^3 + 20\,{\mu_3}^3\,\mu_4 + 30\,{\mu_3}^2\,{\mu_4}^2 + 20\,\mu_3\,{\mu_4}^3, \end{aligned}$$

which is strictly positive, so $\mu_1 = \mu_4$. Similarly, we can factorise

$$\mu_2 \frac{\partial \Phi}{\partial \mu_3} - \mu_3 \frac{\partial \Phi}{\partial \mu_2} = 0,$$

and get

$$(\mu_2 - \mu_3)\, Q = 0,$$

where

$$\begin{aligned} Q = \quad & 120\,\mu_1\,\mu_2\,\mu_3\,\mu_4 + 20\,{\mu_1}^3\,\mu_2 + 20\,{\mu_1}^3\,\mu_3 + 20\,{\mu_1}^3\,\mu_4 + 30\,{\mu_1}^2\,{\mu_2}^2 \\ & +30\,{\mu_1}^2\,{\mu_3}^2 + 30\,{\mu_1}^2\,{\mu_4}^2 + 5\,{\mu_3}^4 + 5\,{\mu_2}^4 + 20\,\mu_1\,{\mu_2}^3 + 20\,\mu_1\,{\mu_3}^3 \\ & +20\,\mu_1\,{\mu_4}^3 + 90\,{\mu_1}^2\,\mu_2\,\mu_3 + 60\,{\mu_1}^2\,\mu_2\,\mu_4 + 60\,{\mu_1}^2\,\mu_3\,\mu_4 \\ & +80\,{\mu_2}^2\,\mu_3\,\mu_4 + 80\,\mu_1\,{\mu_2}^2\,\mu_3 + 60\,\mu_1\,{\mu_2}^2\,\mu_4 + 80\,\mu_1\,\mu_2\,{\mu_3}^2 \\ & +60\,\mu_1\,\mu_2\,{\mu_4}^2 + 60\,\mu_1\,{\mu_3}^2\,\mu_4 + 60\,\mu_1\,\mu_3\,{\mu_4}^2 + 80\,\mu_2\,{\mu_3}^2\,\mu_4 \\ & +90\,\mu_2\,\mu_3\,{\mu_4}^2 + 25\,{\mu_2}^3\,\mu_3 + 20\,{\mu_2}^3\,\mu_4 + 35\,{\mu_2}^2\,{\mu_3}^2 + 30\,{\mu_2}^2\,{\mu_4}^2 \\ & +25\,\mu_2\,{\mu_3}^3 + 20\,\mu_2\,{\mu_4}^3 + 20\,{\mu_3}^3\,\mu_4 + 30\,{\mu_3}^2\,{\mu_4}^2 + 20\,\mu_3\,{\mu_4}^3, \end{aligned}$$

which is also strictly positive, so $\mu_2 = \mu_3$. Finally, we substitute μ_4 by μ_1 and μ_3 by μ_2 in

$$\mu_1 \frac{\partial \Phi}{\partial \mu_2} - \mu_2 \frac{\partial \Phi}{\partial \mu_1} = 0,$$

and then factorise it to obtain

$$(\mu_1 - \mu_2)\, R = 0,$$

where

$$R = 70\,{\mu_1}^4 + 275\,{\mu_1}^3\,\mu_2 + 415\,{\mu_1}^2\,{\mu_2}^2 + 275\,\mu_1\,{\mu_2}^3 + 70\,{\mu_2}^4,$$

which is again strictly positive, so $\mu_1 = \mu_2$. Hence, all the μ_i are equal (and therefore all the m_i are also equal).

Since the m_i are equal and sum to 5/9, each m_i must be equal to 5/36. Substituting this value into any of the equations in (1.31), we deduce that $\mu_i = 2^{-2/5}3^{-4/5}5^{-2/5}$. This shows that the Lagrange multiplier problem has a unique solution. This solution must correspond to the unique stationary point of $\tilde{G}$ in the interior of $\mathcal{M}'(1/9)$, which must then be a maximum.

Finally, (1.36) gives the maximiser of $\hat{G}$ in $\mathcal{D}'(1/9)$ when the $m_{r,t}$ are fixed to be equal and thus the $\mu_{r,t}$ are also equal. The maximum value of $\hat{G}$ is computed from its definition. □

Recall the definition of the nine overlap variables from Section 1.3.2. We observe that (1.11) maps $\mathcal{D}_2$ into a polytope of dimension 4. The vectors $(n^{i,j})$ in this polytope can be expressed in terms of four variables by

$$\begin{aligned} &n^{0,2} = 1/3 - n^{0,0} - n^{0,1}, \quad n^{1,2} = 1/3 - n^{1,0} - n^{1,1}, \quad n^{2,0} = 1/3 - n^{0,0} - n^{1,0}, \\ &n^{2,1} = 1/3 - n^{0,1} - n^{1,1}, \quad n^{2,2} = n^{0,0} + n^{0,1} + n^{1,0} + n^{1,1} - 1/3, \end{aligned} \tag{1.45}$$

where the variables $n^{0,0}$, $n^{0,1}$, $n^{1,0}$ and $n^{1,0}$ take arbitrary non-negative real values such that

$$\begin{aligned} &n^{0,0} + n^{0,1} \le \frac{1}{3}, \quad n^{1,0} + n^{1,1} \le \frac{1}{3}, \quad n^{0,0} + n^{1,0} \le \frac{1}{3}, \\ &n^{0,1} + n^{1,1} \le \frac{1}{3}, \quad n^{0,0} + n^{0,1} + n^{1,0} + n^{1,1} \ge \frac{1}{3}. \end{aligned} \tag{1.46}$$

We are now in a good position to define the function F used in the statement of the Maximum Hypothesis. The domain of F is the set $\mathcal{N}$ of all non-negative real vectors $\boldsymbol{n} = (n^{0,0}, n^{0,1}, n^{1,0}, n^{1,1})$ satisfying (1.46). For each $\boldsymbol{n}$ in $\mathcal{N}$, we compute the nine overlap variables from (1.45) and define $F(\boldsymbol{n})$ to be the maximum of $\hat{F}(\boldsymbol{d})$ over $\mathcal{D}_2$ subject to the constraints in (1.11). This definition of F is repeated in Section 1.7, which also contains an alternative equivalent definition (see also Figure 1.1 in Section 1.7).

Let $\boldsymbol{b} = (b_s^{i,j})_{s\in\mathcal{S}, i,j\in\mathbb{Z}_3}$ be the point in $\mathcal{D}_2$ where $b_s^{i,j} = \frac{\binom{5}{s}}{8100}$ for all i, j, s. Now we return to our main function f_2, which was defined in (1.24).

Lemma 1.5.6. *Under the Maximum Hypothesis, the function f_2 has a unique maximiser in $\mathcal{D}_2$ at $\boldsymbol{b}$. Moreover, $M_2 := \max_{\boldsymbol{d}\in\mathcal{D}_2}\left\{5^{-5/2}e^{f_2(\boldsymbol{d})}\right\} = 25/24$.*

Proof. Recall that $f_2 = \log \hat{F}$. The Maximum Hypothesis implies that any maximiser of $\hat{F}$ on $\mathcal{D}_2$ must satisfy $\sum_{s\in\mathcal{S}} d_s^{i,j} = 1/9$, for all $i, j \in \mathbb{Z}_3$. Let us momentarily relax the constraints in (1.8), and maximise each factor

$$\hat{G}^{i,j}(\boldsymbol{d}) = \left(\prod_s \left(\frac{\binom{5}{s}}{d_s^{i,j}}\right)^{d_s^{i,j}}\right)\left(\prod_{r,t} (m_{r,t}^{i,j})^{\frac{1}{2}m_{r,t}^{i,j}}\right),$$

separately in $\mathcal{D}'(1/9)$. In view of Lemma 1.5.5, $\boldsymbol{b}$ is the unique maximiser and the maximum value of each factor is $(5^{5/2}25/24)^{1/9}$. We observe that the constraints in (1.8) are also satisfied by $\boldsymbol{b}$. Therefore $\boldsymbol{b}$ is the unique maximiser of $\hat{F}$ and the maximum function value is $\left((5^{5/2}25/24)^{1/9}\right)^9 = 5^{5/2}25/24$. □

1.5.2 Subexponential Factors of $\mathbf{E}(X^2)$

Here we complete the computation of the asymptotic expression of $\mathbf{E}(X^2)$ under the Maximum Hypothesis by using a standard Laplace-type integration technique.

First we need the following result:

Lemma 1.5.7. *The following system of* 24 *equations in the variables* $d_s^{i,j}$ *has rank* 23*:*

$$\sum_s s_1^t d_s^{i,j} - \sum_s s_{-1}^{-t} d_s^{i+1,j+t} = 0 \quad \forall i,j,t, \qquad \sum_{j,s} d_s^{i,j} = 1/3 \quad \forall i, \qquad \sum_{i,s} d_s^{i,j} = 1/3 \quad \forall j.$$

Moreover, after relabelling the variables as $d_1, \ldots, d_{324}$*, the solutions can be expressed by*

$$\begin{aligned} &d_1, \ldots, d_{301} \quad \textit{free}, \\ &d_k = L_k(d_1, \ldots, d_{301}, 1/6), \quad k = 302, \ldots, 324, \end{aligned}$$

where L_k *are linear functions with coefficients in* $\mathbb{Z}$.

Proof. Let us call

$$s_1 = \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix}, \quad s_2 = \begin{bmatrix} 0 & 1 \\ 2 & 2 \end{bmatrix}, \quad s_3 = \begin{bmatrix} 0 & 1 \\ 3 & 1 \end{bmatrix}, \quad s_4 = \begin{bmatrix} 0 & 1 \\ 4 & 0 \end{bmatrix},$$

$$s_5 = \begin{bmatrix} 0 & 2 \\ 1 & 2 \end{bmatrix}, \quad s_6 = \begin{bmatrix} 0 & 2 \\ 2 & 1 \end{bmatrix}, \quad s_7 = \begin{bmatrix} 0 & 2 \\ 3 & 0 \end{bmatrix}, \quad s_8 = \begin{bmatrix} 0 & 3 \\ 1 & 1 \end{bmatrix},$$

$$s_9 = \begin{bmatrix} 0 & 3 \\ 2 & 0 \end{bmatrix}, \quad s_{10} = \begin{bmatrix} 0 & 4 \\ 1 & 0 \end{bmatrix}, \quad s_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, \quad s_{12} = \begin{bmatrix} 1 & 0 \\ 1 & 3 \end{bmatrix},$$

$$s_{13} = \begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix}, \quad s_{14} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}, \quad s_{15} = \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix}, \quad s_{16} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix},$$

$$s_{17} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}, \quad s_{18} = \begin{bmatrix} 1 & 1 \\ 3 & 0 \end{bmatrix}, \quad s_{19} = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}, \quad s_{20} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix},$$

$$s_{21} = \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix}, \quad s_{22} = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}, \quad s_{23} = \begin{bmatrix} 1 & 3 \\ 1 & 0 \end{bmatrix}, \quad s_{24} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix},$$

$$s_{25} = \begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}, \quad s_{26} = \begin{bmatrix} 2 & 0 \\ 2 & 1 \end{bmatrix}, \quad s_{27} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \quad s_{28} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix},$$

$$s_{29} = \begin{bmatrix} 2 & 1 \\ 2 & 0 \end{bmatrix}, \quad s_{30} = \begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix}, \quad s_{31} = \begin{bmatrix} 2 & 2 \\ 1 & 0 \end{bmatrix}, \quad s_{32} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix},$$

$$s_{33} = \begin{bmatrix} 3 & 0 \\ 1 & 1 \end{bmatrix}, \quad s_{34} = \begin{bmatrix} 3 & 1 \\ 0 & 1 \end{bmatrix}, \quad s_{35} = \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix}, \quad s_{36} = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then we have the following solution:

$d_{s_{13}}^{2,2} = d_{s_{34}}^{1,1} - d_{s_{14}}^{2,2} - d_{s_{11}}^{2,2} - d_{s_{19}}^{2,2} - d_{s_{20}}^{2,2} - d_{s_{21}}^{2,2} + d_{s_{36}}^{1,1} - 2d_{s_{26}}^{2,2} + d_{s_{33}}^{1,1} - d_{s_{12}}^{2,2} - 3d_{s_{32}}^{2,2} + d_{s_{28}}^{1,1} - 4d_{s_{36}}^{2,2} - d_{s_{22}}^{2,2} - 2d_{s_{27}}^{2,2} - 2d_{s_{25}}^{2,2} + 2d_{s_{32}}^{1,1} - 2d_{s_{28}}^{2,2} - 2d_{s_{29}}^{2,2} - 2d_{s_{30}}^{2,2} - 2d_{s_{31}}^{2,2} + 3d_{s_1}^{1,1} + 2d_{s_2}^{1,1} + d_{s_3}^{1,1} + 2d_{s_5}^{1,1} + d_{s_6}^{1,1} - 3d_{s_{33}}^{2,2} - 3d_{s_{34}}^{2,2} - 3d_{s_{35}}^{2,2} - d_{s_{15}}^{2,2} - d_{s_{16}}^{2,2} - d_{s_{17}}^{2,2} - d_{s_{18}}^{2,2} + 4d_{s_{11}}^{1,1} + 3d_{s_{12}}^{1,1} + 2d_{s_{13}}^{1,1} + d_{s_{14}}^{1,1} + 3d_{s_{15}}^{1,1} + 2d_{s_{16}}^{1,1} + d_{s_{17}}^{1,1} - d_{s_{23}}^{2,2} + 2d_{s_{19}}^{1,1} + d_{s_{20}}^{1,1} + d_{s_{30}}^{1,1} + d_{s_{22}}^{1,1} + 3d_{s_{24}}^{1,1} + 2d_{s_{25}}^{1,1} + d_{s_{26}}^{1,1} + 2d_{s_{27}}^{1,1} - 2d_{s_{24}}^{2,2} + d_{s_8}^{1,1}$,

$d_{s_{26}}^{0,1} = -d_{s_{30}}^{0,1} - d_{s_{28}}^{0,1} - 2d_{s_{25}}^{0,1} - 3d_{s_{12}}^{0,1} - 2d_{s_{19}}^{0,1} - d_{s_{17}}^{0,1} - d_{s_{20}}^{0,1} - 2d_{s_{13}}^{0,1} - 3d_{s_{24}}^{0,1} + d_{s_{11}}^{1,2} + d_{s_{12}}^{1,2} + d_{s_{13}}^{1,2} + d_{s_{14}}^{1,2} + d_{s_{16}}^{1,2} + d_{s_{17}}^{1,2} + d_{s_{18}}^{1,2} - 2d_{s_{27}}^{0,1} + d_{s_{19}}^{1,2} + d_{s_{20}}^{1,2} - d_{s_{22}}^{0,1} + d_{s_{21}}^{1,2} + d_{s_{22}}^{1,2} + d_{s_{23}}^{1,2} + 3d_{s_{32}}^{1,2} + 3d_{s_{33}}^{1,2} -$

$$2d_{s_{16}}^{0,1} - d_{s_{34}}^{0,1} + d_{s_{15}}^{1,2} + 4d_{s_{36}}^{1,2} - 3d_{s_{15}}^{0,1} + 2d_{s_{27}}^{1,2} + 2d_{s_{28}}^{1,2} + 2d_{s_{29}}^{1,2} + 2d_{s_{30}}^{1,2} + 2d_{s_{31}}^{1,2} + 3d_{s_{34}}^{1,2} + 3d_{s_{35}}^{1,2} + 2d_{s_{24}}^{1,2} + 2d_{s_{25}}^{1,2} + 2d_{s_{26}}^{1,2} - 3d_{s_1}^{0,1} - 2d_{s_2}^{0,1} - d_{s_3}^{0,1} - 2d_{s_5}^{0,1} - d_{s_{14}}^{0,1} - d_{s_6}^{0,1} - d_{s_8}^{0,1} - 4d_{s_{11}}^{0,1} - 2d_{s_{32}}^{0,1} - d_{s_{33}}^{0,1} - d_{s_{36}}^{0,1},$$

$$d_{s_{29}}^{2,0} = -3d_{s_{32}}^{2,0} - 3d_{s_{33}}^{2,0} - 2d_{s_{29}}^{1,1} - d_{s_{33}}^{1,1} - d_{s_{35}}^{1,1} - d_{s_{12}}^{2,0} - d_{s_{13}}^{2,0} - d_{s_{14}}^{2,0} - d_{s_{28}}^{1,1} + 4d_{s_{11}}^{1,2} + 3d_{s_{12}}^{1,2} + 2d_{s_{13}}^{1,2} + d_{s_{14}}^{1,2} + 2d_{s_{16}}^{1,2} + d_{s_{17}}^{1,2} + 2d_{s_{19}}^{1,2} + d_{s_{20}}^{1,2} + d_{s_{22}}^{1,2} + 2d_{s_{32}}^{1,2} + d_{s_{33}}^{1,2} + d_{s_1}^{2,0} + d_{s_2}^{2,0} + d_{s_3}^{2,0} + d_{s_4}^{2,0} + 2d_{s_5}^{2,0} + 2d_{s_6}^{2,0} + 2d_{s_7}^{2,0} + 3d_{s_1}^{1,2} + 2d_{s_2}^{1,2} + d_{s_3}^{1,2} + 2d_{s_5}^{1,2} + d_{s_6}^{1,2} - 2d_{s_{24}}^{2,0} - 2d_{s_{25}}^{2,0} - 2d_{s_{26}}^{2,0} - d_{s_1}^{1,1} - 2d_{s_2}^{1,1} - 3d_{s_3}^{1,1} - 4d_{s_4}^{1,1} - d_{s_5}^{1,1} - 2d_{s_6}^{1,1} + d_{s_{19}}^{2,0} + d_{s_{20}}^{2,0} + d_{s_{21}}^{2,0} + 2d_{s_{22}}^{2,0} + 2d_{s_{23}}^{2,0} - d_{s_{27}}^{2,0} + d_{s_8}^{1,2} + 3d_{s_{15}}^{1,2} - 2d_{s_{34}}^{2,0} - d_{s_{12}}^{1,1} - 2d_{s_{13}}^{1,1} - 3d_{s_{14}}^{1,1} - d_{s_{16}}^{1,1} - 2d_{s_{17}}^{1,1} - d_{s_{31}}^{1,1} - 3d_{s_{18}}^{1,1} - 4d_{s_{36}}^{2,0} + d_{s_{36}}^{1,2} - d_{s_{20}}^{1,1} - 2d_{s_{21}}^{1,1} - d_{s_{23}}^{1,1} - d_{s_{25}}^{1,1} - 2d_{s_{26}}^{1,1} - d_{s_{11}}^{2,0} - 3d_{s_7}^{1,1} - d_{s_8}^{1,1} - 2d_{s_9}^{1,1} - 2d_{s_{35}}^{2,0} - d_{s_{28}}^{2,0} + 2d_{s_{27}}^{1,2} + d_{s_{28}}^{1,2} + d_{s_{30}}^{1,2} + d_{s_{34}}^{1,2} + 3d_{s_8}^{2,0} + 3d_{s_9}^{2,0} + 4d_{s_{10}}^{2,0} + 3d_{s_{24}}^{1,2} + 2d_{s_{25}}^{1,2} + d_{s_{26}}^{1,2} - d_{s_{10}}^{1,1},$$

$$d_{s_{33}}^{0,2} = d_{s_{34}}^{1,1} + d_{s_{29}}^{1,1} - 2d_{s_9}^{0,2} - d_{s_{10}}^{0,2} - 4d_{s_4}^{0,2} + d_{s_{35}}^{1,1} + d_{s_{28}}^{1,1} - 2d_{s_{21}}^{0,2} - d_{s_1}^{0,2} - 2d_{s_2}^{0,2} - 3d_{s_3}^{0,2} - d_{s_{25}}^{0,2} - 2d_{s_{26}}^{0,2} + d_{s_1}^{1,1} + d_{s_2}^{1,1} + d_{s_3}^{1,1} + d_{s_4}^{1,1} + 2d_{s_5}^{1,1} + 2d_{s_6}^{1,1} - d_{s_{31}}^{0,2} - d_{s_{35}}^{0,2} - d_{s_{12}}^{0,2} - 2d_{s_{13}}^{0,2} - 3d_{s_{14}}^{0,2} - d_{s_{16}}^{0,2} - 2d_{s_{17}}^{0,2} - d_{s_{20}}^{0,2} - d_{s_{23}}^{0,2} + d_{s_{15}}^{1,1} + d_{s_{16}}^{1,1} + d_{s_{17}}^{1,1} + 2d_{s_{31}}^{1,1} + d_{s_{18}}^{1,1} + 2d_{s_{19}}^{1,1} + 2d_{s_{20}}^{1,1} + 2d_{s_{30}}^{1,1} + 2d_{s_{21}}^{1,1} + 3d_{s_{22}}^{1,1} + 3d_{s_{23}}^{1,1} + d_{s_{27}}^{1,1} - 3d_{s_7}^{0,2} - d_{s_5}^{0,2} - 2d_{s_6}^{0,2} - d_{s_8}^{0,2} + 2d_{s_7}^{1,1} + 3d_{s_8}^{1,1} + 3d_{s_9}^{1,1} - 3d_{s_{18}}^{0,2} - d_{s_{28}}^{0,2} + 4d_{s_{10}}^{1,1} - 2d_{s_{29}}^{0,2},$$

$$d_{s_{15}}^{1,0} = d_{s_{36}}^{1,0} - 2d_{s_3}^{1,0} - 2d_{s_{11}}^{1,0} + d_{s_{34}}^{1,1} - 2d_{s_{14}}^{1,0} + d_{s_{10}}^{1,0} + d_{s_{36}}^{1,1} + d_{s_{35}}^{1,1} - d_{s_{25}}^{1,0} - d_{s_{17}}^{1,0} - d_{s_{18}}^{1,0} - 2d_{s_{11}}^{1,2} - 2d_{s_{12}}^{1,2} - 2d_{s_{13}}^{1,2} - 2d_{s_{14}}^{1,2} - d_{s_{16}}^{1,2} - d_{s_{17}}^{1,2} - d_{s_{18}}^{1,2} + d_{s_{22}}^{1,2} + d_{s_{23}}^{1,2} - d_{s_5}^{1,0} - d_{s_{26}}^{1,0} - 2d_{s_{13}}^{1,0} - 2d_{s_{12}}^{1,0} - 2d_{s_1}^{1,2} - 2d_{s_2}^{1,2} - 2d_{s_3}^{1,2} - 2d_{s_4}^{1,2} - d_{s_5}^{1,2} - d_{s_6}^{1,2} - 2d_{s_1}^{1,1} - 2d_{s_2}^{1,1} - 2d_{s_3}^{1,1} - 2d_{s_4}^{1,1} - d_{s_5}^{1,1} - d_{s_6}^{1,1} - d_{s_7}^{1,0} - d_{s_{16}}^{1,0} + d_{s_{22}}^{1,0} + d_{s_{23}}^{1,0} + d_{s_{31}}^{1,0} + d_{s_{34}}^{1,0} + d_{s_{35}}^{1,0} + d_{s_{30}}^{1,0} - 2d_{s_2}^{1,0} - d_{s_7}^{1,2} + d_{s_{10}}^{1,2} - d_{s_{15}}^{1,2} - 2d_{s_{11}}^{1,1} - 2d_{s_1}^{1,0} - 2d_{s_{12}}^{1,1} - 2d_{s_{13}}^{1,1} - 2d_{s_{14}}^{1,1} - d_{s_{15}}^{1,1} - d_{s_{16}}^{1,1} - d_{s_{17}}^{1,1} + d_{s_{31}}^{1,1} - d_{s_{18}}^{1,1} + d_{s_{36}}^{1,2} + d_{s_{30}}^{1,1} + d_{s_{22}}^{1,1} + d_{s_{23}}^{1,1} - d_{s_{24}}^{1,1} - d_{s_6}^{1,0} - d_{s_{25}}^{1,1} - d_{s_{26}}^{1,1} - d_{s_7}^{1,1} - d_{s_{24}}^{1,0} + d_{s_{30}}^{1,2} + d_{s_{31}}^{1,2} + d_{s_{34}}^{1,2} + 1/6 + d_{s_{35}}^{1,2} - d_{s_{24}}^{1,2} - 2d_{s_4}^{1,0} - d_{s_{25}}^{1,2} - d_{s_{26}}^{1,2} + d_{s_{10}}^{1,1},$$

$$d_{s_{28}}^{1,0} = -4d_{s_{36}}^{1,0} + 5d_{s_3}^{1,0} - d_{s_{33}}^{1,0} + 2d_{s_{11}}^{1,0} - 3d_{s_{34}}^{1,1} + 5d_{s_{14}}^{1,0} - 3d_{s_{10}}^{1,0} - 3d_{s_{36}}^{1,1} - 3d_{s_{35}}^{1,1} + d_{s_{25}}^{1,0} + 2d_{s_{17}}^{1,0} + 3d_{s_{18}}^{1,0} - 2d_{s_{19}}^{1,0} - d_{s_{20}}^{1,0} + 6d_{s_{11}}^{1,2} + 6d_{s_{12}}^{1,2} + 6d_{s_{13}}^{1,2} + 6d_{s_{14}}^{1,2} + 2d_{s_{31}}^{2,1} + 3d_{s_{32}}^{2,1} + 3d_{s_{33}}^{2,1} + 3d_{s_{34}}^{2,1} + 3d_{s_{35}}^{2,1} + 4d_{s_{36}}^{2,1} + 3d_{s_{16}}^{1,2} + 3d_{s_{17}}^{1,2} + 3d_{s_{18}}^{1,2} - 3d_{s_{22}}^{1,2} - 3d_{s_{23}}^{1,2} + d_{s_{12}}^{2,1} + d_{s_{14}}^{2,1} + d_{s_{15}}^{2,1} + d_{s_{16}}^{2,1} + d_{s_{17}}^{2,1} - d_{s_8}^{1,0} + d_{s_5}^{1,0} + 2d_{s_{26}}^{1,0} + 4d_{s_{13}}^{1,0} - 2d_{s_{32}}^{1,0} + 3d_{s_{12}}^{1,0} + 6d_{s_1}^{1,2} + 6d_{s_2}^{1,2} + 6d_{s_3}^{1,2} + 6d_{s_4}^{1,2} + 3d_{s_5}^{1,2} + 3d_{s_6}^{1,2} - 2d_{s_{27}}^{1,0} + 6d_{s_1}^{1,1} + 6d_{s_2}^{1,1} + 6d_{s_3}^{1,1} + 6d_{s_4}^{1,1} + 3d_{s_5}^{1,1} + 3d_{s_6}^{1,1} + 3d_{s_7}^{1,0} + d_{s_{16}}^{1,0} + d_{s_{11}}^{2,1} - 4d_{s_{22}}^{1,0} - 3d_{s_{23}}^{1,0} - 3d_{s_{31}}^{1,0} - 4d_{s_{34}}^{1,0} - 3d_{s_{35}}^{1,0} - 4d_{s_{30}}^{1,0} + 4d_{s_2}^{1,0} + 3d_{s_7}^{1,2} - 3d_{s_{10}}^{1,2} + 3d_{s_{15}}^{1,2} + 6d_{s_{11}}^{1,1} + 3d_{s_1}^{1,0} + 6d_{s_{12}}^{1,1} + 6d_{s_{13}}^{1,1} + 6d_{s_{14}}^{1,1} + 3d_{s_{15}}^{1,1} + 3d_{s_{16}}^{1,1} + 3d_{s_{17}}^{1,1} - 3d_{s_{31}}^{1,1} + 3d_{s_{18}}^{1,1} + 2d_{s_{25}}^{2,1} + 2d_{s_{26}}^{2,1} + 2d_{s_{27}}^{2,1} + 2d_{s_{28}}^{2,1} + 2d_{s_{30}}^{2,1} + d_{s_{13}}^{2,1} - 3d_{s_{36}}^{1,2} - 3d_{s_{30}}^{1,1} - 3d_{s_{22}}^{1,1} - 3d_{s_{23}}^{1,1} + 3d_{s_{24}}^{1,1} + 2d_{s_6}^{1,0} + 3d_{s_{25}}^{1,1} + 3d_{s_{26}}^{1,1} + 3d_{s_7}^{1,1} - 3d_{s_{30}}^{1,2} - 3d_{s_{31}}^{1,2} - 3d_{s_{34}}^{1,2} + d_{s_{18}}^{2,1} + d_{s_{19}}^{2,1} + d_{s_{20}}^{2,1} + d_{s_{21}}^{2,1} + d_{s_{22}}^{2,1} + d_{s_{23}}^{2,1} + 2d_{s_{24}}^{2,1} - 1/2 - 3d_{s_{35}}^{1,2} + 3d_{s_{24}}^{1,2} + 6d_{s_4}^{1,0} + 3d_{s_{25}}^{1,2} + 3d_{s_{26}}^{1,2} - 3d_{s_{10}}^{1,1} + 2d_{s_{29}}^{2,1},$$

$$d_{s_{29}}^{1,0} = 2d_{s_{36}}^{1,0} - 4d_{s_3}^{1,0} - d_{s_{11}}^{1,0} + d_{s_{34}}^{1,1} - d_{s_{29}}^{1,1} - 4d_{s_{14}}^{1,0} + d_{s_{10}}^{1,0} + d_{s_{36}}^{1,1} - d_{s_{33}}^{1,1} + d_{s_{35}}^{1,1} - d_{s_{25}}^{1,0} - 2d_{s_{17}}^{1,0} - 3d_{s_{18}}^{1,0} + d_{s_{19}}^{1,0} - d_{s_{28}}^{1,1} - 5d_{s_{11}}^{1,2} - 5d_{s_{12}}^{1,2} - 5d_{s_{13}}^{1,2} - 5d_{s_{14}}^{1,2} - 2d_{s_{31}}^{2,1} - 3d_{s_{32}}^{2,1} - 3d_{s_{33}}^{2,1} - 3d_{s_{34}}^{2,1} - 3d_{s_{35}}^{2,1} - 4d_{s_{36}}^{2,1} - 3d_{s_{16}}^{1,2} - 3d_{s_{17}}^{1,2} - 3d_{s_{18}}^{1,2} - d_{s_{19}}^{1,2} - d_{s_{20}}^{1,2} - d_{s_{21}}^{1,2} + d_{s_{22}}^{1,2} + d_{s_{23}}^{1,2} - d_{s_{12}}^{2,1} - d_{s_{14}}^{2,1} - d_{s_{15}}^{2,1} - d_{s_{16}}^{2,1} - d_{s_{17}}^{2,1} - d_{s_9}^{1,0} - d_{s_{32}}^{1,1} - d_{s_5}^{1,0} - d_{s_{32}}^{1,2} - d_{s_{33}}^{1,2} - 2d_{s_{26}}^{1,0} - 3d_{s_{13}}^{1,0} + d_{s_{32}}^{1,0} - 2d_{s_{12}}^{1,0} - 5d_{s_1}^{1,2} - 5d_{s_2}^{1,2} - 5d_{s_3}^{1,2} - 5d_{s_4}^{1,2} - 3d_{s_5}^{1,2} - 3d_{s_6}^{1,2} + d_{s_{27}}^{1,0} - 5d_{s_1}^{1,1} - 5d_{s_2}^{1,1} - 5d_{s_3}^{1,1} - 5d_{s_4}^{1,1} - 3d_{s_5}^{1,1} - 3d_{s_6}^{1,1} - 3d_{s_7}^{1,0} - d_{s_{16}}^{1,0} - d_{s_{11}}^{2,1} - d_{s_{21}}^{1,0} + 2d_{s_{22}}^{1,0} + d_{s_{23}}^{1,0} + d_{s_{31}}^{1,0} + 2d_{s_{34}}^{1,0} + d_{s_{35}}^{1,0} + 2d_{s_{30}}^{1,0} - 3d_{s_2}^{1,0} - 3d_{s_7}^{1,2} - d_{s_8}^{1,2} - d_{s_9}^{1,2} + d_{s_{10}}^{1,2} - 3d_{s_{15}}^{1,2} - 5d_{s_{11}}^{1,1} - 2d_{s_1}^{1,0} - 5d_{s_{12}}^{1,1} - 5d_{s_{13}}^{1,1} - 5d_{s_{14}}^{1,1} - 3d_{s_{15}}^{1,1} - 3d_{s_{16}}^{1,1} - 3d_{s_{17}}^{1,1} + d_{s_{31}}^{1,1} - 3d_{s_{18}}^{1,1} - 2d_{s_{25}}^{2,1} - 2d_{s_{26}}^{2,1} - 2d_{s_{27}}^{2,1} - 2d_{s_{28}}^{2,1} - 2d_{s_{30}}^{2,1} - d_{s_{13}}^{2,1} + d_{s_{36}}^{1,2} - d_{s_{19}}^{1,1} - d_{s_{20}}^{1,1} + d_{s_{30}}^{1,1} - d_{s_{21}}^{1,1} + d_{s_{22}}^{1,1} + d_{s_{23}}^{1,1} - 3d_{s_{24}}^{1,1} - 2d_{s_6}^{1,0} - 3d_{s_{25}}^{1,1} - 3d_{s_{26}}^{1,1} - d_{s_{27}}^{1,1} - 3d_{s_7}^{1,1} - d_{s_8}^{1,1} - d_{s_9}^{1,1} - d_{s_{27}}^{1,2} - d_{s_{28}}^{1,2} - d_{s_{29}}^{1,2} + d_{s_{30}}^{1,2} + d_{s_{31}}^{1,2} + d_{s_{34}}^{1,2} - d_{s_{18}}^{2,1} - d_{s_{19}}^{2,1} - d_{s_{20}}^{2,1} - d_{s_{21}}^{2,1} - d_{s_{22}}^{2,1} - d_{s_{23}}^{2,1} - 2d_{s_{24}}^{2,1} + d_{s_{35}}^{1,2} - 3d_{s_{24}}^{1,2} - 5d_{s_4}^{1,0} - 3d_{s_{25}}^{1,2} - 3d_{s_{26}}^{1,2} + d_{s_{10}}^{1,1} - 2d_{s_{29}}^{2,1} + 2/3,$$

$$d_{s_{36}}^{0,0} = 3d_{s_{34}}^{1,1} + 2d_{s_{29}}^{1,1} + 4d_{s_{36}}^{1,1} + 3d_{s_{33}}^{1,1} + 3d_{s_{35}}^{1,1} + 2d_{s_{28}}^{1,1} - 3d_{s_1}^{0,0} - 2d_{s_2}^{0,0} - d_{s_3}^{0,0} - 2d_{s_5}^{0,0} - d_{s_6}^{0,0} - d_{s_8}^{0,0} - 4d_{s_{11}}^{0,0} + 3d_{s_{32}}^{1,1} - 3d_{s_{12}}^{0,0} - 2d_{s_{13}}^{0,0} - d_{s_{14}}^{0,0} - 3d_{s_{15}}^{0,0} - 2d_{s_{16}}^{0,0} - d_{s_{17}}^{0,0} - 2d_{s_{19}}^{0,0} - d_{s_{20}}^{0,0} - d_{s_{22}}^{0,0} - 3d_{s_{24}}^{0,0} - 2d_{s_{25}}^{0,0} - d_{s_{26}}^{0,0} - 2d_{s_{27}}^{0,0} - d_{s_{28}}^{0,0} - d_{s_{30}}^{0,0} - 2d_{s_{32}}^{0,0} - d_{s_{33}}^{0,0} - d_{s_{34}}^{0,0} + d_{s_{11}}^{1,1} + d_{s_{12}}^{1,1} + d_{s_{13}}^{1,1} + d_{s_{14}}^{1,1} + d_{s_{15}}^{1,1} + d_{s_{16}}^{1,1} + d_{s_{17}}^{1,1} + 2d_{s_{31}}^{1,1} + d_{s_{18}}^{1,1} + d_{s_{19}}^{1,1} + d_{s_{20}}^{1,1} + 2d_{s_{30}}^{1,1} + d_{s_{21}}^{1,1} + d_{s_{22}}^{1,1} + d_{s_{23}}^{1,1} + 2d_{s_{24}}^{1,1} + 2d_{s_{25}}^{1,1} + 2d_{s_{26}}^{1,1} + 2d_{s_{27}}^{1,1},$$

$$d_{s_{34}}^{0,2} = -d_{s_{36}}^{1,0} - d_{s_{33}}^{1,0} - d_{s_{11}}^{1,0} - 8d_{s_{34}}^{1,1} - 4d_{s_{29}}^{1,1} - d_{s_{14}}^{1,0} - 3d_{s_9}^{0,2} - 4d_{s_{10}}^{0,2} + 3d_{s_{10}}^{1,0} - d_{s_4}^{0,2} - 12d_{s_{36}}^{1,1} - 8d_{s_{33}}^{1,1} - 8d_{s_{35}}^{1,1} - d_{s_{25}}^{1,0} + d_{s_{19}}^{1,0} + d_{s_{20}}^{1,0} + 12d_{s_4}^{0,0} + 9d_{s_7}^{0,0} + 6d_{s_9}^{0,0} - 4d_{s_{28}}^{1,1} + 3d_{s_{10}}^{0,0} + d_{s_{12}}^{1,2} + 2d_{s_{13}}^{1,2} + 3d_{s_{14}}^{1,2} + 8d_{s_{18}}^{0,0} + 15d_{s_1}^{0,0} + 14d_{s_2}^{0,0} + 13d_{s_3}^{0,0} - 2d_{s_{21}}^{0,2} + 2d_{s_{31}}^{2,1} + 3d_{s_{32}}^{2,1} + 3d_{s_{33}}^{2,1} + 3d_{s_{34}}^{2,1} + 3d_{s_{35}}^{2,1} + 4d_{s_{36}}^{2,1} - 2d_{s_{16}}^{1,2} - d_{s_{17}}^{1,2} - 6d_{s_{19}}^{1,2} - 5d_{s_{20}}^{1,2} - 4d_{s_{21}}^{1,2} - 9d_{s_{22}}^{1,2} - 8d_{s_{23}}^{1,2} + d_{s_{12}}^{2,1} + d_{s_{14}}^{2,1} + d_{s_{15}}^{2,1} + d_{s_{16}}^{2,1} + d_{s_{17}}^{2,1} + 11d_{s_5}^{0,0} + 2d_{s_9}^{1,0} + 10d_{s_6}^{0,0} + 2d_{s_8}^{1,0} + 7d_{s_8}^{0,0} + 15d_{s_{11}}^{0,0} - 8d_{s_{32}}^{1,1} + 14d_{s_{12}}^{0,0} + 13d_{s_{13}}^{0,0} + 4d_{s_{29}}^{0,0} + 5d_{s_{21}}^{0,0} + 2d_{s_{23}}^{0,0} + d_{s_5}^{1,0} + 2d_{s_{32}}^{1,2} + 3d_{s_{33}}^{1,2} - d_{s_1}^{0,2} - d_{s_2}^{0,2} - d_{s_3}^{0,2} - 3d_{s_{22}}^{0,2} - d_{s_{26}}^{1,0} + 12d_{s_{14}}^{0,0} - d_{s_{13}}^{1,0} + 11d_{s_{15}}^{0,0} + 10d_{s_{16}}^{0,0} + 9d_{s_{17}}^{0,0} + 7d_{s_{19}}^{0,0} + 6d_{s_{20}}^{0,0} - d_{s_{32}}^{1,0} + 3d_{s_{22}}^{0,0} - d_{s_{12}}^{1,0} - 3d_{s_1}^{1,2} - 2d_{s_2}^{1,2} - d_{s_3}^{1,2} - 6d_{s_5}^{1,2} - 5d_{s_6}^{1,2} + 4d_{s_1}^{1,1} + 4d_{s_2}^{1,1} + 4d_{s_3}^{1,1} + 4d_{s_4}^{1,1} + 4d_{s_5}^{1,1} + 4d_{s_6}^{1,1} + 10d_{s_{24}}^{0,0} + 9d_{s_{25}}^{0,0} + 8d_{s_{26}}^{0,0} + 6d_{s_{27}}^{0,0} + 5d_{s_{28}}^{0,0} + 2d_{s_{30}}^{0,0} - 2d_{s_{31}}^{0,2} - d_{s_{35}}^{0,2} + d_{s_7}^{1,0} + 5d_{s_{32}}^{0,0} + 4d_{s_{33}}^{0,0} + d_{s_{34}}^{0,0} + d_{s_{11}}^{2,1} + d_{s_{21}}^{1,0} + 2d_{s_{22}}^{1,0} + 2d_{s_{23}}^{1,0} - d_{s_{15}}^{0,2} - d_{s_{16}}^{0,2} - d_{s_{17}}^{0,2} - 2d_{s_{19}}^{0,2} - 2d_{s_{20}}^{0,2} + d_{s_{31}}^{1,0} + d_{s_{30}}^{1,0} - 3d_{s_{23}}^{0,2} - 4d_{s_7}^{1,2} - 9d_{s_8}^{1,2} - 8d_{s_9}^{1,2} - 12d_{s_{10}}^{1,2} - 3d_{s_{15}}^{1,2} - 4d_{s_{31}}^{1,1} + d_{s_{31}}^{0,0} + 2d_{s_{25}}^{2,1} + 2d_{s_{26}}^{2,1} + 2d_{s_{27}}^{2,1} + 2d_{s_{28}}^{2,1} + 2d_{s_{30}}^{2,1} + d_{s_{13}}^{2,1} + 3d_{s_{36}}^{1,2} - 4d_{s_{30}}^{1,1} - 4d_{s_{24}}^{1,1} + d_{s_6}^{1,0} - 4d_{s_{25}}^{1,1} - 4d_{s_{26}}^{1,1} - 4d_{s_{27}}^{1,1} - 2d_{s_7}^{0,2} - 2d_{s_5}^{0,2} - 2d_{s_6}^{0,2} - 3d_{s_8}^{0,2} + 4d_{s_7}^{1,1} + 4d_{s_8}^{1,1} + 4d_{s_9}^{1,1} - d_{s_{18}}^{0,2} - d_{s_{27}}^{0,2} - d_{s_{24}}^{1,0} - d_{s_{28}}^{0,2} - 2d_{s_{30}}^{0,2} - 2d_{s_{27}}^{1,2} - d_{s_{28}}^{1,2} - 5d_{s_{30}}^{1,2} - 4d_{s_{31}}^{1,2} - d_{s_{34}}^{1,2} + d_{s_{18}}^{2,1} + d_{s_{19}}^{2,1} + d_{s_{20}}^{2,1} + d_{s_{21}}^{2,1} + d_{s_{22}}^{2,1} + d_{s_{23}}^{2,1} + 2d_{s_{24}}^{2,1} - 1/2 + d_{s_{24}}^{1,2} + 2d_{s_{25}}^{1,2} + 3d_{s_{26}}^{1,2} + 4d_{s_{10}}^{1,1} + 2d_{s_{29}}^{2,1} - d_{s_{29}}^{0,2},$$

$$d_{s_6}^{2,1} = 9d_{s_{34}}^{1,1} + d_{s_{14}}^{2,2} + 7d_{s_{29}}^{1,1} - 2d_{s_{11}}^{2,2} + d_{s_{20}}^{2,2} + 2d_{s_{21}}^{2,2} - 2d_{s_1}^{2,2} - d_{s_5}^{2,2} - 2d_{s_7}^{2,1} + 12d_{s_{36}}^{1,1} + 2d_{s_{26}}^{2,2} + d_{s_7}^{2,2} + 9d_{s_{33}}^{1,1} - d_{s_{12}}^{2,2} + 10d_{s_{35}}^{1,1} + 2d_{s_{32}}^{2,2} - 12d_{s_4}^{0,0} - 8d_{s_7}^{0,0} - 4d_{s_9}^{0,0} - d_{s_2}^{2,1} - 2d_{s_3}^{2,1} + 6d_{s_{28}}^{1,1} + 4d_{s_{36}}^{2,2} - 3d_{s_4}^{2,1} + d_{s_{11}}^{1,2} + d_{s_{12}}^{1,2} + d_{s_{13}}^{1,2} + d_{s_{14}}^{1,2} - 8d_{s_{18}}^{0,0} - 12d_{s_1}^{0,0} - 12d_{s_2}^{0,0} - 12d_{s_3}^{0,0} + d_{s_{22}}^{2,2} + d_{s_{27}}^{2,2} + d_{s_{25}}^{2,2} + d_{s_{32}}^{2,1} + d_{s_{34}}^{2,1} + d_{s_{36}}^{2,1} + 4d_{s_{16}}^{1,2} + 4d_{s_{17}}^{1,2} + 4d_{s_{18}}^{1,2} + 7d_{s_{19}}^{1,2} + 7d_{s_{20}}^{1,2} + 7d_{s_{21}}^{1,2} + 10d_{s_{22}}^{1,2} + 10d_{s_{23}}^{1,2} - 2d_{s_{14}}^{2,1} + d_{s_{15}}^{2,1} - d_{s_{17}}^{2,1} - 8d_{s_5}^{0,0} - 8d_{s_6}^{0,0} - 4d_{s_8}^{0,0} - 12d_{s_{11}}^{0,0} + 8d_{s_{32}}^{1,1} - 12d_{s_{12}}^{0,0} - 12d_{s_{13}}^{0,0} - 4d_{s_{29}}^{0,0} - 4d_{s_{21}}^{0,0} + d_{s_{32}}^{1,2} + d_{s_{33}}^{1,2} + 2d_{s_{28}}^{2,2} + 3d_{s_{29}}^{2,2} + 2d_{s_{30}}^{2,2} + 3d_{s_{31}}^{2,2} - d_{s_9}^{2,1} - 12d_{s_{14}}^{0,0} - 8d_{s_{15}}^{0,0} - 8d_{s_{16}}^{0,0} - 8d_{s_{17}}^{0,0} - 4d_{s_{19}}^{0,0} - 4d_{s_{20}}^{0,0} + 4d_{s_1}^{1,2} + 4d_{s_2}^{1,2} + 4d_{s_3}^{1,2} + 4d_{s_4}^{1,2} + 7d_{s_5}^{1,2} + 7d_{s_6}^{1,2} - 2d_{s_1}^{1,1} - d_{s_2}^{1,1} + d_{s_4}^{1,1} - d_{s_5}^{1,1} - d_{s_2}^{2,2} - 8d_{s_{24}}^{0,0} - 8d_{s_{25}}^{0,0} - 8d_{s_{26}}^{0,0} - 4d_{s_{27}}^{0,0} - 4d_{s_{28}}^{0,0} + 3d_{s_{33}}^{2,2} + 3d_{s_{34}}^{2,2} + 4d_{s_{35}}^{2,2} - d_{s_{15}}^{2,2} + d_{s_{17}}^{2,2} + 2d_{s_{18}}^{2,2} - 4d_{s_{32}}^{0,0} - 4d_{s_{33}}^{0,0} + d_{s_{11}}^{2,1} + d_{s_4}^{2,2} + d_{s_{10}}^{2,2} + 7d_{s_7}^{1,2} + 10d_{s_8}^{1,2} + 10d_{s_9}^{1,2} + 13d_{s_{10}}^{1,2} + 4d_{s_{15}}^{1,2} + d_{s_{12}}^{1,1} + 2d_{s_{13}}^{1,1} + 3d_{s_{14}}^{1,1} + d_{s_{15}}^{1,1} + 2d_{s_{16}}^{1,1} + 3d_{s_{17}}^{1,1} + 7d_{s_{31}}^{1,1} + 4d_{s_{18}}^{1,1} - d_{s_{26}}^{2,1} + d_{s_{27}}^{2,1} + d_{s_{30}}^{2,1} - d_{s_{13}}^{2,1} + 2d_{s_{23}}^{2,2} + d_{s_{36}}^{1,2} + 2d_{s_{19}}^{1,1} + 3d_{s_{20}}^{1,1} + 6d_{s_{30}}^{1,1} + 4d_{s_{21}}^{1,1} + 3d_{s_{22}}^{1,1} + 4d_{s_{23}}^{1,1} + 4d_{s_{24}}^{1,1} + 5d_{s_{25}}^{1,1} + 6d_{s_{26}}^{1,1} + 5d_{s_{27}}^{1,1} + d_{s_9}^{2,2} + d_{s_7}^{1,1} + d_{s_9}^{1,1} + 4d_{s_{27}}^{1,2} + 4d_{s_{28}}^{1,2} + 4d_{s_{29}}^{1,2} + 7d_{s_{30}}^{1,2} + 7d_{s_{31}}^{1,2} + 4d_{s_{34}}^{1,2} - 2d_{s_{18}}^{2,1} + d_{s_{19}}^{2,1} - d_{s_{21}}^{2,1} + d_{s_{22}}^{2,1} + d_{s_{24}}^{2,1} - 1/3 + 4d_{s_{35}}^{1,2} + d_{s_{24}}^{1,2} + d_{s_{25}}^{1,2} + d_{s_{26}}^{1,2} + d_{s_{10}}^{1,1} - d_{s_{29}}^{2,1},$$

$$d_{s_{31}}^{2,0} = d_{s_{32}}^{2,0} - d_{s_{30}}^{2,0} + 2d_{s_{34}}^{1,1} + d_{s_{33}}^{2,0} + 3d_{s_{29}}^{1,1} + 3d_{s_{36}}^{1,1} + 3d_{s_{33}}^{1,1} + 3d_{s_{35}}^{1,1} + d_{s_{12}}^{2,0} + d_{s_{13}}^{2,0} + d_{s_{14}}^{2,0} - 3d_{s_4}^{0,0} - 2d_{s_7}^{0,0} - d_{s_9}^{0,0} + 2d_{s_{28}}^{1,1} - d_{s_{11}}^{1,2} - d_{s_{12}}^{1,2} - d_{s_{13}}^{1,2} - d_{s_{14}}^{1,2} - 2d_{s_{18}}^{0,0} - 3d_{s_1}^{0,0} - 3d_{s_2}^{0,0} - 3d_{s_3}^{0,0} + d_{s_{19}}^{1,2} + d_{s_{20}}^{1,2} + d_{s_{21}}^{1,2} + 2d_{s_{22}}^{1,2} + 2d_{s_{23}}^{1,2} - 2d_{s_5}^{0,0} - 2d_{s_6}^{0,0} - d_{s_8}^{0,0} - 3d_{s_{11}}^{0,0} + 2d_{s_{32}}^{1,1} - 3d_{s_{12}}^{0,0} - 3d_{s_{13}}^{0,0} - d_{s_{29}}^{0,0} - d_{s_{21}}^{0,0} - d_{s_{32}}^{1,2} - d_{s_{33}}^{1,2} - 3d_{s_{14}}^{0,0} - 2d_{s_{15}}^{0,0} - 2d_{s_{16}}^{0,0} - 2d_{s_{17}}^{0,0} - d_{s_{19}}^{0,0} - d_{s_{20}}^{0,0} - d_{s_5}^{2,0} - d_{s_6}^{2,0} - d_{s_7}^{2,0} + d_{s_5}^{1,2} + d_{s_6}^{1,2} + d_{s_{24}}^{2,0} + d_{s_{25}}^{2,0} + d_{s_{26}}^{2,0} + d_{s_2}^{1,1} + 2d_{s_3}^{1,1} + 3d_{s_4}^{1,1} + d_{s_6}^{1,1} - 2d_{s_{24}}^{0,0} - 2d_{s_{25}}^{0,0} - 2d_{s_{26}}^{0,0} - d_{s_{27}}^{0,0} - d_{s_{28}}^{0,0} - d_{s_{32}}^{0,0} - d_{s_{33}}^{0,0} - d_{s_{19}}^{2,0} - d_{s_{20}}^{2,0} - d_{s_{21}}^{2,0} - 2d_{s_{22}}^{2,0} - 2d_{s_{23}}^{2,0} + d_{s_7}^{1,2} + 2d_{s_8}^{1,2} + 2d_{s_9}^{1,2} + 3d_{s_{10}}^{1,2} + d_{s_{12}}^{1,1} + 2d_{s_{13}}^{1,1} + 3d_{s_{14}}^{1,1} + d_{s_{16}}^{1,1} + 2d_{s_{17}}^{1,1} + 2d_{s_{31}}^{1,1} + 3d_{s_{18}}^{1,1} + d_{s_{36}}^{2,0} - d_{s_{36}}^{1,2} + d_{s_{20}}^{1,1} + d_{s_{30}}^{1,1} + 2d_{s_{21}}^{1,1} + d_{s_{23}}^{1,1} + d_{s_{24}}^{1,1} + 2d_{s_{25}}^{1,1} + 3d_{s_{26}}^{1,1} + d_{s_{27}}^{1,1} + d_{s_{11}}^{2,0} + 2d_{s_7}^{1,1} + d_{s_9}^{1,1} + d_{s_{30}}^{1,2} + d_{s_{31}}^{1,2} - 2d_{s_8}^{2,0} - 2d_{s_9}^{2,0} - 3d_{s_{10}}^{2,0} - d_{s_{24}}^{1,2} - d_{s_{25}}^{1,2} - d_{s_{26}}^{1,2},$$

$$d_{s_{35}}^{0,0} = -4d_{s_4}^{0,0} - 3d_{s_7}^{0,0} - 2d_{s_9}^{0,0} - d_{s_{10}}^{0,0} - 3d_{s_{18}}^{0,0} - d_{s_1}^{0,0} - 2d_{s_2}^{0,0} - 3d_{s_3}^{0,0} + d_{s_{16}}^{1,2} + d_{s_{17}}^{1,2} + d_{s_{18}}^{1,2} + 2d_{s_{19}}^{1,2} + 2d_{s_{20}}^{1,2} + 2d_{s_{21}}^{1,2} + 3d_{s_{22}}^{1,2} + 3d_{s_{23}}^{1,2} - d_{s_5}^{0,0} - 2d_{s_6}^{0,0} - d_{s_8}^{0,0} - d_{s_{12}}^{0,0} - 2d_{s_{13}}^{0,0} - 2d_{s_{29}}^{0,0} - 2d_{s_{21}}^{0,0} - d_{s_{23}}^{0,0} - 3d_{s_{14}}^{0,0} - d_{s_{16}}^{0,0} - 2d_{s_{17}}^{0,0} - d_{s_{20}}^{0,0} + d_{s_1}^{1,2} + d_{s_2}^{1,2} + d_{s_3}^{1,2} + d_{s_4}^{1,2} + 2d_{s_5}^{1,2} + 2d_{s_6}^{1,2} - d_{s_{25}}^{0,0} - 2d_{s_{26}}^{0,0} - d_{s_{28}}^{0,0} - d_{s_{33}}^{0,0} + 2d_{s_7}^{1,2} + 3d_{s_8}^{1,2} + 3d_{s_9}^{1,2} + 4d_{s_{10}}^{1,2} + d_{s_{15}}^{1,2} - d_{s_{31}}^{0,0} + d_{s_{27}}^{1,2} + d_{s_{28}}^{1,2} + d_{s_{29}}^{1,2} + 2d_{s_{30}}^{1,2} + 2d_{s_{31}}^{1,2} + d_{s_{34}}^{1,2} + d_{s_{35}}^{1,2},$$

$$d_{s_3}^{2,2} = 21d_{s_{34}}^{1,1} + 2d_{s_{14}}^{2,2} + 18d_{s_{29}}^{1,1} - 4d_{s_{11}}^{2,2} + 2d_{s_{19}}^{2,2} + 4d_{s_{20}}^{2,2} + 6d_{s_{21}}^{2,2} - 5d_{s_1}^{2,2} - 2d_{s_5}^{2,2} + 29d_{s_{36}}^{1,1} + 6d_{s_{26}}^{2,2} + 2d_{s_7}^{2,2} + 22d_{s_{33}}^{1,1} - 2d_{s_{12}}^{2,2} + 25d_{s_{35}}^{1,1} + 8d_{s_{32}}^{2,2} - 24d_{s_4}^{0,0} - 18d_{s_7}^{0,0} - 12d_{s_9}^{0,0} + d_{s_8}^{2,2} + 14d_{s_{28}}^{1,1} - 6d_{s_{10}}^{0,0} + 14d_{s_{36}}^{2,2} + 4d_{s_{11}}^{1,2} + 4d_{s_{12}}^{1,2} + 4d_{s_{13}}^{1,2} + 4d_{s_{14}}^{1,2} - 16d_{s_{18}}^{0,0} - 30d_{s_1}^{0,0} - 28d_{s_2}^{0,0} - 26d_{s_3}^{0,0} + 5d_{s_{22}}^{2,2} + 5d_{s_{27}}^{2,2} + 2d_{s_{31}}^{2,1} + 4d_{s_{25}}^{2,2} + 3d_{s_{32}}^{2,1} + 3d_{s_{33}}^{2,1} + 3d_{s_{34}}^{2,1} + 3d_{s_{35}}^{2,1} + 4d_{s_{36}}^{2,1} + 9d_{s_{16}}^{1,2} + 9d_{s_{17}}^{1,2} + 9d_{s_{18}}^{1,2} + 14d_{s_{19}}^{1,2} + 14d_{s_{20}}^{1,2} + 14d_{s_{21}}^{1,2} + 19d_{s_{22}}^{1,2} + 19d_{s_{23}}^{1,2} + d_{s_{12}}^{2,1} + d_{s_{14}}^{2,1} + d_{s_{15}}^{2,1} + d_{s_{16}}^{2,1} + d_{s_{17}}^{2,1} - 22d_{s_5}^{0,0} - 20d_{s_6}^{0,0} - 14d_{s_8}^{0,0} - 30d_{s_{11}}^{0,0} + 18d_{s_{32}}^{1,1} - 28d_{s_{12}}^{0,0} - 26d_{s_{13}}^{0,0} - 8d_{s_{29}}^{0,0} - 10d_{s_{21}}^{0,0} - 4d_{s_{23}}^{0,0} + 2d_{s_{32}}^{1,2} + 2d_{s_{33}}^{1,2} + 7d_{s_{28}}^{2,2} + 9d_{s_{29}}^{2,2} + 8d_{s_{30}}^{2,2} + 10d_{s_{31}}^{2,2} - 24d_{s_{14}}^{0,0} - 22d_{s_{15}}^{0,0} - 20d_{s_{16}}^{0,0} - 18d_{s_{17}}^{0,0} - 14d_{s_{19}}^{0,0} - 12d_{s_{20}}^{0,0} - 6d_{s_{22}}^{0,0} + 10d_{s_1}^{1,2} + 10d_{s_2}^{1,2} + 10d_{s_3}^{1,2} + 10d_{s_4}^{1,2} + 15d_{s_5}^{1,2} + 15d_{s_6}^{1,2} - 8d_{s_1}^{1,1} - 4d_{s_2}^{1,1} + 4d_{s_4}^{1,1} - 5d_{s_5}^{1,1} - d_{s_6}^{1,1} - 3d_{s_2}^{2,2} - 20d_{s_{24}}^{0,0} - 18d_{s_{25}}^{0,0} - 16d_{s_{26}}^{0,0} -$$

$$12d_{s_{27}}^{0,0} - 10d_{s_{28}}^{0,0} - 4d_{s_{30}}^{0,0} + 10d_{s_{33}}^{2,2} + 11d_{s_{34}}^{2,2} + 13d_{s_{35}}^{2,2} - d_{s_{15}}^{2,2} + d_{s_{16}}^{2,2} + 3d_{s_{17}}^{2,2} + 5d_{s_{18}}^{2,2} - 10d_{s_{32}}^{0,0} -$$
$$8d_{s_{33}}^{0,0} - 2d_{s_{34}}^{0,0} + d_{s_{11}}^{2,1} + d_{s_{4}}^{2,2} + 4d_{s_{10}}^{2,2} + 15d_{s_{7}}^{1,2} + 20d_{s_{8}}^{1,2} + 20d_{s_{9}}^{1,2} + 25d_{s_{10}}^{1,2} + 9d_{s_{15}}^{1,2} - 4d_{s_{11}}^{1,1} + 4d_{s_{13}}^{1,1} +$$
$$8d_{s_{14}}^{1,1} - d_{s_{15}}^{1,1} + 3d_{s_{16}}^{1,1} + 7d_{s_{17}}^{1,1} + 17d_{s_{31}}^{1,1} + 11d_{s_{18}}^{1,1} - 2d_{s_{31}}^{0,0} + 2d_{s_{25}}^{2,1} + 2d_{s_{26}}^{2,1} + 2d_{s_{27}}^{2,1} + 2d_{s_{28}}^{2,1} + 2d_{s_{30}}^{2,1} +$$
$$d_{s_{13}}^{2,1} + 7d_{s_{23}}^{2,2} + d_{s_{36}}^{1,2} + 2d_{s_{19}}^{1,1} + 6d_{s_{20}}^{1,1} + 13d_{s_{30}}^{1,1} + 10d_{s_{21}}^{1,1} + 5d_{s_{22}}^{1,1} + 9d_{s_{23}}^{1,1} + 7d_{s_{24}}^{1,1} + 11d_{s_{25}}^{1,1} + 15d_{s_{26}}^{1,1} +$$
$$10d_{s_{27}}^{1,1} + 3d_{s_{9}}^{2,2} + 2d_{s_{24}}^{2,2} + 3d_{s_{7}}^{1,1} - 2d_{s_{8}}^{1,1} + 2d_{s_{9}}^{1,1} + 8d_{s_{27}}^{1,2} + 8d_{s_{28}}^{1,2} + 8d_{s_{29}}^{1,2} + 13d_{s_{30}}^{1,2} + 13d_{s_{31}}^{1,2} + 7d_{s_{34}}^{1,2} +$$
$$d_{s_{18}}^{2,1} + d_{s_{19}}^{2,1} + d_{s_{20}}^{2,1} + d_{s_{21}}^{2,1} + d_{s_{22}}^{2,1} + d_{s_{23}}^{2,1} + 2d_{s_{24}}^{2,1} - 5/6 + 7d_{s_{35}}^{1,2} + 3d_{s_{24}}^{1,2} + 3d_{s_{25}}^{1,2} + 3d_{s_{26}}^{1,2} + d_{s_{10}}^{1,1} + 2d_{s_{29}}^{2,1},$$

$$d_{s_{21}}^{0,1} = -2d_{s_{30}}^{0,1} - d_{s_{28}}^{0,1} - d_{s_{25}}^{0,1} + 35d_{s_{34}}^{1,1} - 2d_{s_{12}}^{0,1} + 4d_{s_{14}}^{2,2} + 28d_{s_{29}}^{1,1} - 8d_{s_{11}}^{2,2} + 4d_{s_{20}}^{2,2} + 8d_{s_{21}}^{2,2} - 8d_{s_{1}}^{2,2} -$$
$$4d_{s_{5}}^{2,2} - 3d_{s_{19}}^{0,1} + 48d_{s_{36}}^{1,1} + 8d_{s_{26}}^{2,2} - d_{s_{17}}^{0,1} + 4d_{s_{7}}^{2,2} - 2d_{s_{20}}^{0,1} + 36d_{s_{33}}^{1,1} - 4d_{s_{12}}^{2,2} + 40d_{s_{35}}^{1,1} + 8d_{s_{32}}^{2,2} - d_{s_{13}}^{0,1} -$$
$$39d_{s_{4}}^{0,0} - 29d_{s_{7}}^{0,0} - 19d_{s_{9}}^{0,0} + 23d_{s_{28}}^{1,1} - 2d_{s_{24}}^{0,1} - 9d_{s_{10}}^{0,0} + 16d_{s_{36}}^{2,2} + 4d_{s_{11}}^{1,2} + 4d_{s_{12}}^{1,2} + 4d_{s_{13}}^{1,2} + 4d_{s_{14}}^{1,2} - 26d_{s_{18}}^{0,0} -$$
$$48d_{s_{1}}^{0,0} - 45d_{s_{2}}^{0,0} - 42d_{s_{3}}^{0,0} + 4d_{s_{22}}^{2,2} + 4d_{s_{27}}^{2,2} + 2d_{s_{31}}^{2,1} + 4d_{s_{25}}^{2,2} + 3d_{s_{32}}^{2,1} + 3d_{s_{33}}^{2,1} + 3d_{s_{34}}^{2,1} + 3d_{s_{35}}^{2,1} + 4d_{s_{36}}^{2,1} +$$
$$13d_{s_{16}}^{1,2} + 13d_{s_{17}}^{1,2} + 13d_{s_{18}}^{1,2} - 2d_{s_{27}}^{0,1} + 22d_{s_{19}}^{1,2} + 22d_{s_{20}}^{1,2} - 3d_{s_{22}}^{0,1} + 22d_{s_{21}}^{1,2} + 31d_{s_{22}}^{1,2} + 31d_{s_{23}}^{1,2} + d_{s_{12}}^{2,1} +$$
$$d_{s_{14}}^{2,1} + d_{s_{15}}^{2,1} + d_{s_{16}}^{2,1} + d_{s_{17}}^{2,1} - 35d_{s_{5}}^{0,0} - 32d_{s_{6}}^{0,0} - 22d_{s_{8}}^{0,0} - 48d_{s_{11}}^{0,0} + 31d_{s_{32}}^{1,1} - 45d_{s_{12}}^{0,0} - 42d_{s_{13}}^{0,0} - 13d_{s_{29}}^{0,0} -$$
$$16d_{s_{21}}^{0,0} - 6d_{s_{23}}^{0,0} + 4d_{s_{32}}^{1,2} + 4d_{s_{33}}^{1,2} + 8d_{s_{28}}^{2,2} + 12d_{s_{29}}^{2,2} + 8d_{s_{30}}^{2,2} + 12d_{s_{31}}^{2,2} - 39d_{s_{14}}^{0,0} - 35d_{s_{15}}^{0,0} - 32d_{s_{16}}^{0,0} -$$
$$29d_{s_{17}}^{0,0} - 22d_{s_{19}}^{0,0} - 19d_{s_{20}}^{0,0} - 9d_{s_{22}}^{0,0} + 13d_{s_{1}}^{1,2} + 13d_{s_{2}}^{1,2} + 13d_{s_{3}}^{1,2} + 13d_{s_{4}}^{1,2} + 22d_{s_{5}}^{1,2} + 22d_{s_{6}}^{1,2} - d_{s_{31}}^{0,1} -$$
$$11d_{s_{1}}^{1,1} - 6d_{s_{2}}^{1,1} - d_{s_{3}}^{1,1} + 4d_{s_{4}}^{1,1} - 7d_{s_{5}}^{1,1} - 2d_{s_{6}}^{1,1} - 4d_{s_{2}}^{2,2} - 32d_{s_{24}}^{0,0} - 29d_{s_{25}}^{0,0} - 26d_{s_{26}}^{0,0} - 2d_{s_{16}}^{0,1} - 19d_{s_{27}}^{0,0} -$$
$$16d_{s_{28}}^{0,0} - 6d_{s_{30}}^{0,0} + 12d_{s_{33}}^{2,2} - d_{s_{7}}^{0,1} - 2d_{s_{9}}^{0,1} - 3d_{s_{10}}^{0,1} + 12d_{s_{34}}^{2,2} + 16d_{s_{35}}^{2,2} - 4d_{s_{15}}^{2,2} + 4d_{s_{17}}^{2,2} + 8d_{s_{18}}^{2,2} - d_{s_{34}}^{0,1} -$$
$$16d_{s_{32}}^{0,0} - 13d_{s_{33}}^{0,0} - 3d_{s_{34}}^{0,0} + d_{s_{11}}^{2,1} + 4d_{s_{4}}^{2,2} + 4d_{s_{10}}^{2,2} + 22d_{s_{7}}^{1,2} + 31d_{s_{8}}^{1,2} + 31d_{s_{9}}^{1,2} + 40d_{s_{10}}^{1,2} + 13d_{s_{15}}^{1,2} - 3d_{s_{11}}^{1,1} +$$
$$2d_{s_{12}}^{1,1} + 7d_{s_{13}}^{1,1} + 12d_{s_{14}}^{1,1} + d_{s_{15}}^{1,1} + 6d_{s_{16}}^{1,1} + 11d_{s_{17}}^{1,1} + 27d_{s_{31}}^{1,1} + 16d_{s_{18}}^{1,1} - 3d_{s_{31}}^{0,0} + 2d_{s_{25}}^{2,1} + 2d_{s_{26}}^{2,1} + 2d_{s_{27}}^{2,1} +$$
$$2d_{s_{28}}^{2,1} + 2d_{s_{30}}^{2,1} + d_{s_{13}}^{2,1} + 8d_{s_{23}}^{2,2} + 4d_{s_{36}}^{1,2} + 5d_{s_{19}}^{1,1} + 10d_{s_{20}}^{1,1} + 22d_{s_{30}}^{1,1} + 15d_{s_{21}}^{1,1} - 3d_{s_{15}}^{0,1} + 9d_{s_{22}}^{1,1} + 14d_{s_{23}}^{1,1} +$$
$$14d_{s_{24}}^{1,1} + 19d_{s_{25}}^{1,1} + 24d_{s_{26}}^{1,1} + 18d_{s_{27}}^{1,1} + 4d_{s_{9}}^{2,2} + 3d_{s_{7}}^{1,1} - 3d_{s_{8}}^{1,1} + 2d_{s_{9}}^{1,1} + 13d_{s_{27}}^{1,2} + 13d_{s_{28}}^{1,2} + 13d_{s_{29}}^{1,2} +$$
$$22d_{s_{30}}^{1,2} + 22d_{s_{31}}^{1,2} + 13d_{s_{34}}^{1,2} + d_{s_{18}}^{2,1} + d_{s_{19}}^{2,1} + d_{s_{20}}^{2,1} + d_{s_{21}}^{2,1} + d_{s_{22}}^{2,1} + d_{s_{23}}^{2,1} + 2d_{s_{24}}^{2,1} - 5/6 + 13d_{s_{35}}^{1,2} + 4d_{s_{24}}^{1,2} +$$
$$4d_{s_{25}}^{1,2} + 4d_{s_{26}}^{1,2} - 3d_{s_{1}}^{0,1} - 2d_{s_{2}}^{0,1} - d_{s_{3}}^{0,1} - 3d_{s_{5}}^{0,1} - 2d_{s_{6}}^{0,1} - 3d_{s_{8}}^{0,1} + d_{s_{10}}^{1,1} + 2d_{s_{29}}^{2,1} - 2d_{s_{23}}^{0,1} - 3d_{s_{11}}^{0,1} - d_{s_{32}}^{0,1},$$

$$d_{s_{6}}^{2,2} = -11d_{s_{34}}^{1,1} - d_{s_{14}}^{2,2} - 10d_{s_{29}}^{1,1} + 2d_{s_{11}}^{2,2} - 2d_{s_{19}}^{2,2} - 3d_{s_{20}}^{2,2} - 4d_{s_{21}}^{2,2} + 2d_{s_{1}}^{2,2} - 15d_{s_{36}}^{1,1} - 3d_{s_{26}}^{2,2} -$$
$$2d_{s_{7}}^{2,2} - 12d_{s_{33}}^{1,1} + d_{s_{12}}^{2,2} - 13d_{s_{35}}^{1,1} - 4d_{s_{32}}^{2,2} + 12d_{s_{4}}^{0,0} + 9d_{s_{7}}^{0,0} + 6d_{s_{9}}^{0,0} - 2d_{s_{8}}^{2,2} - 8d_{s_{28}}^{1,1} + 3d_{s_{10}}^{0,0} - 7d_{s_{36}}^{2,2} -$$
$$4d_{s_{11}}^{1,2} - 4d_{s_{12}}^{1,2} - 4d_{s_{13}}^{1,2} - 4d_{s_{14}}^{1,2} + 8d_{s_{18}}^{0,0} + 15d_{s_{1}}^{0,0} + 14d_{s_{2}}^{0,0} + 13d_{s_{3}}^{0,0} - 4d_{s_{22}}^{2,2} - 3d_{s_{27}}^{2,2} - 2d_{s_{31}}^{2,1} - 2d_{s_{25}}^{2,2} -$$
$$3d_{s_{32}}^{2,1} - 3d_{s_{33}}^{2,1} - 3d_{s_{34}}^{2,1} - 3d_{s_{35}}^{2,1} - 4d_{s_{36}}^{2,1} - 6d_{s_{16}}^{1,2} - 6d_{s_{17}}^{1,2} - 6d_{s_{18}}^{1,2} - 8d_{s_{19}}^{1,2} - 8d_{s_{20}}^{1,2} - 8d_{s_{21}}^{1,2} - 10d_{s_{22}}^{1,2} -$$
$$10d_{s_{23}}^{1,2} - d_{s_{12}}^{2,1} - d_{s_{14}}^{2,1} - d_{s_{15}}^{2,1} - d_{s_{16}}^{2,1} - d_{s_{17}}^{2,1} + 11d_{s_{5}}^{0,0} + 10d_{s_{6}}^{0,0} + 7d_{s_{8}}^{0,0} + 15d_{s_{11}}^{0,0} - 10d_{s_{32}}^{1,1} + 14d_{s_{12}}^{0,0} +$$
$$13d_{s_{13}}^{0,0} + 4d_{s_{29}}^{0,0} + 5d_{s_{21}}^{0,0} + 2d_{s_{23}}^{0,0} - 2d_{s_{32}}^{1,2} - 2d_{s_{33}}^{1,2} - 4d_{s_{28}}^{2,2} - 5d_{s_{29}}^{2,2} - 5d_{s_{30}}^{2,2} - 6d_{s_{31}}^{2,2} + 12d_{s_{14}}^{0,0} + 11d_{s_{15}}^{0,0} +$$
$$10d_{s_{16}}^{0,0} + 9d_{s_{17}}^{0,0} + 7d_{s_{19}}^{0,0} + 6d_{s_{20}}^{0,0} + 3d_{s_{22}}^{0,0} - 7d_{s_{1}}^{1,2} - 7d_{s_{2}}^{1,2} - 7d_{s_{3}}^{1,2} - 7d_{s_{4}}^{1,2} - 9d_{s_{5}}^{1,2} - 9d_{s_{6}}^{1,2} + 2d_{s_{1}}^{1,1} -$$
$$2d_{s_{3}}^{1,1} - 4d_{s_{4}}^{1,1} + d_{s_{5}}^{1,1} - d_{s_{6}}^{1,1} + d_{s_{2}}^{2,2} + 10d_{s_{24}}^{0,0} + 9d_{s_{25}}^{0,0} + 8d_{s_{26}}^{0,0} + 6d_{s_{27}}^{0,0} + 5d_{s_{28}}^{0,0} + 2d_{s_{30}}^{0,0} - 5d_{s_{33}}^{2,2} - 6d_{s_{34}}^{2,2} -$$
$$7d_{s_{35}}^{2,2} - d_{s_{16}}^{2,2} - 2d_{s_{17}}^{2,2} - 3d_{s_{18}}^{2,2} + 5d_{s_{32}}^{0,0} + 4d_{s_{33}}^{0,0} + d_{s_{34}}^{0,0} - d_{s_{11}}^{2,1} - d_{s_{4}}^{2,2} - 4d_{s_{10}}^{2,2} - 9d_{s_{7}}^{1,2} - 11d_{s_{8}}^{1,2} - 11d_{s_{9}}^{1,2} -$$
$$13d_{s_{10}}^{1,2} - 6d_{s_{15}}^{1,2} - 2d_{s_{12}}^{1,1} - 4d_{s_{13}}^{1,1} - 6d_{s_{14}}^{1,1} - d_{s_{15}}^{1,1} - 3d_{s_{16}}^{1,1} - 5d_{s_{17}}^{1,1} - 9d_{s_{31}}^{1,1} - 7d_{s_{18}}^{1,1} + d_{s_{31}}^{0,0} - 2d_{s_{25}}^{2,1} - 2d_{s_{26}}^{2,1} -$$
$$2d_{s_{27}}^{2,1} - 2d_{s_{28}}^{2,1} - 2d_{s_{30}}^{2,1} - d_{s_{13}}^{2,1} - 5d_{s_{23}}^{2,2} - d_{s_{36}}^{1,2} - 2d_{s_{19}}^{1,1} - 4d_{s_{20}}^{1,1} - 7d_{s_{30}}^{1,1} - 6d_{s_{21}}^{1,1} - 3d_{s_{22}}^{1,1} - 5d_{s_{23}}^{1,1} - 5d_{s_{24}}^{1,1} -$$
$$7d_{s_{25}}^{1,1} - 9d_{s_{26}}^{1,1} - 6d_{s_{27}}^{1,1} - 3d_{s_{9}}^{2,2} - d_{s_{24}}^{2,2} - 3d_{s_{7}}^{1,1} - 2d_{s_{9}}^{1,1} - 5d_{s_{27}}^{1,2} - 5d_{s_{28}}^{1,2} - 5d_{s_{29}}^{1,2} - 7d_{s_{30}}^{1,2} - 7d_{s_{31}}^{1,2} - 4d_{s_{34}}^{1,2} -$$
$$d_{s_{18}}^{2,1} - d_{s_{19}}^{2,1} - d_{s_{20}}^{2,1} - d_{s_{21}}^{2,1} - d_{s_{22}}^{2,1} - d_{s_{23}}^{2,1} - 2d_{s_{24}}^{2,1} + 5/6 - 4d_{s_{35}}^{1,2} - 3d_{s_{24}}^{1,2} - 3d_{s_{25}}^{1,2} - 3d_{s_{26}}^{1,2} - d_{s_{10}}^{1,1} - 2d_{s_{29}}^{2,1},$$

$$d_{s_{32}}^{0,2} = d_{s_{36}}^{1,0} + d_{s_{33}}^{1,0} + d_{s_{11}}^{1,0} + 8d_{s_{34}}^{1,1} + d_{s_{14}}^{2,2} + 6d_{s_{29}}^{1,1} - 2d_{s_{11}}^{2,2} + d_{s_{20}}^{2,2} + 2d_{s_{21}}^{2,2} + d_{s_{14}}^{1,0} + d_{s_{9}}^{0,2} + d_{s_{10}}^{0,2} -$$
$$3d_{s_{10}}^{1,0} - 2d_{s_{1}}^{2,2} - d_{s_{5}}^{2,2} + d_{s_{4}}^{0,2} + 12d_{s_{36}}^{1,1} + 2d_{s_{26}}^{2,2} + d_{s_{7}}^{2,2} + 9d_{s_{33}}^{1,1} - d_{s_{12}}^{2,2} + 9d_{s_{35}}^{1,1} + d_{s_{25}}^{1,0} - d_{s_{19}}^{1,0} + 2d_{s_{32}}^{2,2} - d_{s_{20}}^{1,0} -$$
$$12d_{s_{4}}^{0,0} - 9d_{s_{7}}^{0,0} - 6d_{s_{9}}^{0,0} + 5d_{s_{28}}^{1,1} - 3d_{s_{10}}^{0,0} + 4d_{s_{36}}^{2,2} + d_{s_{11}}^{1,2} + d_{s_{12}}^{1,2} + d_{s_{13}}^{1,2} + d_{s_{14}}^{1,2} - 8d_{s_{18}}^{0,0} - 15d_{s_{1}}^{0,0} - 14d_{s_{2}}^{0,0} -$$
$$13d_{s_{3}}^{0,0} + d_{s_{22}}^{2,2} + d_{s_{27}}^{2,2} + d_{s_{21}}^{0,2} + d_{s_{25}}^{2,2} + 3d_{s_{16}}^{1,2} + 3d_{s_{17}}^{1,2} + 3d_{s_{18}}^{1,2} + 5d_{s_{19}}^{1,2} + 5d_{s_{20}}^{1,2} + 5d_{s_{21}}^{1,2} + 7d_{s_{22}}^{1,2} + 7d_{s_{23}}^{1,2} -$$
$$11d_{s_{5}}^{0,0} - 2d_{s_{9}}^{1,0} - 10d_{s_{6}}^{0,0} - 2d_{s_{8}}^{1,0} - 7d_{s_{8}}^{0,0} - 15d_{s_{11}}^{0,0} + 8d_{s_{32}}^{1,1} - 14d_{s_{12}}^{0,0} - 13d_{s_{13}}^{0,0} - 4d_{s_{29}}^{0,0} - 5d_{s_{21}}^{0,0} - 2d_{s_{23}}^{0,0} -$$
$$d_{s_{5}}^{1,0} - d_{s_{32}}^{1,2} - d_{s_{33}}^{1,2} + 2d_{s_{28}}^{2,2} + 3d_{s_{29}}^{2,2} + 2d_{s_{30}}^{2,2} + 3d_{s_{31}}^{2,2} - 2d_{s_{1}}^{0,2} - d_{s_{2}}^{0,2} - 2d_{s_{24}}^{0,2} - d_{s_{25}}^{0,2} + d_{s_{26}}^{1,0} - 12d_{s_{14}}^{0,0} + d_{s_{13}}^{1,0} -$$
$$11d_{s_{15}}^{0,0} - 10d_{s_{16}}^{0,0} - 9d_{s_{17}}^{0,0} - 7d_{s_{19}}^{0,0} - 6d_{s_{20}}^{0,0} + d_{s_{32}}^{1,0} - 3d_{s_{22}}^{0,0} + d_{s_{12}}^{1,0} + 4d_{s_{1}}^{1,2} + 4d_{s_{2}}^{1,2} + 4d_{s_{3}}^{1,2} + 4d_{s_{4}}^{1,2} + 6d_{s_{5}}^{1,2} +$$
$$6d_{s_{6}}^{1,2} - 3d_{s_{1}}^{1,1} - 2d_{s_{2}}^{1,1} - d_{s_{3}}^{1,1} - 3d_{s_{5}}^{1,1} - 2d_{s_{6}}^{1,1} - d_{s_{2}}^{2,2} - 10d_{s_{24}}^{0,0} - 9d_{s_{25}}^{0,0} - 8d_{s_{26}}^{0,0} - 6d_{s_{27}}^{0,0} - 5d_{s_{28}}^{0,0} - 2d_{s_{30}}^{0,0} +$$
$$d_{s_{31}}^{0,2} + d_{s_{35}}^{0,2} + 3d_{s_{33}}^{2,2} - d_{s_{7}}^{1,0} + 3d_{s_{34}}^{2,2} + 4d_{s_{35}}^{2,2} - d_{s_{15}}^{2,2} + d_{s_{17}}^{2,2} + 2d_{s_{18}}^{2,2} - 5d_{s_{32}}^{0,0} - 4d_{s_{33}}^{0,0} - d_{s_{34}}^{0,0} + d_{s_{4}}^{2,2} + d_{s_{10}}^{2,2} -$$

$$d_{s_{21}}^{1,0} - 2d_{s_{22}}^{1,0} - 2d_{s_{23}}^{1,0} - 2d_{s_{12}}^{0,2} - d_{s_{13}}^{0,2} - 2d_{s_{15}}^{0,2} - d_{s_{16}}^{0,2} - d_{s_{19}}^{0,2} - d_{s_{31}}^{1,0} - d_{s_{30}}^{1,0} + d_{s_{23}}^{0,2} + 6d_{s_7}^{1,2} + 8d_{s_8}^{1,2} + 8d_{s_9}^{1,2} + 10d_{s_{10}}^{1,2} + 3d_{s_{15}}^{1,2} + d_{s_{12}}^{1,1} + 2d_{s_{13}}^{1,1} + 3d_{s_{14}}^{1,1} + d_{s_{16}}^{1,1} + 2d_{s_{17}}^{1,1} + 5d_{s_{31}}^{1,1} + 3d_{s_{18}}^{1,1} - d_{s_{31}}^{0,0} + 2d_{s_{23}}^{2,2} - 2d_{s_{36}}^{1,2} + d_{s_{20}}^{1,1} + 4d_{s_{30}}^{1,1} + 2d_{s_{21}}^{1,1} + d_{s_{23}}^{1,1} + 4d_{s_{24}}^{1,1} - d_{s_6}^{1,0} + 5d_{s_{25}}^{1,1} + 6d_{s_{26}}^{1,1} + 4d_{s_{27}}^{1,1} + d_{s_9}^{2,2} + d_{s_7}^{0,2} - d_{s_5}^{0,2} - 3d_{s_{11}}^{0,2} - d_{s_7}^{1,1} - 3d_{s_8}^{1,1} - 2d_{s_9}^{1,1} + d_{s_{18}}^{0,2} - d_{s_{27}}^{0,2} + d_{s_{24}}^{1,0} + 2d_{s_{27}}^{1,2} + 2d_{s_{28}}^{1,2} + 2d_{s_{29}}^{1,2} + 4d_{s_{30}}^{1,2} + 4d_{s_{31}}^{1,2} + d_{s_{34}}^{1,2} + 1/6 + d_{s_{35}}^{1,2} - 3d_{s_{10}}^{1,1} + d_{s_{29}}^{0,2},$$

$$d_{s_1}^{2,1} = -18d_{s_{34}}^{1,1} - 2d_{s_{14}}^{2,2} - 14d_{s_{29}}^{1,1} + 4d_{s_{11}}^{2,2} - 2d_{s_{20}}^{2,2} - 4d_{s_{21}}^{2,2} + 4d_{s_1}^{2,2} + 2d_{s_5}^{2,2} - 24d_{s_{36}}^{1,1} - 4d_{s_{26}}^{2,2} - 2d_{s_7}^{2,2} - 18d_{s_{33}}^{1,1} + 2d_{s_{12}}^{2,2} - 20d_{s_{35}}^{1,1} - 4d_{s_{32}}^{2,2} + 18d_{s_4}^{0,0} + 14d_{s_7}^{0,0} + 10d_{s_9}^{0,0} - d_{s_2}^{2,1} - d_{s_3}^{2,1} - 12d_{s_{28}}^{1,1} + 6d_{s_{10}}^{0,0} - 8d_{s_{36}}^{2,2} - d_{s_4}^{2,1} - 2d_{s_{11}}^{1,2} - 3d_{s_{12}}^{1,2} - 4d_{s_{13}}^{1,2} - 5d_{s_{14}}^{1,2} + 12d_{s_{18}}^{0,0} + 24d_{s_1}^{0,0} + 22d_{s_2}^{0,0} + 20d_{s_3}^{0,0} - 2d_{s_{22}}^{2,2} - 2d_{s_{27}}^{2,2} - 2d_{s_{25}}^{2,2} - 2d_{s_{32}}^{2,1} - 2d_{s_{33}}^{2,1} - d_{s_{34}}^{2,1} - d_{s_{35}}^{2,1} - 2d_{s_{36}}^{2,1} - 7d_{s_{16}}^{1,2} - 8d_{s_{17}}^{1,2} - 9d_{s_{18}}^{1,2} - 10d_{s_{19}}^{1,2} - 11d_{s_{20}}^{1,2} - 12d_{s_{21}}^{1,2} - 14d_{s_{22}}^{1,2} - 15d_{s_{23}}^{1,2} - 2d_{s_{12}}^{2,1} - 2d_{s_{14}}^{2,1} - d_{s_{15}}^{2,1} - d_{s_{16}}^{2,1} - d_{s_{17}}^{2,1} + 18d_{s_5}^{0,0} + 16d_{s_6}^{0,0} + 12d_{s_8}^{0,0} + 24d_{s_{11}}^{0,0} - 16d_{s_{32}}^{1,1} + 22d_{s_{12}}^{0,0} + 20d_{s_{13}}^{0,0} + 6d_{s_{29}}^{0,0} + 8d_{s_{21}}^{0,0} + 4d_{s_{23}}^{0,0} - 2d_{s_{32}}^{1,2} - 3d_{s_{33}}^{1,2} - 4d_{s_{28}}^{2,2} - 6d_{s_{29}}^{2,2} - 4d_{s_{30}}^{2,2} - 6d_{s_{31}}^{2,2} + d_{s_8}^{2,1} + d_{s_9}^{2,1} + 18d_{s_{14}}^{0,0} + 18d_{s_{15}}^{0,0} + 16d_{s_{16}}^{0,0} + 14d_{s_{17}}^{0,0} + 12d_{s_{19}}^{0,0} + 10d_{s_{20}}^{0,0} + 6d_{s_{22}}^{0,0} - 7d_{s_1}^{1,2} - 8d_{s_2}^{1,2} - 9d_{s_3}^{1,2} - 10d_{s_4}^{1,2} - 11d_{s_5}^{1,2} - 12d_{s_6}^{1,2} + 4d_{s_1}^{1,1} + 2d_{s_2}^{1,1} - 2d_{s_4}^{1,1} + 2d_{s_5}^{1,1} + 2d_{s_2}^{2,2} + 16d_{s_{24}}^{0,0} + 14d_{s_{25}}^{0,0} + 12d_{s_{26}}^{0,0} + 10d_{s_{27}}^{0,0} + 8d_{s_{28}}^{0,0} + 4d_{s_{30}}^{0,0} - 6d_{s_{33}}^{2,2} - 6d_{s_{34}}^{2,2} - 8d_{s_{35}}^{2,2} + 2d_{s_{15}}^{2,2} - 2d_{s_{17}}^{2,2} - 4d_{s_{18}}^{2,2} + 8d_{s_{32}}^{0,0} + 6d_{s_{33}}^{0,0} + 2d_{s_{34}}^{0,0} - 2d_{s_{11}}^{2,1} - 2d_{s_4}^{2,2} - 2d_{s_{10}}^{2,2} - 13d_{s_7}^{1,2} - 15d_{s_8}^{1,2} - 16d_{s_9}^{1,2} - 19d_{s_{10}}^{1,2} - 6d_{s_{15}}^{1,2} - 2d_{s_{12}}^{1,1} - 4d_{s_{13}}^{1,1} - 6d_{s_{14}}^{1,1} - 2d_{s_{15}}^{1,1} - 4d_{s_{16}}^{1,1} - 6d_{s_{17}}^{1,1} - 14d_{s_{31}}^{1,1} - 8d_{s_{18}}^{1,1} + 2d_{s_{31}}^{0,0} - 2d_{s_{25}}^{2,1} - 2d_{s_{26}}^{2,1} - d_{s_{27}}^{2,1} - d_{s_{28}}^{2,1} - 2d_{s_{13}}^{2,1} + 2d_{s_{10}}^{2,1} - 4d_{s_{23}}^{2,2} - 2d_{s_{36}}^{1,2} - 4d_{s_{19}}^{1,1} - 6d_{s_{20}}^{1,1} - 12d_{s_{30}}^{1,1} - 8d_{s_{21}}^{1,1} - 6d_{s_{22}}^{1,1} - 8d_{s_{23}}^{1,1} - 8d_{s_{24}}^{1,1} - 10d_{s_{25}}^{1,1} - 12d_{s_{26}}^{1,1} - 10d_{s_{27}}^{1,1} - 2d_{s_9}^{2,2} - 2d_{s_7}^{1,1} - 2d_{s_9}^{1,1} - 6d_{s_{27}}^{1,2} - 7d_{s_{28}}^{1,2} - 8d_{s_{29}}^{1,2} - 10d_{s_{30}}^{1,2} - 11d_{s_{31}}^{1,2} - 6d_{s_{34}}^{1,2} - d_{s_{18}}^{2,1} + d_{s_{22}}^{2,1} + d_{s_{23}}^{2,1} - 2d_{s_{24}}^{2,1} - 7d_{s_{35}}^{1,2} - 2d_{s_{24}}^{1,2} - 3d_{s_{25}}^{1,2} - 4d_{s_{26}}^{1,2} - 2d_{s_{10}}^{1,1} - d_{s_{29}}^{2,1} + 2/3,$$

$$d_{s_{16}}^{2,0} = 1/2 + d_{s_{36}}^{1,0} - 3d_{s_{32}}^{2,0} + d_{s_{33}}^{1,0} + d_{s_{11}}^{1,0} - d_{s_{30}}^{2,0} + 2d_{s_{34}}^{1,1} - 2d_{s_{33}}^{2,0} - d_{s_{29}}^{1,1} + d_{s_{14}}^{1,0} - 3d_{s_{10}}^{1,0} + 3d_{s_{36}}^{1,1} + d_{s_{33}}^{1,1} + d_{s_{35}}^{1,1} - 2d_{s_{12}}^{2,0} - d_{s_{13}}^{2,0} + d_{s_{25}}^{1,0} - d_{s_{19}}^{1,0} - d_{s_{20}}^{1,0} - 3d_{s_4}^{0,0} - 3d_{s_7}^{0,0} - 3d_{s_9}^{0,0} - 3d_{s_{10}}^{0,0} + 3d_{s_{11}}^{1,2} + 2d_{s_{12}}^{1,2} + d_{s_{13}}^{1,2} - 2d_{s_{18}}^{0,0} - 6d_{s_1}^{0,0} - 5d_{s_2}^{0,0} - 4d_{s_3}^{0,0} - 2d_{s_{31}}^{2,1} - 3d_{s_{32}}^{2,1} - 3d_{s_{33}}^{2,1} - 3d_{s_{34}}^{2,1} - 3d_{s_{35}}^{2,1} - 4d_{s_{36}}^{2,1} + 2d_{s_{16}}^{1,2} + d_{s_{17}}^{1,2} + 3d_{s_{19}}^{1,2} + 2d_{s_{20}}^{1,2} + d_{s_{21}}^{1,2} + 3d_{s_{22}}^{1,2} + 2d_{s_{23}}^{1,2} - d_{s_{12}}^{2,1} - d_{s_{14}}^{2,1} - d_{s_{15}}^{2,1} - d_{s_{16}}^{2,1} - d_{s_{17}}^{2,1} - 5d_{s_5}^{0,0} - 2d_{s_9}^{1,0} - 4d_{s_6}^{0,0} - 2d_{s_8}^{1,0} - 4d_{s_8}^{0,0} - 6d_{s_{11}}^{0,0} + 2d_{s_{32}}^{1,1} - 5d_{s_{12}}^{0,0} - 4d_{s_{13}}^{0,0} - d_{s_{29}}^{0,0} - 2d_{s_{21}}^{0,0} - 2d_{s_{23}}^{0,0} - d_{s_5}^{1,0} + d_{s_{32}}^{1,2} + d_{s_{26}}^{1,0} - 3d_{s_{14}}^{0,0} + d_{s_{13}}^{1,0} - 5d_{s_{15}}^{0,0} - 4d_{s_{16}}^{0,0} - 3d_{s_{17}}^{0,0} - 4d_{s_{19}}^{0,0} - 3d_{s_{20}}^{0,0} + d_{s_{32}}^{1,0} - 3d_{s_{22}}^{0,0} - d_{s_1}^{2,0} + d_{s_3}^{2,0} + 2d_{s_4}^{2,0} + d_{s_6}^{2,0} + 2d_{s_7}^{2,0} + d_{s_{12}}^{1,0} + 3d_{s_1}^{1,2} + 2d_{s_2}^{1,2} + d_{s_3}^{1,2} + 3d_{s_5}^{1,2} + 2d_{s_6}^{1,2} - 3d_{s_{24}}^{2,0} - 2d_{s_{25}}^{2,0} - d_{s_{26}}^{2,0} - 2d_{s_1}^{1,1} - 3d_{s_2}^{1,1} - 4d_{s_3}^{1,1} - 5d_{s_4}^{1,1} - 2d_{s_5}^{1,1} - 3d_{s_6}^{1,1} - 4d_{s_{24}}^{0,0} - 3d_{s_{25}}^{0,0} - 2d_{s_{26}}^{0,0} - 3d_{s_{27}}^{0,0} - 2d_{s_{28}}^{0,0} - 2d_{s_{30}}^{0,0} - d_{s_7}^{1,0} - 2d_{s_{32}}^{0,0} - d_{s_{33}}^{0,0} - d_{s_{34}}^{0,0} - d_{s_{11}}^{2,1} - d_{s_{21}}^{1,0} - 2d_{s_{22}}^{1,0} - 2d_{s_{23}}^{1,0} + d_{s_{18}}^{2,0} - d_{s_{19}}^{2,0} + d_{s_{21}}^{2,0} + d_{s_{23}}^{2,0} - 2d_{s_{27}}^{2,0} - d_{s_{31}}^{1,0} - d_{s_{30}}^{1,0} + d_{s_7}^{1,2} + 3d_{s_8}^{1,2} + 2d_{s_9}^{1,2} + 3d_{s_{10}}^{1,2} + 3d_{s_{15}}^{1,2} - 2d_{s_{34}}^{2,0} - d_{s_{12}}^{1,1} - 2d_{s_{13}}^{1,1} - 3d_{s_{14}}^{1,1} - d_{s_{16}}^{1,1} - 2d_{s_{17}}^{1,1} - 3d_{s_{18}}^{1,1} - d_{s_{31}}^{0,0} - 2d_{s_{25}}^{2,1} - 2d_{s_{26}}^{2,1} - 3d_{s_{36}}^{2,0} - 2d_{s_{27}}^{2,1} - 2d_{s_{28}}^{2,1} - 2d_{s_{30}}^{2,1} - d_{s_{13}}^{2,1} - d_{s_{20}}^{1,1} + d_{s_{30}}^{1,1} - 2d_{s_{21}}^{1,1} - d_{s_{23}}^{1,1} + d_{s_{24}}^{1,1} - d_{s_6}^{1,0} - d_{s_{26}}^{1,1} + d_{s_{27}}^{1,1} - 3d_{s_{11}}^{2,0} - 4d_{s_7}^{1,1} - 2d_{s_8}^{1,1} - 3d_{s_9}^{1,1} - d_{s_{35}}^{2,0} - d_{s_{28}}^{2,0} + d_{s_{24}}^{1,0} + 2d_{s_{27}}^{1,2} + d_{s_{28}}^{1,2} + 2d_{s_{30}}^{1,2} + d_{s_{31}}^{1,2} + d_{s_{34}}^{1,2} - d_{s_{18}}^{2,1} - d_{s_{19}}^{2,1} - d_{s_{20}}^{2,1} - d_{s_{21}}^{2,1} - d_{s_{22}}^{2,1} - d_{s_{23}}^{2,1} - 2d_{s_{24}}^{2,1} + d_{s_8}^{2,0} + 2d_{s_9}^{2,0} + 2d_{s_{10}}^{2,0} - 2d_{s_{15}}^{2,0} + 2d_{s_{24}}^{1,2} + d_{s_{25}}^{1,2} - 2d_{s_{10}}^{1,1} - 2d_{s_{29}}^{2,1},$$

$$d_{s_{29}}^{0,1} = 4d_{s_{30}}^{0,1} + 2d_{s_{28}}^{0,1} - d_{s_{36}}^{1,0} - d_{s_{33}}^{1,0} + 3d_{s_{25}}^{0,1} - d_{s_{11}}^{1,0} - 44d_{s_{34}}^{1,1} + 5d_{s_{12}}^{0,1} - 5d_{s_{14}}^{2,2} - 35d_{s_{29}}^{1,1} + 10d_{s_{11}}^{2,2} - 5d_{s_{20}}^{2,2} - 10d_{s_{21}}^{2,2} - d_{s_{14}}^{1,0} + 3d_{s_{10}}^{1,0} + 10d_{s_1}^{2,2} + 5d_{s_5}^{2,2} + 6d_{s_{19}}^{0,1} - 60d_{s_{36}}^{1,1} - 10d_{s_{26}}^{2,2} + d_{s_{17}}^{0,1} - 5d_{s_7}^{2,2} + 3d_{s_{20}}^{0,1} - 45d_{s_{33}}^{1,1} + 5d_{s_{12}}^{2,2} - 50d_{s_{35}}^{1,1} - d_{s_{25}}^{1,0} + d_{s_{19}}^{1,0} - 10d_{s_{32}}^{2,2} + d_{s_{20}}^{1,0} + 2d_{s_{13}}^{0,1} + 48d_{s_4}^{0,0} + 36d_{s_7}^{0,0} + 24d_{s_9}^{0,0} - 29d_{s_{28}}^{1,1} + 6d_{s_{24}}^{0,1} + 12d_{s_{10}}^{0,0} - 20d_{s_{36}}^{2,2} - 7d_{s_{11}}^{1,2} - 7d_{s_{12}}^{1,2} - 7d_{s_{13}}^{1,2} - 7d_{s_{14}}^{1,2} + 32d_{s_{18}}^{0,0} + 60d_{s_1}^{0,0} + 56d_{s_2}^{0,0} + 52d_{s_3}^{0,0} - 5d_{s_{22}}^{2,2} - 5d_{s_{27}}^{2,2} - 2d_{s_{31}}^{2,1} - 5d_{s_{25}}^{2,2} - 3d_{s_{32}}^{2,1} - 3d_{s_{33}}^{2,1} - 3d_{s_{34}}^{2,1} - 3d_{s_{35}}^{2,1} - 4d_{s_{36}}^{2,1} - 18d_{s_{16}}^{1,2} - 18d_{s_{17}}^{1,2} - 18d_{s_{18}}^{1,2} + 5d_{s_{27}}^{0,1} - 29d_{s_{19}}^{1,2} - 29d_{s_{20}}^{1,2} + 5d_{s_{22}}^{0,1} - 29d_{s_{21}}^{1,2} - 40d_{s_{22}}^{1,2} - 40d_{s_{23}}^{1,2} - d_{s_{12}}^{2,1} - d_{s_{14}}^{2,1} - d_{s_{15}}^{2,1} - d_{s_{16}}^{2,1} - d_{s_{17}}^{2,1} + 44d_{s_5}^{0,0} + 2d_{s_9}^{1,0} + 40d_{s_6}^{0,0} + 2d_{s_8}^{1,0} + 28d_{s_8}^{0,0} + 60d_{s_{11}}^{0,0} - 39d_{s_{32}}^{1,1} + 56d_{s_{12}}^{0,0} + 52d_{s_{13}}^{0,0} + 16d_{s_{29}}^{0,0} + 20d_{s_{21}}^{0,0} + 8d_{s_{23}}^{0,0} + d_{s_5}^{1,0} - 9d_{s_{32}}^{1,2} - 9d_{s_{33}}^{1,2} - 10d_{s_{28}}^{2,2} - 15d_{s_{29}}^{2,2} - 10d_{s_{30}}^{2,2} - 15d_{s_{31}}^{2,2} - d_{s_{26}}^{1,0} + 48d_{s_{14}}^{0,0} - d_{s_{13}}^{1,0} + 44d_{s_{15}}^{0,0} + 40d_{s_{16}}^{0,0} + 36d_{s_{17}}^{0,0} + 28d_{s_{19}}^{0,0} + 24d_{s_{20}}^{0,0} - d_{s_{32}}^{1,0} + 12d_{s_{22}}^{0,0} - d_{s_{12}}^{1,0} - 17d_{s_1}^{1,2} - 17d_{s_2}^{1,2} - 17d_{s_3}^{1,2} - 17d_{s_4}^{1,2} - 28d_{s_5}^{1,2} - 28d_{s_6}^{1,2} + d_{s_{31}}^{0,1} + 13d_{s_1}^{1,1} + 7d_{s_2}^{1,1} + d_{s_3}^{1,1} - 5d_{s_4}^{1,1} + 8d_{s_5}^{1,1} + 2d_{s_6}^{1,1} + 5d_{s_2}^{2,2} + 40d_{s_{24}}^{0,0} + 36d_{s_{25}}^{0,0} + 32d_{s_{26}}^{0,0} + 4d_{s_{16}}^{0,1} + 24d_{s_{27}}^{0,0} + 20d_{s_{28}}^{0,0} + 8d_{s_{30}}^{0,0} - 15d_{s_{33}}^{2,2} - d_{s_7}^{0,1} + d_{s_7}^{1,0} + d_{s_9}^{0,1} + 3d_{s_{10}}^{0,1} - 15d_{s_{34}}^{2,2} - 20d_{s_{35}}^{2,2} + 5d_{s_{15}}^{2,2} - 5d_{s_{17}}^{2,2} - 10d_{s_{18}}^{2,2} + 3d_{s_{34}}^{0,1} + 20d_{s_{32}}^{0,0} + 16d_{s_{33}}^{0,0} + 4d_{s_{34}}^{0,0} - d_{s_{11}}^{2,1} - 5d_{s_4}^{2,2} - 5d_{s_{10}}^{2,2} + d_{s_{21}}^{1,0} + 2d_{s_{22}}^{1,0} + 2d_{s_{23}}^{1,0} + d_{s_{31}}^{1,0} + d_{s_{30}}^{1,0} - 28d_{s_7}^{1,2} - 39d_{s_8}^{1,2} - 39d_{s_9}^{1,2} - 50d_{s_{10}}^{1,2} - 18d_{s_{15}}^{1,2} + 3d_{s_{11}}^{1,1} - 3d_{s_{12}}^{1,1} - 9d_{s_{13}}^{1,1} - 15d_{s_{14}}^{1,1} - 2d_{s_{15}}^{1,1} - 8d_{s_{16}}^{1,1} - 14d_{s_{17}}^{1,1} - 34d_{s_{31}}^{1,1} - 20d_{s_{18}}^{1,1} + 4d_{s_{31}}^{0,0} - 2d_{s_{25}}^{2,1} - 2d_{s_{26}}^{2,1} - 2d_{s_{27}}^{2,1} - 2d_{s_{28}}^{2,1} - 2d_{s_{30}}^{2,1} - d_{s_{13}}^{2,1} - 10d_{s_{23}}^{2,2} - 10d_{s_{36}}^{1,2} - 7d_{s_{19}}^{1,1} - 13d_{s_{20}}^{1,1} - 28d_{s_{30}}^{1,1} - 19d_{s_{21}}^{1,1} + 7d_{s_{15}}^{0,1} - 12d_{s_{22}}^{1,1} - 18d_{s_{23}}^{1,1} - 18d_{s_{24}}^{1,1} + d_{s_6}^{1,0} - 24d_{s_{25}}^{1,1} - 30d_{s_{26}}^{1,1} -$$

$$2d_{s_{18}}^{0,1} - 23d_{s_{27}}^{1,1} - 5d_{s_9}^{2,2} - 4d_{s_7}^{1,1} + 3d_{s_8}^{1,1} - 3d_{s_9}^{1,1} - d_{s_{24}}^{1,0} - 19d_{s_{27}}^{1,2} - 19d_{s_{28}}^{1,2} - 19d_{s_{29}}^{1,2} - 30d_{s_{30}}^{1,2} - 30d_{s_{31}}^{1,2} - 20d_{s_{34}}^{1,2} - d_{s_{18}}^{2,1} - d_{s_{19}}^{2,1} - d_{s_{20}}^{2,1} - d_{s_{21}}^{2,1} - d_{s_{22}}^{2,1} - d_{s_{23}}^{2,1} - 2d_{s_{24}}^{2,1} - 20d_{s_{35}}^{1,2} - 8d_{s_{24}}^{1,2} - 8d_{s_{25}}^{1,2} - 8d_{s_{26}}^{1,2} + 6d_{s_1}^{0,1} + 3d_{s_2}^{0,1} + 5d_{s_5}^{0,1} - d_{s_{14}}^{0,1} + 2d_{s_6}^{0,1} + 4d_{s_8}^{0,1} - 2d_{s_{10}}^{1,1} - 2d_{s_{29}}^{2,1} + 2d_{s_{23}}^{0,1} - 3d_{s_4}^{0,1} + 8d_{s_{11}}^{0,1} + 4d_{s_{32}}^{0,1} + d_{s_{33}}^{0,1} + 2d_{s_{36}}^{0,1} + 7/6,$$

$$d_{s_5}^{2,1} = d_{s_7}^{2,1} + 3d_{s_4}^{0,0} + d_{s_7}^{0,0} - d_{s_9}^{0,0} + d_{s_2}^{2,1} + 2d_{s_3}^{2,1} - 3d_{s_{10}}^{0,0} + 3d_{s_4}^{2,1} + d_{s_{12}}^{1,2} + 2d_{s_{13}}^{1,2} + 3d_{s_{14}}^{1,2} + 2d_{s_{18}}^{0,0} + d_{s_2}^{0,0} + 2d_{s_3}^{0,0} - d_{s_{31}}^{2,1} + d_{s_{33}}^{2,1} - d_{s_{34}}^{2,1} + d_{s_{17}}^{1,2} + 2d_{s_{18}}^{1,2} - 2d_{s_{19}}^{1,2} - d_{s_{20}}^{1,2} - 3d_{s_{22}}^{1,2} - 2d_{s_{23}}^{1,2} + d_{s_{12}}^{2,1} + 3d_{s_{14}}^{2,1} - d_{s_{15}}^{2,1} + d_{s_{17}}^{2,1} - d_{s_5}^{0,0} - 2d_{s_8}^{0,0} + d_{s_{12}}^{0,0} + 2d_{s_{13}}^{0,0} + d_{s_{29}}^{0,0} - 2d_{s_{23}}^{0,0} + d_{s_{33}}^{1,2} - 2d_{s_8}^{2,1} - d_{s_9}^{2,1} + 3d_{s_{14}}^{0,0} - d_{s_{15}}^{0,0} + d_{s_{17}}^{0,0} - 2d_{s_{19}}^{0,0} - d_{s_{20}}^{0,0} - 3d_{s_{22}}^{0,0} + d_{s_2}^{1,2} + 2d_{s_3}^{1,2} + 3d_{s_4}^{1,2} - d_{s_5}^{1,2} + d_{s_{25}}^{0,0} + 2d_{s_{26}}^{0,0} - d_{s_{27}}^{0,0} - 2d_{s_{30}}^{0,0} + d_{s_{33}}^{0,0} - d_{s_{34}}^{0,0} + d_{s_7}^{1,2} - 2d_{s_8}^{1,2} - d_{s_9}^{1,2} - 3d_{s_{10}}^{1,2} - d_{s_{15}}^{1,2} - d_{s_{31}}^{0,0} + d_{s_{25}}^{2,1} + 2d_{s_{26}}^{2,1} - d_{s_{27}}^{2,1} - 2d_{s_{30}}^{2,1} + 2d_{s_{13}}^{2,1} - 3d_{s_{10}}^{2,1} - d_{s_{27}}^{1,2} + d_{s_{29}}^{1,2} - 2d_{s_{30}}^{1,2} - d_{s_{31}}^{1,2} - d_{s_{34}}^{1,2} + 2d_{s_{18}}^{2,1} - 2d_{s_{19}}^{2,1} - d_{s_{20}}^{2,1} - 3d_{s_{22}}^{2,1} - 2d_{s_{23}}^{2,1} + d_{s_{25}}^{1,2} + 2d_{s_{26}}^{1,2} + d_{s_{29}}^{2,1},$$

$$d_{s_{17}}^{2,0} = -d_{s_{36}}^{1,0} + 4d_{s_{32}}^{2,0} - d_{s_{33}}^{1,0} - d_{s_{11}}^{1,0} + d_{s_{30}}^{2,0} - 6d_{s_{34}}^{1,1} + 3d_{s_{33}}^{2,0} - d_{s_{29}}^{1,1} - d_{s_{14}}^{1,0} + 3d_{s_{10}}^{1,0} - 9d_{s_{36}}^{1,1} - 5d_{s_{33}}^{1,1} - 5d_{s_{35}}^{1,1} + d_{s_{12}}^{2,0} - d_{s_{14}}^{2,0} - d_{s_{25}}^{1,0} + d_{s_{19}}^{1,0} + d_{s_{20}}^{1,0} + 9d_{s_4}^{0,0} + 7d_{s_7}^{0,0} + 5d_{s_9}^{0,0} - 2d_{s_{28}}^{1,1} + 3d_{s_{10}}^{0,0} - 5d_{s_{11}}^{1,2} - 3d_{s_{12}}^{1,2} - d_{s_{13}}^{1,2} + d_{s_{14}}^{1,2} + 6d_{s_{18}}^{0,0} + 12d_{s_1}^{0,0} + 11d_{s_2}^{0,0} + 10d_{s_3}^{0,0} + 2d_{s_{31}}^{2,1} + 3d_{s_{32}}^{2,1} + 3d_{s_{33}}^{2,1} + 3d_{s_{34}}^{2,1} + 3d_{s_{35}}^{2,1} + 4d_{s_{36}}^{2,1} - 4d_{s_{16}}^{1,2} - 2d_{s_{17}}^{1,2} - 7d_{s_{19}}^{1,2} - 5d_{s_{20}}^{1,2} - 3d_{s_{21}}^{1,2} - 8d_{s_{22}}^{1,2} - 6d_{s_{23}}^{1,2} + d_{s_{12}}^{2,1} + d_{s_{14}}^{2,1} + d_{s_{15}}^{2,1} + d_{s_{16}}^{2,1} + d_{s_{17}}^{2,1} + 9d_{s_5}^{0,0} + 2d_{s_9}^{1,0} + 8d_{s_6}^{0,0} + 2d_{s_8}^{1,0} + 6d_{s_8}^{0,0} + 12d_{s_{11}}^{0,0} - 6d_{s_{32}}^{1,1} + 11d_{s_{12}}^{0,0} + 10d_{s_{13}}^{0,0} + 3d_{s_{29}}^{0,0} + 4d_{s_{21}}^{0,0} + 2d_{s_{23}}^{0,0} + d_{s_5}^{1,0} - d_{s_{32}}^{1,2} + d_{s_{33}}^{1,2} - d_{s_{26}}^{1,0} + 9d_{s_{14}}^{0,0} - d_{s_{13}}^{1,0} + 9d_{s_{15}}^{0,0} + 8d_{s_{16}}^{0,0} + 7d_{s_{17}}^{0,0} + 6d_{s_{19}}^{0,0} + 5d_{s_{20}}^{0,0} - d_{s_{32}}^{1,0} + 3d_{s_{22}}^{0,0} - d_{s_1}^{2,0} - 2d_{s_2}^{2,0} - 3d_{s_3}^{2,0} - 4d_{s_4}^{2,0} - 2d_{s_5}^{2,0} - 3d_{s_6}^{2,0} - 4d_{s_7}^{2,0} - d_{s_{12}}^{1,0} - 6d_{s_1}^{1,2} - 4d_{s_2}^{1,2} - 2d_{s_3}^{1,2} - 7d_{s_5}^{1,2} - 5d_{s_6}^{1,2} + 3d_{s_{24}}^{2,0} + 2d_{s_{25}}^{2,0} + d_{s_{26}}^{2,0} + 4d_{s_1}^{1,1} + 5d_{s_2}^{1,1} + 6d_{s_3}^{1,1} + 7d_{s_4}^{1,1} + 4d_{s_5}^{1,1} + 5d_{s_6}^{1,1} + 8d_{s_{24}}^{0,0} + 7d_{s_{25}}^{0,0} + 6d_{s_{26}}^{0,0} + 5d_{s_{27}}^{0,0} + 4d_{s_{28}}^{0,0} + 2d_{s_{30}}^{0,0} + d_{s_7}^{1,0} + 4d_{s_{32}}^{0,0} + 3d_{s_{33}}^{0,0} + d_{s_{34}}^{0,0} + d_{s_{11}}^{2,1} + d_{s_{21}}^{1,0} + 2d_{s_{22}}^{1,0} + 2d_{s_{23}}^{1,0} - 2d_{s_{18}}^{2,0} - d_{s_{20}}^{2,0} - 2d_{s_{21}}^{2,0} - d_{s_{22}}^{2,0} - 2d_{s_{23}}^{2,0} + 2d_{s_{27}}^{2,0} + d_{s_{31}}^{1,0} + d_{s_{30}}^{1,0} - 3d_{s_7}^{1,2} - 8d_{s_8}^{1,2} - 6d_{s_9}^{1,2} - 9d_{s_{10}}^{1,2} - 6d_{s_{15}}^{1,2} + 3d_{s_{34}}^{2,0} + d_{s_{12}}^{1,1} + 2d_{s_{13}}^{1,1} + 3d_{s_{14}}^{1,1} + d_{s_{16}}^{1,1} + 2d_{s_{17}}^{1,1} - 2d_{s_{31}}^{1,1} + 3d_{s_{18}}^{1,1} + d_{s_{31}}^{0,0} + 2d_{s_{25}}^{2,1} + 2d_{s_{26}}^{2,1} + 5d_{s_{36}}^{2,0} + 2d_{s_{27}}^{2,1} + 2d_{s_{28}}^{2,1} + 2d_{s_{30}}^{2,1} + d_{s_{13}}^{2,1} + d_{s_{36}}^{1,2} + d_{s_{20}}^{1,1} - 3d_{s_{30}}^{1,1} + 2d_{s_{21}}^{1,1} + d_{s_{23}}^{1,1} - 3d_{s_{24}}^{1,1} + d_{s_6}^{1,0} - 2d_{s_{25}}^{1,1} - d_{s_{26}}^{1,1} - 3d_{s_{27}}^{1,1} + 2d_{s_{11}}^{2,0} + 6d_{s_7}^{1,1} + 4d_{s_8}^{1,1} + 5d_{s_9}^{1,1} + 2d_{s_{35}}^{2,0} + d_{s_{28}}^{2,0} - d_{s_{24}}^{1,0} - 4d_{s_{27}}^{1,2} - 2d_{s_{28}}^{1,2} - 5d_{s_{30}}^{1,2} - 3d_{s_{31}}^{1,2} - 2d_{s_{34}}^{1,2} + d_{s_{18}}^{2,1} + d_{s_{19}}^{2,1} + d_{s_{20}}^{2,1} + d_{s_{21}}^{2,1} + d_{s_{22}}^{2,1} + d_{s_{23}}^{2,1} + 2d_{s_{24}}^{2,1} - 3d_{s_8}^{2,0} - 4d_{s_9}^{2,0} - 4d_{s_{10}}^{2,0} + d_{s_{15}}^{2,0} - 1/2 - 3d_{s_{24}}^{1,2} - d_{s_{25}}^{1,2} + d_{s_{26}}^{1,2} + 4d_{s_{10}}^{1,1} + 2d_{s_{29}}^{2,1},$$

$$d_{s_{35}}^{0,1} = -2d_{s_{30}}^{0,1} - d_{s_{28}}^{0,1} + d_{s_{36}}^{1,0} + d_{s_{33}}^{1,0} - d_{s_{25}}^{0,1} + d_{s_{11}}^{1,0} + 17d_{s_{34}}^{1,1} - d_{s_{12}}^{0,1} + 2d_{s_{14}}^{2,2} + 13d_{s_{29}}^{1,1} - 4d_{s_{11}}^{2,2} + 2d_{s_{20}}^{2,2} + 4d_{s_{21}}^{2,2} + d_{s_{14}}^{1,0} - 3d_{s_{10}}^{1,0} - 4d_{s_1}^{2,2} - 2d_{s_5}^{2,2} - 2d_{s_{19}}^{0,1} + 23d_{s_{36}}^{1,1} + 4d_{s_{26}}^{2,2} + 2d_{s_7}^{2,2} - d_{s_{20}}^{0,1} + 17d_{s_{33}}^{1,1} - 2d_{s_{12}}^{2,2} + 19d_{s_{35}}^{1,1} + d_{s_{25}}^{1,0} - d_{s_{19}}^{1,0} + 4d_{s_{32}}^{2,2} - d_{s_{20}}^{1,0} - 18d_{s_4}^{0,0} - 14d_{s_7}^{0,0} - 10d_{s_9}^{0,0} + 11d_{s_{28}}^{1,1} - 2d_{s_{24}}^{0,1} - 6d_{s_{10}}^{0,0} + 8d_{s_{36}}^{2,2} + 3d_{s_{11}}^{1,2} + 3d_{s_{12}}^{1,2} + 3d_{s_{13}}^{1,2} + 3d_{s_{14}}^{1,2} - 12d_{s_{18}}^{0,0} - 24d_{s_1}^{0,0} - 22d_{s_2}^{0,0} - 20d_{s_3}^{0,0} + 2d_{s_{22}}^{2,2} + 2d_{s_{27}}^{2,2} + 2d_{s_{25}}^{2,2} + 7d_{s_{16}}^{1,2} + 7d_{s_{17}}^{1,2} + 7d_{s_{18}}^{1,2} - 2d_{s_{27}}^{0,1} + 11d_{s_{19}}^{1,2} + 11d_{s_{20}}^{1,2} - 2d_{s_{22}}^{0,1} + 11d_{s_{21}}^{1,2} + 15d_{s_{22}}^{1,2} + 15d_{s_{23}}^{1,2} - 18d_{s_5}^{0,0} - 2d_{s_9}^{1,0} - 16d_{s_6}^{0,0} - 2d_{s_8}^{1,0} - 12d_{s_8}^{0,0} - 24d_{s_{11}}^{0,0} + 15d_{s_{32}}^{1,1} - 22d_{s_{12}}^{0,0} - 20d_{s_{13}}^{0,0} - 6d_{s_{29}}^{0,0} - 8d_{s_{21}}^{0,0} - 4d_{s_{23}}^{0,0} - d_{s_5}^{1,0} + 3d_{s_{32}}^{1,2} + 3d_{s_{33}}^{1,2} + 4d_{s_{28}}^{2,2} + 6d_{s_{29}}^{2,2} + 4d_{s_{30}}^{2,2} + 6d_{s_{31}}^{2,2} + d_{s_{26}}^{1,0} - 18d_{s_{14}}^{0,0} + d_{s_{13}}^{1,0} - 18d_{s_{15}}^{0,0} - 16d_{s_{16}}^{0,0} - 14d_{s_{17}}^{0,0} - 12d_{s_{19}}^{0,0} - 10d_{s_{20}}^{0,0} + d_{s_{32}}^{1,0} - 6d_{s_{22}}^{0,0} + d_{s_{12}}^{1,0} + 7d_{s_1}^{1,2} + 7d_{s_2}^{1,2} + 7d_{s_3}^{1,2} + 7d_{s_4}^{1,2} + 11d_{s_5}^{1,2} + 11d_{s_6}^{1,2} - d_{s_{31}}^{0,1} - 5d_{s_1}^{1,1} - 3d_{s_2}^{1,1} - d_{s_3}^{1,1} + d_{s_4}^{1,1} - 3d_{s_5}^{1,1} - d_{s_6}^{1,1} - 2d_{s_2}^{2,2} - 16d_{s_{24}}^{0,0} - 14d_{s_{25}}^{0,0} - 12d_{s_{26}}^{0,0} - d_{s_{16}}^{0,1} - 10d_{s_{27}}^{0,0} - 8d_{s_{28}}^{0,0} - 4d_{s_{30}}^{0,0} + 6d_{s_{33}}^{2,2} + d_{s_7}^{0,1} - d_{s_7}^{1,0} - d_{s_{10}}^{0,1} + 6d_{s_{34}}^{2,2} + 8d_{s_{35}}^{2,2} - 2d_{s_{15}}^{2,2} + 2d_{s_{17}}^{2,2} + 4d_{s_{18}}^{2,2} - 2d_{s_{34}}^{0,1} - 8d_{s_{32}}^{0,0} - 6d_{s_{33}}^{0,0} - 2d_{s_{34}}^{0,0} + 2d_{s_4}^{2,2} + 2d_{s_{10}}^{2,2} - d_{s_{21}}^{1,0} - 2d_{s_{22}}^{1,0} - 2d_{s_{23}}^{1,0} - d_{s_{31}}^{1,0} - d_{s_{30}}^{1,0} + 11d_{s_7}^{1,2} + 15d_{s_8}^{1,2} + 15d_{s_9}^{1,2} + 19d_{s_{10}}^{1,2} + 7d_{s_{15}}^{1,2} - d_{s_{11}}^{1,1} + d_{s_{12}}^{1,1} + 3d_{s_{13}}^{1,1} + 5d_{s_{14}}^{1,1} + d_{s_{15}}^{1,1} + 3d_{s_{16}}^{1,1} + 5d_{s_{17}}^{1,1} + 13d_{s_{31}}^{1,1} + 7d_{s_{18}}^{1,1} - 2d_{s_{31}}^{0,0} + 4d_{s_{23}}^{2,2} + 3d_{s_{36}}^{1,2} + 3d_{s_{19}}^{1,1} + 5d_{s_{20}}^{1,1} + 11d_{s_{30}}^{1,1} + 7d_{s_{21}}^{1,1} - 2d_{s_{15}}^{0,1} + 5d_{s_{22}}^{1,1} + 7d_{s_{23}}^{1,1} + 7d_{s_{24}}^{1,1} - d_{s_6}^{1,0} + 9d_{s_{25}}^{1,1} + 11d_{s_{26}}^{1,1} + d_{s_{18}}^{0,1} + 9d_{s_{27}}^{1,1} + 2d_{s_9}^{2,2} + d_{s_7}^{1,1} - d_{s_8}^{1,1} + d_{s_9}^{1,1} + d_{s_{24}}^{1,0} + 7d_{s_{27}}^{1,2} + 7d_{s_{28}}^{1,2} + 7d_{s_{29}}^{1,2} + 11d_{s_{30}}^{1,2} + 11d_{s_{31}}^{1,2} + 7d_{s_{34}}^{1,2} - 1/3 + 7d_{s_{35}}^{1,2} + 3d_{s_{24}}^{1,2} + 3d_{s_{25}}^{1,2} + 3d_{s_{26}}^{1,2} - d_{s_1}^{0,1} + d_{s_3}^{0,1} - d_{s_5}^{0,1} + d_{s_{14}}^{0,1} - d_{s_8}^{0,1} + d_{s_{10}}^{1,1} - d_{s_{23}}^{0,1} + 2d_{s_4}^{0,1} - 2d_{s_{11}}^{0,1} - 2d_{s_{32}}^{0,1} - d_{s_{33}}^{0,1} - 2d_{s_{36}}^{0,1},$$

$$d_{s_{36}}^{0,2} = -12d_{s_{34}}^{1,1} - 2d_{s_{14}}^{2,2} - 11d_{s_{29}}^{1,1} + 4d_{s_{11}}^{2,2} - 2d_{s_{20}}^{2,2} - 4d_{s_{21}}^{2,2} + 3d_{s_9}^{0,2} + 3d_{s_{10}}^{0,2} + 4d_{s_1}^{2,2} + 2d_{s_5}^{2,2} + 3d_{s_4}^{0,2} - 15d_{s_{36}}^{1,1} - 4d_{s_{26}}^{2,2} - 2d_{s_7}^{2,2} - 12d_{s_{33}}^{1,1} + 2d_{s_{12}}^{2,2} - 14d_{s_{35}}^{1,1} - 4d_{s_{32}}^{2,2} + 12d_{s_4}^{0,0} + 9d_{s_7}^{0,0} + 6d_{s_9}^{0,0} - 9d_{s_{28}}^{1,1} + 3d_{s_{10}}^{0,0} - 8d_{s_{36}}^{2,2} - 2d_{s_{11}}^{1,2} - 3d_{s_{12}}^{1,2} - 4d_{s_{13}}^{1,2} - 5d_{s_{14}}^{1,2} + 8d_{s_{18}}^{0,0} + 15d_{s_1}^{0,0} + 14d_{s_2}^{0,0} + 13d_{s_3}^{0,0} - 2d_{s_{22}}^{2,2} - 2d_{s_{27}}^{2,2} + 2d_{s_{21}}^{0,2} - 2d_{s_{31}}^{2,1} - 2d_{s_{25}}^{2,2} - 3d_{s_{32}}^{2,1} - 3d_{s_{33}}^{2,1} - 3d_{s_{34}}^{2,1} - 3d_{s_{35}}^{2,1} - 4d_{s_{36}}^{2,1} - 5d_{s_{16}}^{1,2} - 6d_{s_{17}}^{1,2} - 7d_{s_{18}}^{1,2} - 6d_{s_{19}}^{1,2} - 7d_{s_{20}}^{1,2} - 8d_{s_{21}}^{1,2} - 8d_{s_{22}}^{1,2} - 9d_{s_{23}}^{1,2} - d_{s_{12}}^{2,1} - d_{s_{14}}^{2,1} - d_{s_{15}}^{2,1} - d_{s_{16}}^{2,1} - d_{s_{17}}^{2,1} + 11d_{s_5}^{0,0} + 10d_{s_6}^{0,0} + 7d_{s_8}^{0,0} + 15d_{s_{11}}^{0,0} - 10d_{s_{32}}^{1,1} + 14d_{s_{12}}^{0,0} + 13d_{s_{13}}^{0,0} + 4d_{s_{29}}^{0,0} + 5d_{s_{21}}^{0,0} + 2d_{s_{23}}^{0,0} - 2d_{s_{32}}^{1,2} - 3d_{s_{33}}^{1,2} - 4d_{s_{28}}^{2,2} - 6d_{s_{29}}^{2,2} - 4d_{s_{30}}^{2,2} - 6d_{s_{31}}^{2,2} + 3d_{s_1}^{0,2} + 3d_{s_2}^{0,2} + 3d_{s_3}^{0,2} +$$

$2d_{s_{22}}^{0,2}+d_{s_{24}}^{0,2}+d_{s_{25}}^{0,2}+d_{s_{26}}^{0,2}+12d_{s_{14}}^{0,0}+11d_{s_{15}}^{0,0}+10d_{s_{16}}^{0,0}+9d_{s_{17}}^{0,0}+7d_{s_{19}}^{0,0}+6d_{s_{20}}^{0,0}+3d_{s_{22}}^{0,0}-5d_{s_1}^{1,2}-6d_{s_2}^{1,2}-7d_{s_3}^{1,2}-8d_{s_4}^{1,2}-7d_{s_5}^{1,2}-8d_{s_6}^{1,2}+d_{s_1}^{1,1}-d_{s_2}^{1,1}-3d_{s_3}^{1,1}-5d_{s_4}^{1,1}-d_{s_5}^{1,1}-3d_{s_6}^{1,1}+2d_{s_2}^{2,2}+10d_{s_{24}}^{0,0}+9d_{s_{25}}^{0,0}+8d_{s_{26}}^{0,0}+6d_{s_{27}}^{0,0}+5d_{s_{28}}^{0,0}+2d_{s_{30}}^{0,0}+d_{s_{31}}^{0,2}-6d_{s_{33}}^{2,2}-6d_{s_{34}}^{2,2}-8d_{s_{35}}^{2,2}+2d_{s_{15}}^{2,2}-2d_{s_{17}}^{2,2}-4d_{s_{18}}^{2,2}+5d_{s_{32}}^{0,0}+4d_{s_{33}}^{0,0}+d_{s_{34}}^{0,0}-d_{s_{11}}^{2,1}-2d_{s_4}^{2,2}-2d_{s_{10}}^{2,2}+2d_{s_{12}}^{0,2}+2d_{s_{13}}^{0,2}+2d_{s_{14}}^{0,2}+2d_{s_{15}}^{0,2}+2d_{s_{16}}^{0,2}+2d_{s_{17}}^{0,2}+2d_{s_{19}}^{0,2}+2d_{s_{20}}^{0,2}+2d_{s_{23}}^{0,2}-9d_{s_7}^{1,2}-9d_{s_8}^{1,2}-10d_{s_9}^{1,2}-11d_{s_{10}}^{1,2}-4d_{s_{15}}^{1,2}-2d_{s_{12}}^{1,1}-4d_{s_{13}}^{1,1}-6d_{s_{14}}^{1,1}-2d_{s_{15}}^{1,1}-4d_{s_{16}}^{1,1}-6d_{s_{17}}^{1,1}-11d_{s_{31}}^{1,1}-8d_{s_{18}}^{1,1}+d_{s_{31}}^{0,0}-2d_{s_{25}}^{2,1}-2d_{s_{26}}^{2,1}-2d_{s_{27}}^{2,1}-2d_{s_{28}}^{2,1}-2d_{s_{30}}^{2,1}-d_{s_{13}}^{2,1}-4d_{s_{23}}^{2,2}-2d_{s_{36}}^{1,2}-4d_{s_{19}}^{1,1}-6d_{s_{20}}^{1,1}-9d_{s_{30}}^{1,1}-8d_{s_{21}}^{1,1}-6d_{s_{22}}^{1,1}-8d_{s_{23}}^{1,1}-5d_{s_{24}}^{1,1}-7d_{s_{25}}^{1,1}-9d_{s_{26}}^{1,1}-7d_{s_{27}}^{1,1}-2d_{s_9}^{2,2}+3d_{s_7}^{0,2}+3d_{s_5}^{0,2}+3d_{s_6}^{0,2}+3d_{s_8}^{0,2}+2d_{s_{11}}^{0,2}-5d_{s_7}^{1,1}-3d_{s_8}^{1,1}-5d_{s_9}^{1,1}+2d_{s_{18}}^{0,2}+d_{s_{27}}^{0,2}+d_{s_{28}}^{0,2}+d_{s_{30}}^{0,2}-4d_{s_{27}}^{1,2}-5d_{s_{28}}^{1,2}-6d_{s_{29}}^{1,2}-6d_{s_{30}}^{1,2}-7d_{s_{31}}^{1,2}-4d_{s_{34}}^{1,2}-d_{s_{18}}^{2,1}-d_{s_{19}}^{2,1}-d_{s_{20}}^{2,1}-d_{s_{21}}^{2,1}-d_{s_{22}}^{2,1}-d_{s_{23}}^{2,1}-2d_{s_{24}}^{2,1}-5d_{s_{35}}^{1,2}-2d_{s_{24}}^{1,2}-3d_{s_{25}}^{1,2}-4d_{s_{26}}^{1,2}-5d_{s_{10}}^{1,1}-2d_{s_{29}}^{2,1}+2/3+d_{s_{29}}^{0,2}.$

The remaining variables are free. □

Hereinafter, we relabel $d_s^{i,j}$ as $d_1,\dots,d_{324}$ in the sense of Lemma 1.5.7. The $b_s^{i,j}$ are also relabelled as $b_1,\dots,b_{324}$ accordingly. (Recall that $b_s^{i,j}$ was defined as $\binom{5}{s}/8100$.) For a point $\boldsymbol{d}=(d_1,\dots,d_{324})\in\mathcal{D}_2$, we call $\tilde{\boldsymbol{d}}=(d_1,\dots,d_{301})$.

Let $\epsilon>0$ be fixed but small enough. We consider the cube of side 2ϵ centred on $\tilde{\boldsymbol{b}}$

$$\tilde{\mathcal{Q}}_2=\{(d_1,\dots,d_{301})\in\mathbb{R}^{301}\ :\ d_k\in[b_k-\epsilon,b_k+\epsilon],\ \forall k\}$$

and the discrete subset

$$\tilde{\mathcal{J}}_2=\tilde{\mathcal{Q}}_2\cap\left(\frac{1}{n}\mathbb{Z}^{301}\right).$$

Let us define their extension to higher dimensions:

$$\begin{aligned}\mathcal{Q}_2=\{(d_1,\dots,d_{324})\in\mathbb{R}^{324}\ :\ &(d_1,\dots,d_{301})\in\tilde{\mathcal{Q}},\\ &d_k=L_k(d_1,\dots,d_{301},1/6),\ \forall k=302,\dots,324\},\end{aligned}$$

where the L_k's are as in Lemma 1.5.7, and

$$\mathcal{J}_2=\mathcal{Q}_2\cap\left(\frac{1}{n}\mathbb{Z}^{324}\right).$$

Note that $\boldsymbol{b}$ is an interior point of $\mathcal{D}_2$, and that for each k the function $L_k(\cdot,1/6)$ is continuous. Then, if ϵ is chosen small enough, we can ensure that for some $\delta>0$

$$\forall\boldsymbol{d}\in\mathcal{Q}_2,\quad d_k>\delta\text{ and }|d_k-b_k|<\delta,\quad k=1,\dots,324,\tag{1.47}$$

and hence $\mathcal{Q}_2\subset\mathcal{D}_2$. Moreover, since n is always divisible by 6, for each k the function $L_k(\cdot,1/6)$ maps points from $\frac{1}{n}\mathbb{Z}^{301}$ into $\frac{1}{n}\mathbb{Z}$, and so $\mathcal{J}_2\subset\mathcal{I}_2$.

Recalling the definitions of f_2, g_2 and h_2 in (1.24), we define for any $(d_1,\dots,d_{301})\in\tilde{\mathcal{Q}}_2$

$$\left.\begin{array}{l}\tilde{f}_2(d_1,\dots,d_{301})=f_2(d_1,\dots,d_{324})\\ \tilde{g}_2(d_1,\dots,d_{301})=g_2(d_1,\dots,d_{324})\end{array}\right.,\quad\text{where }d_k=L_k(d_1,\dots,d_{301},1/6),\ \forall k=302,\dots,324.$$

From Lemma 1.5.6 and by straightforward computations we obtain the following:

Lemma 1.5.8. *The following statements hold:*

- *Under the Maximum Hypothesis, f_2 has a unique maximum in $\mathcal{D}_2$ at $\boldsymbol{b}$.*
- *Under the Maximum Hypothesis, $\tilde{f}_2$ has a unique maximum in $\tilde{\mathcal{Q}}_2$ at $\tilde{\boldsymbol{b}}$, with $e^{f_2(\boldsymbol{b})}=e^{\tilde{f}_2(\tilde{b})}=\frac{25}{24}5^{5/2}\approx 58.2309$.*

- *The Hessian* $\tilde{H}_2$ *of* $\tilde{f}_2$ *at* $\tilde{\boldsymbol{b}}$ *is negative definite, and* $\det \tilde{H}_2 = -2^{175}3^{1078}5^{310}7^{12}11^{14}13\ 17\ 79^4$.
- $\tilde{g}_2(\tilde{\boldsymbol{b}}) = 2^{90}3^{558}5^{171} \neq 0$.
- *Both* $\tilde{f}_2$ *and* $\tilde{g}_2$ *are of class* $\mathcal{C}^\infty$ *in* $\tilde{\mathcal{Q}}_2$.

We compute the contribution to $\mathbf{E}(X^2)$ of the terms around $\boldsymbol{b}$ to get the following result.

Lemma 1.5.9. *Under the Maximum Hypothesis,*

$$\sum_{\boldsymbol{d}\in\mathcal{J}_2} q_2(n,\boldsymbol{d})e^{f_2(\boldsymbol{d})n} \sim \frac{(2\pi n)^{301/2}}{\sqrt{|\det \tilde{H}_2|}}\tilde{g}_2(\tilde{\boldsymbol{b}})e^{n\tilde{f}_2(\tilde{\boldsymbol{b}})} = \frac{2^3 3^{19}5^{16}(2\pi n)^{301/2}}{7^6 11^7 79^2\sqrt{2\ 13\ 17}}\left(\frac{25}{24}\right)^n 5^{5n/2}.$$

Proof. From (1.47), we see that for all $\boldsymbol{d} \in \mathcal{J}_2 \subset \mathcal{Q}_2$ we must have $d_k > \delta\ \forall k$. Thus, from their definition, all the $m_{r,t}^{i,j}$ are bounded away from 0, $q_2(n,\boldsymbol{d}) \sim g_2(\boldsymbol{d})$ and we can write

$$\sum_{\boldsymbol{d}\in\mathcal{J}_2} q_2(n,\boldsymbol{d})e^{f_2(\boldsymbol{d})n} \sim \sum_{\mathcal{J}_2} g_2(\boldsymbol{d})e^{nf_2(\boldsymbol{d})} = \sum_{\tilde{\mathcal{J}}_2} \tilde{g}_2(\tilde{\boldsymbol{d}})e^{n\tilde{f}_2(\tilde{\boldsymbol{d}})}. \tag{1.48}$$

The remaining of the argument is analogous to that in the proof of Lemma 1.4.8 □

Now we deal with the remaining terms of the sum.

Lemma 1.5.10. *Under the Maximum Hypothesis, there exists some positive real* $\alpha < e^{f_2(\boldsymbol{b})}$ *s.t.* $\sum_{\mathcal{I}_2\setminus\mathcal{J}_2} q_2(n,\boldsymbol{d})e^{f_2(\boldsymbol{d})n} = o(\alpha^n)$.

Proof. Let B be the topological closure of $\mathcal{D}_2 \setminus \mathcal{Q}_2$. We recall from Lemma 1.5.8 that f_2 has a unique maximum in $\mathcal{D}_2$ at point $\boldsymbol{b} \notin B$. Then, since B is a compact set and f_2 is continuous, there must be some real $\beta < f_2(\boldsymbol{b})$ such that $f_2(\boldsymbol{x}) \leq \beta\ \forall \boldsymbol{x} \in B$. Now we observe that all terms in the sum $\sum_{\mathcal{I}_2\setminus\mathcal{J}_2} q_2(n,\boldsymbol{d})e^{f_2(\boldsymbol{d})n}$ can be uniformly bounded by $Cn^{162}e^{\beta n}$, for some fixed constant C. Note furthermore that there are at most $(n+1)^{324}$ terms in the sum. Hence, the result holds by taking for instance $\alpha = (e^\beta + e^{f_2(\boldsymbol{b})})/2$. □

From Lemmata 1.5.9 and 1.5.10, we get

$$\sum_{\mathcal{I}_2} q_2(\boldsymbol{d})e^{f_2(\boldsymbol{d})n} \sim \frac{2^3 3^{19}5^{16}(2\pi n)^{301/2}}{7^6 11^7 79^2\sqrt{2\ 13\ 17}}\left(\frac{25}{24}\right)^n 5^{5n/2},$$

which together with Lemma 1.5.1, gives us the following result.

Theorem 1.5.11. *Under the maximum hypothesis,*

$$\mathbf{E}(X^2) \sim \frac{2^2 3^{19}5^{16}}{7^6 11^7 79^2\sqrt{13\ 17}}\frac{1}{(2\pi n)^2}\left(\frac{25}{24}\right)^n.$$

1.6 From Configurations to Graphs

In order to transfer the results obtained in Sections 1.4 and 1.5 to $\mathcal{G}(n,5)$ and prove Theorem 1.2.1, we need to consider the restriction of $\mathcal{P}(n,5)$ to *simple* pairings (i.e. those without loops and multiple edges). We write $\mathbf{P}^*$ and $\mathbf{E}^*$ to denote probability and expectation conditional to the event "$\mathcal{P}(n,5)$ is simple".

By using similar techniques to the ones developed in [13] (see also Theorem 2.6 in [78]), we get:

Lemma 1.6.1. *Let C be any fixed balanced 3-colouring of n vertices. Then,*

$$\mathbf{P}(\mathcal{P}(n,5) \text{ is simple} \mid C \in \mathcal{R}_{\mathcal{P}(n,5)})$$

is bounded away from 0, independently of C and n.

Proof. We first observe that this probability does not depend of C, since all balanced colourings are essentially the same. We can safely relabel the cells and assume without loss of generality that C is such that the first $n/3$ cells have colour 0, the next $n/3$ ones have colour 1, and the last $n/3$ ones have colour 2.

We condition upon the event that $C \in \mathcal{R}_{\mathcal{P}(n,5)}$. In this new probability space, we aim to prove that

$$\mathbf{P}(\mathcal{P}(n,5) \text{ is simple}) \sim e^{-\frac{392}{75}}.$$

A parallel couple is a set of two different pairs which connect the same two cells. The pairs of the parallel couple are not ordered. A colourable pairing cannot have any loop, since this would violate the colourability. Thus, if $\mathcal{P}(n,5)$ satisfies $C \in \mathcal{R}_{\mathcal{P}(n,5)}$ without any parallel couple, the $\mathcal{P}(n,5)$ must be simple, since the corresponding multigraph has no loops and no multiple edges.

A parallel couple is determined by giving two cells of different colour, two points in each one, and one way of matching them. Let Y be the number of parallel couples. Let A be the class of possible parallel couples. Then

$$Y = \sum_{a \in A} Y_a,$$

where Y_a is the indicator variable of the event "$\mathcal{P}(n,5)$ has a as a parallel couple".

For a fixed k,

$$\mathbf{E}[Y]_k = \sum_{a_1,\dots,a_k \in A}^{*} \mathbf{P}\Big(\bigwedge_{i=1}^{k} (Y_{a_i} = 1)\Big), \tag{1.49}$$

where $\sum^{*}$ denotes a sum along k-tuples of pairwise different indices.

For some k-tuples $(a_1, \dots, a_k)$, the term $\mathbf{P}(\bigwedge_{i=1}^{k} (X_{a_i} = 1))$ equals 0. In fact, there might be incompatible pairs involved in different parallel couples of the k-tuple or the locally rainbow property might be violated. We call these k-tuples non-feasible (or of type 1). When there are not such incompatibilities, we call the k-tuples feasible. Furthermore, for feasible k-tuples, sometimes one same pair is involved in more than one parallel couple a_i. We call this a repetition. Those feasible k-tuples with some repetitions are called of type 2, and those ones with no repetitions are called of type 3.

By definition, the terms corresponding to k-tuples of type 1 have no weight in the sum in (1.49).

Next, we deal with terms corresponding to k-tuples of type 3. First, we compute the number of such k-tuples. Observe that the number of possible parallel couples is

$$|A| = 3 \cdot (n/3)^2 \cdot 10^2 \cdot 2,$$

since we choose which 2 colours are involved (3 ways), one cell for each of these colours ($(n/3)^2$ ways), 2 points for each of these vertices ($\binom{5}{2}^2 = 10^2$ ways), and the way of matching the points creating a parallel couple (2 ways). Thus an upper bound for the number of k-tuples of type 3 would be

$$|A|^k = \left[6 \cdot (10n/3)^2\right]^k . \tag{1.50}$$

We obtain a lower bound as follows. When we choose each a_i, we do not to use any cell involved in $a_1, \ldots, a_{i-1}$ to insure the compatibility of the pairs involved in $a_1, \ldots, a_k$ and to avoid repetitions. Then, the number of k-tuples of type 3 is at least

$$\left[6 \cdot (10(n/3 - 2k))^2\right]^k . \tag{1.51}$$

As k is fixed, both quantities in (1.50) and (1.51) coincide asymptotically when n grows large.

We are interested in bounding $\mathbf{P}(\bigwedge_{i=1}^k (Y_{a_i} = 1))$ for k-tuples of type 3. Let us denote the event $\bigwedge_{i=1}^k (Y_{a_i} = 1)$ by $\mathcal{E}$.

An arrangement of the points (arrangement for short) consists in specifying for each point the colour of the cell of the point to which it is matched. Of course, we only consider those arrangements which lead to locally rainbow colourings. Let H be the set of all arrangements which are compatible with the event $\mathcal{E}$. We have

$$\mathbf{P}(\mathcal{E}) = \sum_{h \in H} \mathbf{P}(\mathcal{E} \mid h)\, \mathbf{P}(h).$$

As long as h is compatible with $\mathcal{E}$, this $\mathbf{P}(\mathcal{E} \mid h)$ can be bounded between $\left(\frac{1}{5n/6}\right)^{2k}$ and $\left(\frac{1}{5n/6-2k}\right)^{2k}$, independently of the specific h. Thus,

$$\mathbf{P}(\mathcal{E}) \sim \left(\frac{1}{5n/6}\right)^{2k} \mathbf{P}(h \in H), \tag{1.52}$$

where $\mathbf{P}(h \in H)$ is the probability that the arrangement of the points is compatible with the k-tuple of parallel couples.

In order to compute $\mathbf{P}(h \in H)$, we first observe the following: Let us just consider the cells of any specific colour. Then, almost all arrangements of the points have the typical proportions of cells with each type of spectrum. Thus, we can safely condition to this event.

Let us assume w.l.o.g. that the first parallel pair a_1 consists of two points a and b in a 0-coloured cell u, and two points c and d in a 1-coloured cell v, so that a and c must be matched and so must b and d. For h to be compatible with this, cell u must have at least 2 points matched to points in 1-coloured cells and a, b must be among them. Moreover cell v

must have at least 2 points matched with points in 0-coloured cells and c, d must be among them. This holds with probability

$$\left(\frac{1}{10}\mathbf{P}(2)+\frac{3}{10}\mathbf{P}(3)+\frac{3}{5}\mathbf{P}(4)\right)^2,$$

where $\mathbf{P}(i)$ is the probability that a given cell has spectrum $(i, 5-i)$.

More generally, taking into account that the typical proportions of cells of spectrum $(1,4)$, $(2,3)$, $(3,2)$, $(4,1)$ are respectively $1/6$, $1/3$, $1/3$, $1/6$, we conclude that the probability $\mathbf{P}(h \in H)$ that an arrangement h is compatible with the k-tuple $(a_1, \ldots, a_k)$ is bounded between $\left(\frac{1}{10}\frac{n/9}{n/3}+\frac{3}{10}\frac{n/9}{n/3}+\frac{3}{5}\frac{n/18}{n/3}\right)^{2k}$ and $\left(\frac{1}{10}\frac{n/9}{n/3-k}+\frac{3}{10}\frac{n/9}{n/3-k}+\frac{3}{5}\frac{n/18}{n/3-k}\right)^{2k}$. Hence,

$$\mathbf{P}(h \in H) \sim \left(\frac{1}{10}\frac{1}{3}+\frac{3}{10}\frac{1}{3}+\frac{3}{5}\frac{1}{6}\right)^{2k} = \left(\frac{7}{30}\right)^{2k}.$$

Then, from (1.52), we get:

$$\mathbf{P}(\mathcal{E}) = \left(\frac{7}{25n}\right)^{2k}.$$

We conclude that the weight in the sum in (1.49) due to the terms corresponding to k-tuples of type 3 is asymptotically

$$\sim \left[6\cdot(10n/3)^2\right]^k \left(\frac{7}{25n}\right)^{2k} = \left(\frac{392}{75}\right)^k.$$

We claim that the terms corresponding to k-tuples of type 2 have negligible weight in the sum in (1.49) asymptotically for n growing large. The calculations are analogous to those for tuples of type 3 and we omit them here. The terms $\mathbf{P}(\bigwedge_{i=1}^{k}(X_{a_i}=1))$ happen to be larger in this situation, but the number of terms in the sum is much smaller.

Then, we conclude that

$$\mathbf{E}[Y]_k = \sum_{a_1,\ldots,a_k\in A}^{*} \mathbf{P}(\bigwedge_{i=1}^{k}(Y_{a_i}=1)) \sim \left(\frac{392}{75}\right)^k, \tag{1.53}$$

and Y is asymptotically Poisson of parameter $\frac{392}{75}$ (see Theorem 1.23 in [15]). Thus,

$$\mathbf{P}(\mathcal{P}(n,5) \text{ is simple} \mid C \in \mathcal{R}_{\mathcal{P}(n,5)}) \sim e^{-\frac{392}{75}}. \qquad \square$$

We are now ready to prove the main result of this chapter: The fact that under the Maximum Hypothesis, for n divisible by 6 the chromatic number of $\mathcal{G}(n,5)$ is 3, with probability bounded away from 0.

Proof of Theorem 1.2.1. From Lemma 1.6.1 and from the fact that $\mathbf{P}(\mathcal{P}(n,5)$ is simple$)$ is also bounded away from 0 independently of n (see, e.g., [78]), we obtain:

$$\mathbf{E}^*X = \frac{\left|\{(P,C)\ :\ P\in\mathbb{P}(n,5) \text{ is simple},\ C\in\mathcal{R}_P\}\right|}{\left|\{P\ :\ P\in\mathbb{P}(n,5) \text{ is simple}\}\right|}$$

$$
\begin{aligned}
&= \frac{\big|\{(P,C)\ :\ P \in \mathbb{P}(n,5),\ C \in \mathcal{R}_P\}\big|\ \mathbf{P}(P \text{ is simple} \mid C \in \mathcal{R}_P)}{\big|\{P\ :\ P \in \mathbb{P}(n,5)\}\big|\ \mathbf{P}(\mathcal{P}(n,5) \text{ is simple})} \\
&= \Theta(\mathbf{E}X) \qquad (1.54)
\end{aligned}
$$

For the second moment, it is sufficient the following weaker bound:

$$
\begin{aligned}
\mathbf{E}^*(X^2) &= \frac{\big|\{(P,C_1,C_2)\ :\ P \in \mathbb{P}(n,5) \text{ is simple},\ C_1,C_2 \in \mathcal{R}_P\}\big|}{\big|\{P\ :\ P \in \mathbb{P}(n,5) \text{ is simple}\}\big|} \\
&\le \frac{\big|\{(P,C_1,C_2)\ :\ P \in \mathbb{P}(n,5),\ C_1,C_2 \in \mathcal{R}_P\}\big|}{\big|\{P \mid P \in \mathbb{P}(n,5)\}\big|\ \mathbf{P}(\mathcal{P}(n,5) \text{ is simple})} \\
&= \Theta(\mathbf{E}(X^2)) \qquad (1.55)
\end{aligned}
$$

Therefore, we have that $\mathbf{E}^*X = \Theta(\mathbf{E}X)$ and $\mathbf{E}^*(X^2) = O\left(\mathbf{E}(X^2)\right)$. The result follows from (1.1) and in view of Theorems 1.4.10 and 1.5.11. □

1.7 The Maximum Hypothesis and its Empirical Validation

In this section we describe the evidence which supports the Maximum Hypothesis. Recall that the hypothesis asserts that for a certain four-variable function $F(\boldsymbol{n})$ on a bounded domain $\mathcal{N}$, $F(\boldsymbol{n})$ has a unique global maximum at the point $(1/9, 1/9, 1/9, 1/9)$. There are two equivalent definitions for the function F, which give two possible approaches to numerical verification of the Maximum Hypothesis. All the relevant definitions and equations are repeated here, so that the definition of F in this section is self-contained.

We first define the domain $\mathcal{N}$ of F. This is the set of all non-negative real vectors $\boldsymbol{n} = (n_1, \ldots, n_4)$ satisfying

$$
n_1+n_2 \le \frac{1}{3}, \quad n_3+n_4 \le \frac{1}{3}, \quad n_1+n_3 \le \frac{1}{3}, \quad n_2+n_4 \le \frac{1}{3}, \quad n_1+n_2+n_3+n_4 \ge \frac{1}{3}. \qquad (1.56)
$$

For each $\boldsymbol{n}$ in $\mathcal{N}$, we define the following nine values

$$
\begin{array}{lll}
n^{0,0} = n_1, & n^{0,1} = n_2, & n^{0,2} = 1/3 - n_1 - n_2 \\
n^{1,0} = n_3, & n^{1,1} = n_4, & n^{1,2} = 1/3 - n_3 - n_4 \\
n^{2,0} = 1/3 - n_1 - n_3, & n^{2,1} = 1/3 - n_2 - n_4, & n^{2,2} = n_1 + n_2 + n_3 + n_4 - 1/3
\end{array} \qquad (1.57)
$$

We need some more definitions before stating how to compute F at any point in its domain.

A *spectrum* s is a 2×2 non-negative integer matrix such that each row and column sum is at least 1, and the sum of all four entries is 5. We index the rows and columns by -1 and 1, with -1 for the first row or column. So

$$
s = \begin{bmatrix} s_{-1,-1} & s_{-1,1} \\ s_{1,-1} & s_{1,1} \end{bmatrix}.
$$

Let $\mathcal{S}$ denote the set of all spectra (cf. (1.9)). Note that $|\mathcal{S}| = 36$. This definition of spectrum is the same as the one presented in Section 1.3.2.

For each ordered pair (i,j), $i,j \in \mathbb{Z}_3$, and spectrum $s \in \mathcal{S}$, introduce a real variable $d_s^{i,j}$, called a *spectral* variable. Also define matrices $m^{i,j}$ (cf. (1.10)) by

$$m^{i,j} = \sum_{s \in \mathcal{S}} d_s^{i,j}\, s. \tag{1.58}$$

Consider the following as constraints for all i and j:

$$\sum_{s \in \mathcal{S}} d_s^{i,j} = n^{i,j}, \quad d_s^{i,j} \geq 0\ \forall s \in \mathcal{S}, \tag{1.59}$$

where the constants $n^{i,j}$ are defined in (1.57), and

$$m_{r,t}^{i,j} = m_{-r,-t}^{i+r,j+t}, \quad \text{for } i,j \in \{0,1,2\} \text{ and } r,t \in \{-1,1\}, \tag{1.60}$$

where the arithmetic in the indices is modulo 3.

Let $\mathcal{D}(\boldsymbol{n})$ be the set of tuples $\boldsymbol{d} = (d_s^{i,j})_{i,j \in \mathbb{Z}_3, s \in \mathcal{S}_2}$ satisfying the above constraints. For each $\boldsymbol{d} \in \mathcal{D}(\boldsymbol{n})$, let $\hat{F}(\boldsymbol{d})$ be the function defined as

$$\hat{F}(\boldsymbol{d}) = \left(\prod_{i,j,s} \left(\frac{\binom{5}{s}}{d_s^{i,j}} \right)^{d_s^{i,j}} \right) \left(\prod_{i,j,r,t} (m_{r,t}^{i,j})^{\frac{1}{2} m_{r,t}^{i,j}} \right). \tag{1.61}$$

Note that $\hat{F}$ is a function of 9×36 constrained variables. Since $\hat{F}$ is continuous in the compact domain defined by the constraints, it must have a maximum. Set $F(\boldsymbol{n})$ to be the value of this maximum.

In Section 1.5 we defined the same function $\hat{F}(\boldsymbol{d})$ but extended it to the larger domain $\mathcal{D}_2$ where the $n^{i,j}$ are not fixed but rather take any value in (1.57).

For the second definition of F, define the matrix function (also defined in (1.30))

$$\Phi \begin{bmatrix} x & y \\ z & w \end{bmatrix} = (x+y+z+w)^5 - (x+y)^5 - (x+z)^5 - (y+w)^5 - (z+w)^5 + x^5 + y^5 + z^5 + w^5. \tag{1.62}$$

For each of the nine possible pairs (i,j), $i,j \in \mathbb{Z}_3$, let $\mu^{i,j}$ and $m^{i,j}$ be 2×2 matrices whose rows and columns are indexed by -1 and 1 (as in the first definition of F). For each such (i,j), consider the 4×4 system (cf. (1.31))

$$\frac{\partial\ \Phi \mu^{i,j}}{\partial \mu_{r,t}^{i,j}}\ \mu_{r,t}^{i,j} = m_{r,t}^{i,j}, \quad r,t = -1, 1. \tag{1.63}$$

As in (1.27), we consider the following constraints, for all such i and j, and all $r,t \in \{-1,1\}$,

$$\begin{array}{rcl} m_{r,t}^{i,j} & \geq & 0, \\ m_{r,t}^{i,j} + m_{r,-t}^{i,j} & \leq & 4(m_{-r,t}^{i,j} + m_{-r,-t}^{i,j}), \\ m_{r,t}^{i,j} + m_{-r,t}^{i,j} & \leq & 4(m_{r,-t}^{i,j} + m_{-r,-t}^{i,j}). \end{array} \tag{1.64}$$

For each $\boldsymbol{n}$ in $\mathcal{N}$, we define $\mathcal{M}(\boldsymbol{n})$ to be the set of all vectors $\boldsymbol{m} = (m^{i,j})_{i,j \in \mathbb{Z}_3}$ of 2×2 matrices $m^{i,j}$ satisfying (1.64), (1.60), and also

$$\sum_{r,t} m_{r,t}^{i,j} = 5 n^{i,j}, \tag{1.65}$$

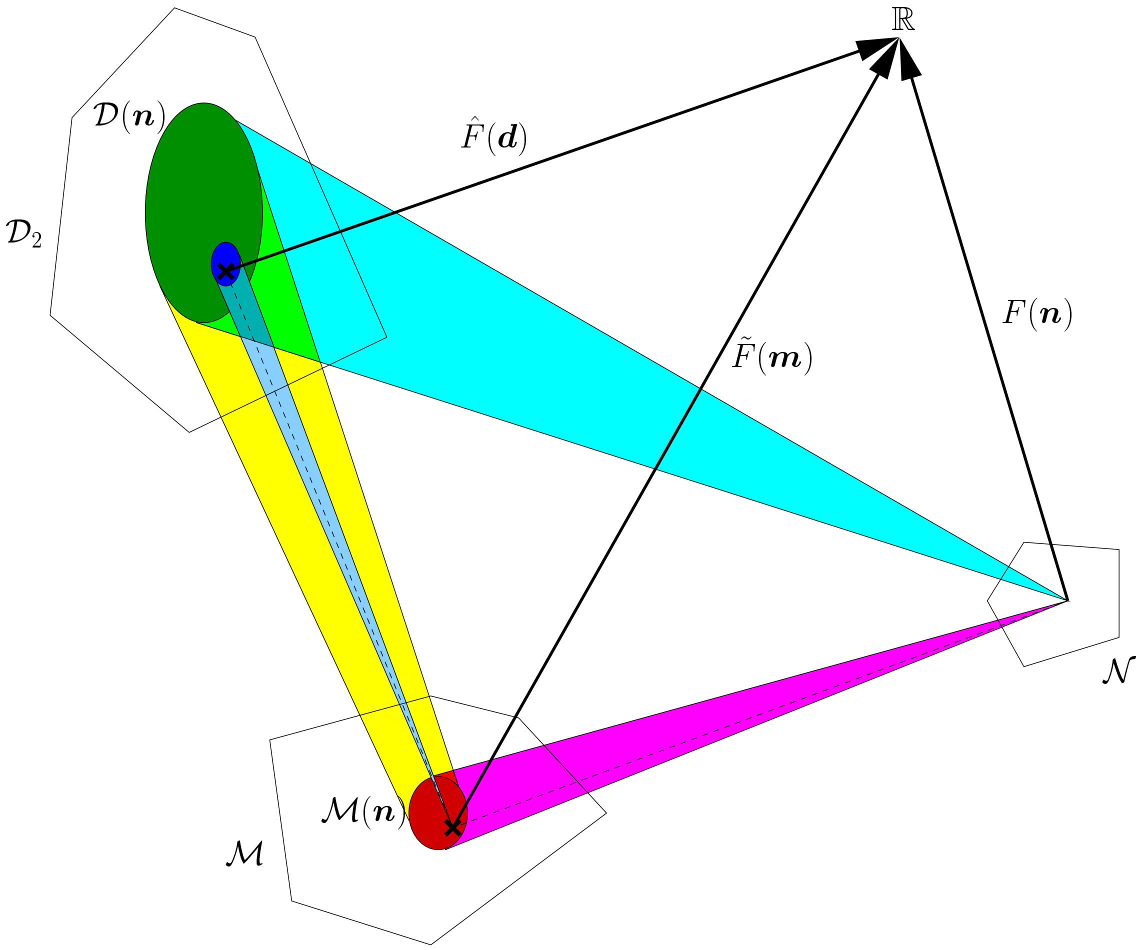


Figure 1.1: Relationships between the sets involved in the definition of F. The crosses denote the maximisers of $\hat{F}$ and $\tilde{F}$ over $\mathcal{D}(\boldsymbol{n})$ and $\mathcal{M}(\boldsymbol{n})$, respectively.

where the constants $n^{i,j}$ are defined in (1.57). We observe that $\mathcal{M}(\boldsymbol{n})$ is a polytope of dimension 9. Given a vector $\boldsymbol{m}$ of matrices $(m^{i,j})_{i,j\in\mathbb{Z}_3}$ in the interior of $\mathcal{M}(\boldsymbol{n})$, define

$$\tilde{F}(\boldsymbol{m}) = \prod_{i,j,r,t} \left(\frac{(m^{i,j}_{r,t})^{\frac{1}{2}}}{\mu^{i,j}_{r,t}} \right)^{m^{i,j}_{r,t}}, \tag{1.66}$$

with the $\mu^{i,j}_{r,t}$ given in terms of the $m^{i,j}_{r,t}$ by (1.63) and required to be strictly positive. In Section 1.5, we show that for $\boldsymbol{m}$ in the interior of $\mathcal{M}(\boldsymbol{n})$ the $\mu^{i,j}_{r,t}$ variables are determined uniquely, and that $\tilde{F}$ can be continuously extended to the boundary points of the polytope.

Our second definition of $F(\boldsymbol{n})$ is the maximum of $\tilde{F}(\boldsymbol{m})$ over all $\boldsymbol{m}$ lying in $\mathcal{M}(\boldsymbol{n})$. This is well defined by continuity of the function and compactness of the domain.

We observe that Lemma 1.5.3 shows the equivalence of these two alternative definitions of F. Figure 1.1 shows the different domains involved in the definition of F.

One important piece of evidence supporting the Maximum Hypothesis is the following theorem.

Theorem 1.7.1. *The function $F(\boldsymbol{n})$ has a local maximum at the point $(1/9, 1/9, 1/9, 1/9)$.*

Proof. By Lemma 1.5.5, $\hat{G}(\boldsymbol{d})$ takes its maximum in $D'(1/9)$ uniquely at the point where all the d_s are equal to $\binom{5}{s}/8100$. It follows by continuity of $\hat{F}$ that the only values of $\hat{F}$ that can contribute to the maximum value of F in a neighbourhood of $(1/9, 1/9, 1/9, 1/9)$ must come from $\boldsymbol{d}$ in a neighbourhood of $(\binom{5}{s}/8100)_{s\in\mathcal{S}}$. The Hessian, computed using Maple, shows that $\hat{F}$ has a local maximum here, so the local maximum of F at $(1/9, 1/9, 1/9, 1/9)$ follows. □

Next, we describe the empirical evidence that F has a unique maximum at $(\frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9})$.

Let $\boldsymbol{n}$ be any fixed vector in $\mathcal{N}$. Recall the definition of $\mathcal{M}'(c)$ from Section 1.5.1. We observe that the projection of $\mathcal{M}(\boldsymbol{n})$ to the (i, j) coordinate is $\mathcal{M}'(n^{i,j})$. Let us momentarily relax the constraints in (1.60), and consider separately each factor

$$\tilde{G}^{i,j} = \prod_{r,t} \left(\frac{(m^{i,j}_{r,t})^{\frac{1}{2}}}{\mu^{i,j}_{r,t}} \right)^{m^{i,j}_{r,t}},$$

to be defined in $\mathcal{M}'(n^{i,j})$. We note that $\mathcal{M}'(n^{i,j})$ is a polytope of dimension 3, so $G^{i,j}$ can be written in terms of three free variables. In order to show that $\log \tilde{G}^{i,j}$ is concave, it is sufficient to see that the 3×3 Hessian matrix is negative definite over the domain. This was numerically confirmed by direct computation over a fine grid of points in the domain. Having experimentally confirmed the concavity of the logarithm of each factor of $\tilde{F}$, we conclude the concavity of $\log \tilde{F}$. Moreover, this concavity is not affected by adding the constraints in (1.60), previously relaxed.

The procedure we use is based on the concavity of $\log \tilde{F}$. We sweep the domain $\mathcal{N}$ of F. Variables n_1, n_2, n_3, n_4 take all non-negative values satisfying (1.56) in a grid of 200 steps per dimension. For each point $\boldsymbol{n} = (n_1, \ldots, n_4)$ we compute $F(\boldsymbol{n})$ as follows.

Procedure for computing $F(\boldsymbol{n})$.

1. We compute the nine overlap variables $n^{i,j}$ from (1.57). (The sweep avoids a fine layer of width 1/1000 around the boundary.)

2. We set $\boldsymbol{m}_0$ to be an interior point of $\mathcal{M}(\boldsymbol{n})$.

3. Starting from $\boldsymbol{m}_0$, we numerically maximise $\log \tilde{F}$ in $\mathcal{M}(\boldsymbol{n})$ by a Newton-like iterative method. This works well due to the concavity of $\log \tilde{F}$. As observed before, the maximisation domain has dimension 9. In fact, the elements in $\mathcal{M}(\boldsymbol{n})$ can be expressed in terms of the nine coordinates $m^{i,j}_{1,1}$ by

$$\begin{aligned} m^{i,j}_{-1,-1} &= m^{i-1,j-1}_{1,1} \\ m^{i,j}_{1,-1} &= \tfrac{1}{2}\left(a^{i,j} + a^{i+1,j-1} - a^{i-1,j+1}\right) \\ m^{i,j}_{-1,1} &= \tfrac{1}{2}\left(a^{i,j} + a^{i-1,j+1} - a^{i+1,j-1}\right), \end{aligned}$$

where

$$a^{i,j} = 5n^{i,j} - m^{i,j}_{1,1} - m^{i-1,j-1}_{1,1}.$$

Then we must write $\log \tilde{F}$ in terms of the $m_{1,1}^{i,j}$ and, at each step of the maximisation, compute the gradient with respect to these nine variables. From the proof of Lemma 1.5.5 and in view of (1.42), we can get rid of the derivatives of the $\mu_{r,t}^{i,j}$ and express this gradient just in terms of the $m_{r,t}^{i,j}$ and $\mu_{r,t}^{i,j}$. Hence, each iteration of the maximisation algorithm requires the solution of the nine 4×4 systems in (1.63), which are known to have a unique positive solution.

The maximum obtained is $F(\boldsymbol{n})$.

Recall from Lemma 1.5.6 that $F(\frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}) = 5^{5/2}25/24 \approx 58.2309$. The values of F we obtained by this procedure for each $\boldsymbol{n}$ were always below $F(\frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9})$. There were some points in the domain where a value over 58 was obtained. These points were all near $(\frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9})$. Around these points we made an additional scan of the neighbourhood with step-size 1/8000. The values obtained were always less than $F(\frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9})$.

The sweeping procedure was carried out by using the IBM's Mare-Nostrum supercomputer in the Barcelona Supercomputing Center, which consists of 2.268 dual 64-bit processor blade nodes with a total of 4.536 2.2 GHz PPC970FX processors.

To get rid of the Maximum Hypothesis, all it remains is bounding the derivatives of F in the interior of its domain, which remains an open issue at the time of this work.

1.8 Extensions

This section contains a brief overview about two extensions of Theorem 1.2.1 which are not part of this thesis. The reader should refer to [51] and [23] for further details.

It can be shown by using the small subgraph conditioning method (see [78]) that Theorem 1.2.1 can be extended to hold a.a.s. The main obstacle for proving that the number of locally rainbow balanced colourings is a.a.s. non-zero is the large variance of X, which is a constant times the square of $\mathbf{E}X$. However it turns out that the distribution of X is sensibly affected by the number of cycles of given constant lengths. In fact, if we divide the graphs into groups according to their short cycle counts, most of the variance of X comes from the fact that the graphs of different groups have a different expected number of such colourings. The small subgraph conditioning method overrides this fact by conditioning upon the number of short cycles, in order to get the ratio between the $\mathbf{Var}(X)$ and $\mathbf{E}(X)$ arbitrarily small.

In [23] there is also an argument which covers the case n is even but not necessarily divisible by 6. Loosely, two (or four) vertices are removed from the original graph, and the edges of their neighbours readjusted so that we obtain a 5-regular graph with the number of vertices divisible by 6 and with a set of distinguished edges. The distribution of the resulting graph is not exactly uniform over the set of 5-regular graphs with that particular number of distinguished edges, but it is close to it. Then, by an argument analogous to the one for 5-regular graphs, it is shown that these graphs with a set of distinguished edges admit a.a.s. some locally rainbow balanced colouring with some prefixed colours on the vertices of those distinguished edges. By choosing the appropriate colours on these vertices, this colouring can be extended to a legal 3-colouring of the original graph.

2

Walkers on the Cycle and the Grid

2.1 Mobile Ad-Hoc Networks

A *wireless ad-hoc network* (cf. [7, 41, 81]) is an autonomous and decentralised collection of agents, also denoted as nodes, placed in some geographical region, and able to receive and transmit information by means of radio-frequency or an equivalent wireless device. Each agent is equipped with an omnidirectional antenna of limited power, and thus can stablish communication in one hop with any other agent which is located close enough. In addition, two agents can communicate in several hops by means of intermediate transmitters. In many applications (e.g. in sensor networks), the number of nodes in such a network is large and they are placed without any prescribed strategy, thus at random. Then, it is natural to model the agents as points in a finite random-point process over some domain, and represent the network by a random graph defined the obvious way, with the agents being the vertices and setting an edge between any two agents falling within the communication range of each other. The resulting random graph is of a rather different nature than the classical Erdős and Rényi model of random graphs, in which edges are selected independently with some common probability. In fact, for a wireless ad-hoc network, the presence of wireless links between nodes presents a significant local correlation. This fact motivates an alternative model of random graph due to Gilbert in 1961 [35]: *"To construct a random plane network, pick points from the plane by a Poisson process with density D points per unit area. Next joint each each pair of points by a line if they are at distance less that R."* Observe that in such a network the nodes are selected from an infinite random-point process over the real plane, and that each agent has the same transmission power. Several variants of this model have been widely studied and are usually referred to as *random geometric graphs*. We give a brief outlook on these graphs in Section 2.2. The theoretical properties of these graphs, have been broadly used by many researchers to design algorithms for more efficient coverage, communication and energy savings in ad-hoc networks, and in particular for sensor networks. Various aspects of such networks have been studied in the static case, i.e. in which the agents are randomly placed but fixed [6, 53]. Algorithms for computing connectivity properties of such a network have been studied [59], and simulation results for randomly placed agents were reported there and in [43, 48, 47, 72]. Furthermore, several of the above-mentioned studies mention the dynamic situation, i.e. in which the agents are mobile.

Recently, there has been an increasing interest for MANETs (*mobile ad-hoc networks*). A MANET is a type of wireless ad-hoc network in which the agents are allowed to move freely around some prescribed environment. Several models of mobility have been proposed in the literature — for an excellent survey of those models we refer to [47]. In all these models, the connections in the network are created and destroyed as the vertices move closer together or further apart, and therefore this gives rise to a random-graph process. In all previous work, the authors performed empirical studies on network topology and routing performance. The paper [39] also deals with the problem of maintaining connectivity of mobile vertices communicating by radio, but from an orthogonal perspective to the one in the present dissertation — it describes a *kinetic data structure* to maintain the connected components of the union of unit-radius disks moving in the plane.

The particular mobility model used in this dissertation is often called the *random walk* model in the literature. It was introduced by Guerin [38], and it can be seen as the foundation for most of the mobility models developed afterwards [47]. In the random walk model, each vertex selects uniformly at random a direction (angle) in which to travel. The vertices select their velocities from a given distribution of velocities, and then each vertex moves in its selected direction at its selected velocity. After some randomly chosen period of time, each vertex halts, selects a new direction and velocity, and the process repeats. An experimental study of the connectivity of the resulting ad-hoc network for this particular model is presented in [73]. As is stated in the same paper, in many applications which are not life-critical, connectivity is an important parameter: *"Temporary network disconnections can be tolerated, especially if this goes along with a significant decrease of energy consumption."* In [62] the authors use a similar model to the one used in the present paper to prove that if the vertices are initially distributed uniformly at random, the distribution remains uniform at any time.

To the best of our knowledge, this work presents the first analytical study on the maintenance of connectivity of a MANET under the random walk model. As a first approximation to the problem, in Section 2.3 we introduce the *walkers* model, in which the agents, denoted *walkers*, move through a fixed environment modelled by a motion graph. The particular topologies of the environment covered there are the cycle and the toroidal grid. On the other hand in Chapter 3, we analyse a setting in which the agents move freely around the unit torus $[0, 1)^2$. This can be regarded as a continuous counterpart of the walkers model on the toroidal grid, and indeed the two models bear some resemblance. However, some aspects of the technical developments in each one are quite different, and each model has mathematical interest on its own.

2.2 Brief Survey on Random Geometric Graphs and Related Models

Recall from last section the model of random graph proposed by Gilbert in 1961 [35], in which the vertices are selected from an infinite random-point process over the real plane, and there is an edge between any two vertices at distance less than some parameter. A variant of this model in which the number of vertices is finite and the point process is taken over a bounded region $\mathcal{S}$ has also been widely studied (see, e.g., [40, 64]). Usually, the point process is assumed to be uniform over $\mathcal{S}$, and $\mathcal{S}$ is taken to be the unit disk, the unit square or their higher dimensional counterparts. Also, in order to avoid boundary effects, it is also useful to consider $\mathcal{S}$ to be the d-torus which results from identifying the opposite faces of the unit d-cube $[0, 1]^d$. Moreover, distances in $\mathcal{S}$ are usually measured with respect to the Euclidean distance, but the model can be extended without much effort to any arbitrary ℓ_p-normed distance, with $1 \leq p \leq \infty$. In general, using a different distance just changes a constant multiplicative factor in all the expressions. All of these variations from the original model are also usually referred to as *random geometric graphs*. Figure 2.1 illustrates an instance of a random geometric graph of 125 vertices over the unit square.

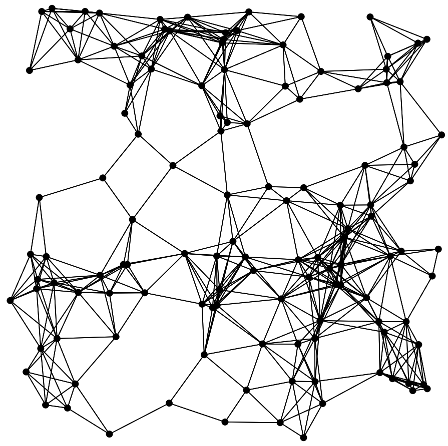


Figure 2.1: Random geometric graph of size 125 and radius 0.155 (by J. Petit)

Rigorously, given a metric space $\mathcal{S}$, a natural number n and a positive real r, a random geometric graph over $\mathcal{S}$ of size n and radius r is constructed by selecting independently and u.a.r. n points $X_1, \ldots, X_n \in \mathcal{S}$ and joining by an edge any two of them at distance at most r. We denote the random set of points by $\mathcal{X} = \mathcal{X}(n) = \bigcup_{i=1}^{n}\{X_i\}$, and call $G(\mathcal{X}; r)$ to the resulting random geometric graph. In Section 2.3 of the present chapter, we deal with random graphs over a discrete space $\mathcal{S}$, which consists in the set of vertices of a connected graph (the cycle and the toroidal grid). For this model, several distances are considered. On the other hand, the random geometric graphs treated in Chapters 3 and 4, are respectively over the unit torus $[0, 1)^2$ under the Euclidean distance, and over the unit square $[0, 1]^2$under any ℓ_p-normed distance.

Now we survey the main known results on random geometric graphs, in particular the ones which have some relevance to the work described in this dissertation. We often denote by $G(\mathcal{X}; r)$ a random geometric graph of size n and radius r without specifying the precise model considered. In addition, r is regarded as a function of the number of vertices n and, unless otherwise stated, all the presented results are asymptotic as n tends to infinity. For an exhaustive exposition on random geometric graphs in full generality, see the monograph by Penrose [66].

One of the most important topics when studying a family of random graphs is their connectivity. We wonder for which values of the parameter r of a random geometric graph $G(\mathcal{X};r)$ we can guarantee that it is a.a.s. connected, and more accurately how the connectivity of $G(\mathcal{X};r)$ evolves as r grows. This property turns out to undergo a phase transition for some threshold value r_c. Namely, for values of r below r_c (sub-critical case) $G(\mathcal{X};r)$ is a.a.s. disconnected, and for values of r above r_c (super-critical case) it is a.a.s. connected. Several authors independently deduced the connectivity threshold value for related variants of the random geometric graph model. Penrose [64] obtained r_c for random geometric graphs over the unit square $[0,1]^2$ and the d-torus $[0,1)^d$ under the Euclidean norm. Gupta and Kumar [40] obtained similar results for the unit disk. Also, Appel and Russo [8] computed r_c for the unit square under the ℓ_∞ distance. Finally, Penrose [65] gave a stronger result, which in particular yields the threshold for k-connectivity in the unit d-cube $[0,1]^d$ under any ℓ_p normed distance ($1 \leq p \leq \infty$), and in Chapter 13 of [66] covered this topic in full generality. In all these cases, isolated vertices turn out to play a major role in connectivity. In fact, r_c coincides with the value of the parameter r which makes the expected number of isolated vertices be $\Theta(1)$. This corresponds to $r_c = \sqrt{(\log n \pm O(1))/(\alpha_p n)}$, where α_p is the area of the unit disk in the ℓ_p norm, in the case that we are considering random geometric graphs over the unit square or the unit torus with the ℓ_p distance (see Lemma 3.2.1 and Theorem 3.2.2 for more details in the Euclidean case).

Another issue extensively studied in Chapter 3 of [66] is the number of components in $G(\mathcal{X};r)$ isomorphic to a fixed graph, and thus the probability of finding components in $G(\mathcal{X};r)$ of a given size. However the range of r covered there does not exceed $\Theta(\sqrt{1/n})$, below the connectivity threshold r_c. In fact, a percolation argument in [66] shows that with probability $1-o(1)$ no components other than isolated vertices and the giant one exist at the connectivity threshold, without giving accurate bounds on this probability. In Chapter 3 we compute for $r = r_c$ the probability of having components other than isolated vertices according to their size.

Now recall that a graph property Π is said to be *monotone* if it is preserved when edges are added to the graph. For example, connectivity and hamiltonicity are monotone graph properties. Given a non-trivial monotone property Π, let $r_\Pi(n,p)$ be the radius at which a random geometric graph has property Π with probability exactly p. For $0 < p < 1/2$ define the *threshold width* $\delta_\Pi(n,p)$ of Π by

$$\delta_\Pi(n,p) = r_\Pi(n,1-p) - r_\Pi(n,p).$$

A property Π has a *sharp threshold* if $\delta_\Pi(n,p) = o(1)$. Goel, Rai and Krishnamachari proved that for random geometric graphs, any non-trivial monotone property has a sharp threshold [36].

Other important graph invariants are the *clique number* ω and the *chromatic number* χ. In applications, the chromatic number of a random geometric graph $G(\mathcal{X};r)$ gives an indication of how many different channels we have to use in an ad-hoc network to avoid interference. Elementary computations show that, when the radius $r = r_c$, the expected number of neighbours of a vertex in $G(\mathcal{X};r_c)$ is $\Theta(\log n)$. Moreover, it is not hard to prove that the clique number $\omega\big(G(\mathcal{X};r_c)\big)$ and the chromatic number $\chi\big(G(\mathcal{X};r_c)\big)$ are a.a.s. also $\Theta(\log n)$ (see, e.g., Chapter 6 in [66]). McDiarmid [58] computed the asymptotic ratio between these two invariants. In fact, he showed that in the sub-critical case,

$$\frac{\chi\big(G(\mathcal{X};r)\big)}{\omega\big(G(\mathcal{X};r)\big)} \to 1 \qquad \text{a.a.s.},$$

while in the super-critical case,

$$\frac{\chi\big(G(\mathcal{X};r)\big)}{\omega\big(G(\mathcal{X};r)\big)} \to 1.103 \qquad \text{a.a.s.}$$

The cover time of a graph G is the expected time taken by a simple random walk on G to visit all vertices in G. Avin and Ercal [10] showed for some critical parameter $r_{\text{cov}} = \Theta(r_c)$ that if $r > r_{\text{cov}}$ then $G(\mathcal{X};r)$ has a.a.s. optimal cover time of $\Theta(n \log n)$. Note that if r is below r_c then $G(\mathcal{X};r)$ is a.a.s. disconnected and it has infinite cover time.

Another natural issue to study is the existence of Hamiltonian cycles in $G(\mathcal{X};r)$. Penrose in Chapter 13 of his book [66] poses it as an open problem whether exactly at the point where $G(\mathcal{X};r)$ gets 2-connected, the graph a.a.s. also becomes Hamiltonian. Petit in [68] proved that for $r = \omega(\sqrt{\log n/n})$, $G(\mathcal{X};r)$ is Hamiltonian a.a.s. and he also gave a distributed algorithm to find a Hamiltonian cycle in $G(\mathcal{X};r)$ with his choice of radius. In Chapter 4 of the present dissertation, we find the sharp threshold for this property for any ℓ_p metric.

There have been other types of random graphs also used to model communication networks, and that sometimes tend to be mistaken with random geometric graphs. The *Random Euclidean graphs* are constructed as follows: choose a sequence $\mathcal{X} = (X_1, \dots, X_n)$ of independently and uniformly distributed (i.u.d.) points in $[0,1]^d$, and consider the weighted complete graph on $\mathcal{X}$, where the weight of an edge is its Euclidean length. A lot of theoretical work has been done on this model (see, e.g., the books of Steel [76] and Yukich [82]). Probably, the most celebrated result in this area is the Beardwood-Halton-Hammersley Theorem [12]: Let $\mathcal{X}$ a sequence of i.u.d. points in $[0,1]^d$, and let L_n be the optimal solution of the Travelling Salesman Problem (TSP) on the Euclidean graph defined on $\mathcal{X}$. Then there exists $\beta(d)$, $0 < \beta(d) < \infty$, such that

$$\frac{L_n}{n^{(d-1)/d}} \to \beta(d) \qquad \text{a.a.s.,}$$

where later on, it was experimentally obtained that $0.70 \le \beta(2) \le 0.73$.

Another related model of random graph is the *random proximity graph*: Given n labelled nodes distributed in $\mathbb{R}^2$ according to a Poisson point process, and given a fixed $\phi(n) \in \mathbb{Z}^+$, let $\mathcal{G}(n,\phi(n))$ be the graph formed when each node is connected with its $\phi(n)$ nearest neighbours [80]. The main result in this area is the following characterisation of the connectivity of $\mathcal{G}(n,\phi(n))$ by Xue and Kumar in the previous paper: If $\phi(n) \le 0.0074 \log n$ then $\mathcal{G}(n,\phi(n))$ is a.a.s. disconnected; and if $\phi(n) > 5.\log n$ then $\mathcal{G}(n,\phi(n))$ is a.a.s. connected. See [28] for further comparisons between these models and random geometric graphs.

Finally one world about one of the main techniques used in this dissertation when dealing with random geometric graphs over the unit square with radius r around the connectivity threshold. The *dissection technique* consists of first tessellating the unit square into $\Theta(1/r^2)$ small square cells, such that each contains around $(\Theta(\log n))$ vertices a.a.s., then solving the problem on each small cell and finally compounding the solution (or an approximation) of the whole problem. Karp used it, together with dynamic programming, to give an approximated solution to the Travelling Salesman Problem on Euclidean graphs [50]. Since then, variants of the dissection technique have been widely used extensively in the realm of random Euclidean graphs and random geometric graphs.

2.3 The Walkers Model

Consider a setting in which a large number of mobile agents can perform concurrent basic movements: ahead/behind/left/right, determining a grid pattern, or left/right, describing a line. Each agent can communicate directly with any other agents which are within a given distance d. This enables communications with agents at a further distance using several intermediate agents. At each step in time there is an ad-hoc network defined by the dynamic graph whose vertex set consists of the agents, with an edge between any two agents iff they are within the distance d of each other.

We propose what we call the *walkers model*, defined as follows: A connected graph $M_N = (V, E)$ with $|V| = N$, a number w of *walkers* (agents) and a "distance" d are given. We call M_N the *motion graph.* A set $W = \{1, \dots, w\}$ of walkers are placed independently and uniformly at random on the vertices of the motion graph M_N, allowing several walkers to lie on the same vertex. Each walker has a range d for communication; that is, two walkers can communicate in one hop if the distance, in M_N, between the position of the walkers is at most d. Two walkers can communicate if they can reach each other by a sequence of such hops. In addition, each walker takes an independent standard random walk on M_N, i.e. moves at each time step to a neighbouring vertex, which is chosen with equal probability.

Let $\mathcal{V} = \mathcal{V}(M_N, w)$ denote the tuple $(v_1, \dots, v_w) \in V^w$ of positions of the walkers on M_N. The interesting features of the walkers model are encapsulated by the *random graph of walkers*, $G\big(\mathcal{V}(M_N, w); d\big)$. The vertices of $G\big(\mathcal{V}(M_N, w); d\big)$ are the vertices in M_N that receive at least one walker, i.e. those ones which appear at least once in $\mathcal{V}$. Two vertices in $G\big(\mathcal{V}(M_N, w); d\big)$ are joined by an edge iff they are at distance at most d in M_N. Notice that this ad-hoc graph generated by the placement of the walkers on M_N is a very particular type of random geometric graph. We are interested in the probability of $G\big(\mathcal{V}(M_N, w); d\big)$ being connected, or in the number of components and their sizes, with certain mild asymptotic restrictions on w and d. We generally abbreviate $G\big(\mathcal{V}(M_N, w); d\big)$ to $G(\mathcal{V})$ when M_N, w and d are understood.

Our primary goal with the walkers model is to characterise the dynamics of the connectivity of the communication network as represented by the random graph process $\big(G(\mathcal{V}_t)\big)_{t\in\mathbb{Z}}$. Here $\mathcal{V}_t$ denotes the vector of positions of the walkers at time t, and $G(\mathcal{V}_t)$ is the corresponding graph of walkers at that time, constructed from $\mathcal{V}_t$. Formally, $\mathcal{V}_t$ is obtained as follows: Start from an initial configuration of walkers $\mathcal{V}_0$ in which each walker chooses independently and u.a.r. one vertex of M_N. Then $\mathcal{V}_{t+1}$ is obtained from $\mathcal{V}_t$ by making each walker move simultaneously one step to a randomly selected neighbour vertex in M_N. The process is time-reversible and $\mathcal{V}_t$ can be obtained from $\mathcal{V}_{t+1}$ in the same way. This allows us to consider also negative times, by looking at the process backwards. In order to study $\big(G(\mathcal{V}_t)\big)_{t\in\mathbb{Z}}$, we first examine $G(\mathcal{V})$, which we call the *static* model. This is a snapshot of the process at one point in time: we are interested in the distribution of the number of components, as well as some other information which helps to answer the dynamic questions. In particular, we are interested in studying the birth and death of components, and the sudden connection and disconnection of $\big(G(\mathcal{V}_t)\big)_{t\in\mathbb{Z}}$.

We consider a sequence of graphs M_N with increasing numbers of vertices N, for N tending to infinity. The parameters w and d are functions of N, and unless otherwise indicated all asymptotic notation is with respect to $N \to \infty$. We restrict to the case $w \to \infty$ in order to avoid considering small-case effects. Of course we take $d \geq 1$. We make further restrictions on w and d in order to rule out non-interesting cases, such as values of the parameters in which the network is a.a.s. disconnected or a.a.s. connected. In this chapter, we study the walkers model for two particular sequences of graphs M_N: the cycle C_N of length N and the $n \times n$ toroidal grid T_N of size $N = n^2$. In the case of the grid, we use the ℓ^p distance, for any $1 \le p \le \infty$. The two cases have an essential difference that prevents a unified treatment: for the interesting values of w and d, disconnectedness of the graph of walkers for the cycle is basically caused by the presence of at least two large “gaps” between the walkers around the cycle, whereas for the grid, it is caused by the presence of one or more isolated walkers.

We now give a brief outline of some of our main results. First we consider the cycle C_N on N vertices. To study connectedness in this case, we introduce the concept of a *d-hole*, which informally is a sequence of at least d consecutive vertices containing no walkers between two vertices containing some walker. A parameter μ closely related to the expected number of d-holes is introduced, and the connectedness of $G(\mathcal{V})$ is characterised in terms of μ (see Theorem 2.5.2 and Corollary 2.5.3). In fact $\mu = \Theta(1)$ is a threshold for this property, and when this condition holds we give the asymptotic distribution of the number of components of $G(\mathcal{V})$. In the dynamic setting we study the creation, evolution and destruction of these d-holes, and in particular obtain the expected time that a given d-hole will live from the moment it is created (see Theorems 2.5.8 and 2.5.9). Note that we restrict the attention to the case $\mu = \Theta(1)$, since we wish to study only the non-trivial dynamic situations. One of our main results concerns the expected time that the graph of walkers remains (dis)connected, after a point in time at which it becomes (dis)connected (see Theorem 2.5.12). It must be noted that this quantity is affected by the particular initial state $\mathcal{V}_0$. This motivates an alternative notion of the average time that $\big(G(\mathcal{V}_t)\big)_{t\in\mathbb{Z}}$ is (dis)connected. See the discussion in Subsection 2.5.2 and Theorem 2.5.13 for further details.

We turn now to the toroidal grid T_N with $N = n^2$ vertices, for which our results apply with any normed ℓ^p distance, for $1 \le p \le \infty$. A role similar to that of the d-holes in the cycle case is played here by the isolated vertices of $G(\mathcal{V})$, also called *simple* components. In fact, the connectedness of $G(\mathcal{V})$ is determined in terms of μ, where μ is redefined to an expression closely related to the expected number of simple components (see Theorem 2.6.6 and Corollary 2.6.7). Moreover $\mu = \Theta(1)$ is a threshold for this property, and when this condition holds a.a.s. $G(\mathcal{V})$ consists only of a “giant” component and a Poisson number of simple components. In the dynamic setting we study properties analogous to those covered for the cycle. In particular, we obtain asymptotic expressions for the expected lifespan of a simple component (see Theorem 2.6.13) and the expected time that $\big(G(\mathcal{V}_t)\big)_{t\in\mathbb{Z}}$ remains (dis)connected from the moment it becomes (dis)connected (see Theorem 2.6.14).

The remainder of the chapter gives proofs of these theorems as well as stating and proving related ones. In Section 2.4 we give basic definitions and technical lemmas to be used throughout the chapter. In Section 2.5 we deal with the cycle C_N, and in Section 2.6, the toroidal grid T_N. One of the main differences between this case and the cycle is the need for a geometric lemma (Lemma 2.6.3) that may be of independent interest. This bounds the size of the set of non-occupied vertices at distance at most d from the boundary of any connected component in $G(\mathcal{V})$. The last section contains some discussion and related problems.

2.4 General Definitions and Basic Results

We begin with some definitions and results which are common for all graphs M_N. Let K denote the number of components in $G(\mathcal{V})$. We call a component *simple* if it consists of only one isolated vertex. We define the ratio $\varrho = w/N$, which is the expected number of walkers at any given vertex. Most of the statements in the chapter are in terms of ϱ rather than w. For any $v \in V(M_N)$, define h_v to be the number of vertices in $V(M_N)$ different from v and at distance at most d from v, and define $h = \min_{v \in V} h_v$. We say that a vertex or set of vertices is *empty of walkers* (*e.o.w.*), or simply *empty*, if it contains no walkers, and *occupied* otherwise. Note that there must be h empty vertices in $V(M_N)$ within distance d of a simple component.

By considering the coupon collector's problem, we observe that if $w = N \log N + \omega(N)$ then $G(\mathcal{V})$ is trivially a.a.s. connected due to every vertex being occupied. Moreover, for the graphs M_N which we consider in this chapter, if $h \in \Omega(N/\sqrt{w})$ then $G(\mathcal{V})$ is a.a.s. connected as well. This last claim will be seen in Observations 2.5.1 and 2.6.1. Thus, we consider throughout the chapter $w \leq N \log N + O(N)$ and $h = o(N/\sqrt{w})$. In fact, our proofs will just assume h to be $o(N)$. Note that, for the cycle, $h = 2d$.

Often in this chapter we will need to compute the probability of certain configurations of walkers involving two consecutive time steps t and $t+1$, in order to record the event that walkers jump to the appropriate place at that step. There is a convenient way to formulate this in terms of occupancy of *arcs* (directed edges). Let us regard M_N as a directed graph, by considering each edge as a pair of anti-parallel arcs, and let us denote by $A = A(M_N)$ the set of arcs of M_N. For any arc $e = (u, v) \in A(M_N)$, we say that a walker is placed on e between time steps t and $t+1$ if the walker is on u at time step t and jumps onto v within one step. This determines an arrangement of the set of walkers on the arcs of M_N in which the position of each walker is chosen independently and with probability inversely proportional to the degree of the origin of the arc. This way we can encode dynamic transitions between t and $t + 1$ in terms of static configurations of walkers over $A(M_N)$. Often t and $t + 1$ are not explicitly mentioned when they are understood from the context.

There is an alternative formulation in terms of cells. Each vertex is divided into as many cells as its degree, and each cell is associated with one of arcs stemming from the vertex. Then, the transition of the system between two consecutive time steps can be described by the placement of the walkers in the cells (see Figure 2.2). We use this representation mainly in figures for sake of simplicity and visual clarity.

Assign *size* 1 to all vertices in $V(M_N)$. For a given arc stemming from a vertex v with degree δ_v, its *size* will be $1/\delta_v$. Given a set $\mathcal{S}$ of vertices or arcs, we define $\mathsf{Size}(\mathcal{S})$ to be the sum of the sizes of its elements. Observe that the probability that one walker lies in $\mathcal{S}$ is $\mathsf{Size}(\mathcal{S})/N$.

Throughout the chapter, we are often interested in the probability of events which can be characterised by the fact that some subset of $V(M_N)$ or $A(M_N)$ is e.o.w. and some other ones are occupied. The following results provide asymptotic bounds on these probabilities, and are used in most of the computations, sometimes without an explicit mention. The first lemma is stated in a more general setting and is also used in Chapter 3.

Lemma 2.4.1. *Consider a setting with n balls and $k + 1$ disjoint bins $\mathcal{U}_0, \dots, \mathcal{U}_k$, where each ball is either placed into at most one of the bins or possibly remains outside all of them. Suppose that each ball is assigned to bin $\mathcal{U}_i$ with probability $p_i = p_i(n)$, independently from*

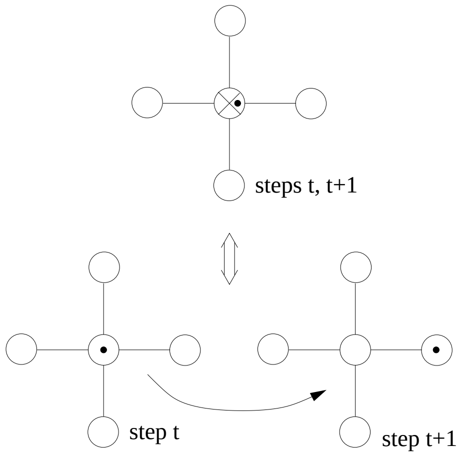


Figure 2.2: The walker jumps to a neighbour according to which cell it is placed on

the choices of the other balls. Moreover suppose that for all i ($0 \le i \le k$) *we have* $p_i = o(1)$, *where asymptotics are with respect to* $n \to \infty$ *and where k is assumed to be fixed. Then, the probability P that* $\mathcal{U}_0$ *contains no balls but for all* $i \in \{1, \ldots, k\}$ $\mathcal{U}_i$ *receives at least one is*

$$P \sim (1 - p_0)^n \prod_{i=1}^{k} (1 - e^{-np_i}).$$

Proof. By reordering the labels of the bins except for $\mathcal{U}_0$, assume that $np_i = o(1)$ if $1 \le i \le r$ and $np_i = \Omega(1)$ if $r + 1 \le i \le k$. For any expression $f = o(1)$, we define

$$P_f = \sum_{a_1,\ldots,a_r} (-1)^{\sum_{i=1}^r a_i} \left(1 - \frac{\sum_{i=1}^r a_i p_i}{1-f}\right)^n,$$

where the summation indices $a_1, \ldots, a_r$ run from 0 to 1. We can write

$$P_f = \sum_{a_1,\ldots,a_r} (-1)^{\sum_{i=1}^r a_i} \sum_{\substack{m_0,\ldots,m_r \ge 0 \\ m_0+\cdots+m_r=n}} \binom{n}{m_0, \ldots, m_r} \prod_{i=1}^{r} \left(\frac{-a_i p_i}{1-f}\right)^{m_i},$$

with the convention $0^0 = 1$. A changing of the order of summation converts this expression to

$$P_f = \sum_{\substack{m_0,\ldots,m_r \ge 0 \\ m_0+\cdots+m_r=n}} \binom{n}{m_0, \ldots, m_r} \prod_{i=1}^{r} \left(\frac{-p_i}{1-f}\right)^{m_i} \sum_{a_1,\ldots,a_r} (-1)^{\sum_{i=1}^r a_i} \prod_{i=1}^{r} a_i^{m_i}.$$

All terms in this sum with some $m_i = 0$ ($1 \le i \le r$) cancel, since each of these terms is equal but has opposite sign to the one obtained by switching the value of a_i. So, only the terms with all $m_1, \ldots, m_r \ge 1$ remain, and among these we can remove the ones with some $a_i = 0$. Hence,

$$P_f = \sum_{\substack{m_0 \ge 0 \\ m_1,\ldots,m_r \ge 1 \\ m_0+\cdots+m_r=n}} (-1)^r \binom{n}{m_0, \ldots, m_r} \prod_{i=1}^{r} \left(\frac{-p_i}{1-f}\right)^{m_i}.$$

Since $np_i = o(1)$, the main asymptotic weight in this sum corresponds to the term $m_0 = n-r$ and $m_1, \dots, m_r = 1$, so

$$P_f \sim \frac{[n]_r}{(1-f)^r} \prod_{i=1}^{r} p_i \sim \prod_{i=1}^{r} np_i \sim \prod_{i=1}^{r} (1-e^{-np_i}). \tag{2.1}$$

By an inclusion-exclusion argument, the probability in the statement can be written as

$$P = \sum_{a_1,\dots,a_k} (-1)^{\sum_{i=1}^k a_i} \left(1 - p_0 - \sum_{i=1}^{k} a_i p_i\right)^n,$$

where the summation indices $a_1, \dots, a_k$ run from 0 to 1. Then, if we define

$$P_{a_{r+1},\dots,a_k} = \sum_{a_1,\dots,a_r} (-1)^{\sum_{i=1}^r a_i} \left(1 - \frac{\sum_{i=1}^r a_i p_i}{1 - p_0 - \sum_{i=r+1}^k a_i p_i}\right)^n,$$

we can write

$$\begin{aligned} P &= (1-p_0)^n \sum_{a_{r+1},\dots,a_k} (-1)^{\sum_{i=r+1}^k a_i} \left(1 - \frac{\sum_{i=r+1}^k a_i p_i}{1-p_0}\right)^n P_{a_{r+1},\dots,a_k} \\ &= (1-p_0)^n \sum_{a_{r+1},\dots,a_k} (-1)^{\sum_{i=r+1}^k a_i} \exp\left(-(1+o(1)) \sum_{i=r+1}^{k} a_i np_i\right) P_{a_{r+1},\dots,a_k}. \end{aligned} \tag{2.2}$$

Note that for each $a_{r+1}, \dots, a_k \in \{0,1\}$, in view of (2.1) and setting $f = p_0 + \sum_{i=r+1}^k a_i p_i$, we have

$$P_{a_{r+1},\dots,a_k} \sim \prod_{i=1}^{r} (1-e^{-np_i}). \tag{2.3}$$

The fact that $np_i = \Omega(1)$ for $r+1 \le i \le k$ prevents the leading term of the sum in (2.2) from cancelling out. Thus, from (2.2) and (2.3), we obtain

$$P \sim (1-p_0)^n \prod_{i=1}^{k} (1-e^{-np_i}).$$

□

As an immediate consequence of this result, by regarding walkers as balls and sets of vertices (or arcs) in M_N as bins, we obtain

Lemma 2.4.2. *For any fixed integer $k \ge 0$, let $\mathcal{S}_0, \dots, \mathcal{S}_k$ be pairwise disjoint sets of vertices (or arcs) in M_N, with sizes $s_0, \dots, s_k$ respectively. If $\sum_{i=0}^k s_i = o(N)$, then*

$$\mathbf{P}\left(\mathcal{S}_0 \text{ is e.o.w. } \wedge \bigwedge_{i=1}^{k} (\mathcal{S}_i \text{ is occupied})\right) \sim \left(1 - \frac{s_0}{N}\right)^w \prod_{i=1}^{k} \left(1 - e^{-s_i \varrho}\right).$$

To cover large sizes s, not necessarily $o(N)$, and $k = k(N)$ not necessarily fixed, we need the following variation.

Lemma 2.4.3. *Let $\mathcal{S} \subset V(M_N)$ be a set of vertices of size s, and $v_1, \dots, v_k \in V(M_N)$ be vertices not in $\mathcal{S}$, with $1 \le k \le N$. Assume that $N - s \to \infty$. Then the probability that $\mathcal{S}$ is e.o.w. and $v_1, \dots, v_k$ are all occupied is at most $p_0 p^{k-1} \alpha^w (1 + o(1))$ where $p_0 = 1 - e^{-\varrho/\alpha}$, $\alpha = 1 - s/N$ and $p = \min(1, \varrho/\alpha)$. Here the asymptotics in the $(1+o(1))$ factor are uniform over all k.*

Proof. The probability of the event $\mathcal{E}$ that $\mathcal{S}$ is e.o.w. is α^w. The probability that v_1 is occupied conditional upon $\mathcal{E}$ is $1-(1-(N-s)^{-1})^w$, which is asymptotic to p_0. The lemma follows for $\varrho > \alpha$, i.e. $w > N - s$. Otherwise, conditional upon $\mathcal{E}$ and the event that $v_1, \dots, v_i$ are occupied, the probability that the next is occupied is clearly decreasing with i and is thus at most $w/(N-s) = \varrho/\alpha$. □

2.5 The Cycle

Here $M_N = C_N$, the cycle with N vertices.

Observation 2.5.1. Cover C_N with $\lceil N \lceil d/2 \rceil^{-1} \rceil$ paths of $\lceil d/2 \rceil$ vertices. If $d = \Omega(N/\sqrt{w})$, then the probability that some path is e.o.w. is at most

$$\left\lceil \frac{N}{\lceil d/2 \rceil} \right\rceil \left(1 - \frac{\lceil d/2 \rceil}{N}\right)^w \le O(\sqrt{w}) e^{-\Omega(\sqrt{w})} \to 0.$$

Thus, a.a.s. each of these paths is occupied (by at least one walker), and $G(\mathcal{V})$ is connected.

In view of this observation, we assume for the rest of the section that $d = o(N)$. If $d = \Omega(N)$, then $G(\mathcal{V})$ is a.a.s. connected.

To study connectedness of $G(\mathcal{V})$, we introduce the concept of a *hole*. There is a *hole* between two vertices u and v if u and v each contain at least one walker, but no vertex in the clockwise path from u to v contains a walker. We say that such a hole *follows* u, or that u is the *start vertex* of the hole. The number of internal vertices in a hole is its *exact size*. A k-*hole* is a hole whose exact size is at least k. Let H be the random variable counting the number of d-holes in the walkers model for C_N. Notice that at least two d-holes are needed to disconnect the walkers on C_N. To be precise,

$$G(\mathcal{V}) \text{ is connected iff } H \le 1 \quad \text{and} \quad K = \max\{1, H\}. \tag{2.4}$$

2.5.1 Static Properties

Here, we study the connectedness of the graph of walkers $G(\mathcal{V})$ in the static situation, by analysing the behaviour of H. In view of (2.4), if $\mathbf{E}H \to 0$ then $G(\mathcal{V})$ is a.a.s. connected.

We define a new parameter $\mu = N\,(1 - e^{-\varrho})\, e^{-d\varrho}$, which plays a key role in characterising the connectedness of $G(\mathcal{V})$. Notice that

$$\mu \sim \begin{cases} w e^{-d\varrho} & \text{if } \varrho = o(1), \\ N\,(1 - e^{-\varrho})\, e^{-d\varrho} & \text{if } \varrho = \Theta(1), \\ N e^{-d\varrho} & \text{if } \varrho = \omega(1). \end{cases}$$

Regarding the behaviour of H, and connectedness, we have the following.

Theorem 2.5.2. *The expected number of holes satisfies*

$$\mathbf{E}H \sim N\left(1 - e^{-\varrho}\right)(1 - d/N)^{w}.$$

Furthermore,

(i). if $\mu = o(1)$, then a.a.s. $G(\mathcal{V})$ is connected,

(ii). if $\mu = \omega(1)$, then a.a.s. $G(\mathcal{V})$ is disconnected,

(iii). if $\mu = \Theta(1)$, then H is asymptotically Poisson with mean μ.

Proof. For any vertex v in $V(C_N)$, let H_v be an indicator random variable such that $H_v = 1$ iff there is a d-hole following vertex v. Then,

$$H = \sum_{v \in V(C_N)} H_v \qquad \text{and} \qquad \mathbf{E}H = \sum_{v \in V(C_N)} \mathbf{P}(H_v = 1). \tag{2.5}$$

Let $\boldsymbol{\mathcal{T}}$ bet the set of all k-tuples $\boldsymbol{v} = (v_1, \ldots, v_k)$ of pairwise different vertices $v_1, \ldots, v_k \in V(C_N)$. We compute the k^{th} factorial moment of H:

$$\mathbf{E}[H]_k = \sum_{\boldsymbol{v} \in \boldsymbol{\mathcal{T}}} \mathbf{P}\big((H_{v_1} = 1) \wedge \cdots \wedge (H_{v_k} = 1)\big). \tag{2.6}$$

Let $\boldsymbol{\mathcal{T}}_1$ denote the set of tuples $\boldsymbol{v} \in \boldsymbol{\mathcal{T}}$ such that each v_i and v_j, $i \neq j$, have distance at least $d+1$ around the cycle. For $\boldsymbol{v} \in \boldsymbol{\mathcal{T}} \setminus \boldsymbol{\mathcal{T}}_1$, the probability in (2.6) is 0 since one of the v_i "lies in" the hole following some v_j, and yet v_i must be occupied. For $\boldsymbol{v} \in \boldsymbol{\mathcal{T}}_1$, the probability in (2.6) is easily computed by applying Lemma 2.4.2:

$$\mathbf{P}\big((H_{v_1} = 1) \wedge \cdots \wedge (H_{v_k} = 1)\big) \sim \left(1 - \frac{kd}{N}\right)^{w} \left(1 - e^{-\varrho}\right)^{k}.$$

Since $d = o(N)$ we have $|\boldsymbol{\mathcal{T}}_1| \sim N^k$, and thus from (2.6)

$$\mathbf{E}[H]_k \sim \left(N\left(1 - e^{-\varrho}\right) e^{-d\varrho - O(d^2 w/N^2)}\right)^{k}. \tag{2.7}$$

In particular,

$$\begin{aligned}\mathbf{E}H &\sim N\left(1 - e^{-\varrho}\right)\left(1 - \frac{d}{N}\right)^{w} \\ &\sim N\left(1 - e^{-\varrho}\right) e^{-d\varrho - O(d^2 w/N^2)}.\end{aligned} \tag{2.8}$$

In the case $\mu \to 0$, we have also $\mathbf{E}H \to 0$, since $(1 - d/N)^w \leq e^{-d\varrho}$. Then, $\mathbf{P}(H = 0) \to 1$, and $G(\mathcal{V})$ is connected a.a.s. In the case that μ is bounded away from 0, taking logarithms,

$$d\varrho = \log N\left(1 - e^{-\varrho}\right) - \log \mu. \tag{2.9}$$

Considering separately the cases when $\varrho \to 0$ and when $\varrho = \Omega(1)$, we get from (2.9) that $d^2 w/N^2 = o(1)$, so we can ignore the term $O(d^2 w/N^2)$ in the expression (2.7) and obtain

$$\mathbf{E}[H]_k \sim \left(N\left(1 - e^{-\varrho}\right) e^{-d\varrho}\right)^{k} = \mu^k. \tag{2.10}$$

Moreover, if μ is bounded, then it follows, from (2.10) and from Theorem 1.23 in [15], that H is asymptotically Poisson. For $\mu \to \infty$, we have that $\mathbf{E}[H]_2 \sim \mu^2$, and so it follows from Chebyshev's inequality that a.a.s. $H > \mu/2$, so $G(\mathcal{V})$ is disconnected. □

The following corollary gives the asymptotic probability that $G(\mathcal{V})$ is connected. It follows immediately from the theorem in view of (2.4).

Corollary 2.5.3. $\mathbf{P}\big(G(\mathcal{V})$ *is connected*$\big) = e^{-\mu}(1+\mu)+o(1)$.

2.5.2 Dynamic Properties

Assume that from an initial random placement of the walkers, at each step, every walker moves from its current position to one of its neighbours, with probability 1/2 of going either way. This is a standard random walk on the cycle for each walker. To study the connectivity properties of the dynamic graph of walkers we need to introduce some notation.

A *configuration* or *state* is an arrangement of the w walkers on the vertices of C_N. If we regard the vertices of C_N as elements of $\mathbb{Z}_N$, then each configuration can be represented by a vector $\mathcal{V} = (v_1, \ldots, v_w) \in (\mathbb{Z}_N)^w$, where v_i indicates the vertex being occupied by walker i. Consider the graph of configurations, where the vertices are the N^w different configurations. Given a configuration $\mathcal{V}$, there exists an edge between $\mathcal{V}$ and all configurations $(v_1 \pm 1, \ldots, v_w \pm 1)$. Thus, any configuration has 2^w neighbours, and the relationship of being neighbours is symmetric. The dynamic process can be viewed as a random walk on the graph of configurations, in particular, a Markov chain $\mathscr{M} = (\mathcal{V}_t)_{t\in\mathbb{Z}}$, where $\mathcal{V}_t$ denotes the state of the process at time step t.

For N even, given any two configurations $\mathcal{U}$ and $\mathcal{V}$, we say that they have the *same parity* if for all i and j, $u_i - u_j \equiv v_i - v_j \mod 2$. There are 2^{w-1} different parities. Note that the initial parity stays invariant during the dynamic process. The following lemma is straightforward and the proof is left to the reader.

Lemma 2.5.4. *Let $\mathcal{U}$ and $\mathcal{V}$ be any two configurations and let $h_{\mathcal{U},\mathcal{V}}$ denote the hitting time from $\mathcal{U}$ to $\mathcal{V}$. If N is odd, then $\mathcal{V}$ is reachable from $\mathcal{U}$ and $h_{\mathcal{U},\mathcal{V}}$ is finite for any $\mathcal{U}$ and $\mathcal{V}$. If N is even then $\mathcal{V}$ is reachable from $\mathcal{U}$ provided that $\mathcal{U}$ and $\mathcal{V}$ have the same parity, and in this case $h_{\mathcal{U},\mathcal{V}}$ is finite.*

Proof. Suppose first that N is odd. For every walker i, there are two paths connecting vertices u_i and v_i, whose lengths are different $\mod 2$. Choose that path with even length (for instance), and call this length d_i. Let us suppose that $d_m \geq d_i$ for all $i \in \{1, \ldots, w\}$. We can get from state $\mathcal{U}$ to state $\mathcal{V}$ within d_m steps in the following way: for each walker i at position u_i move counter-clockwise for $\frac{d_m - d_i}{2}$ steps, and return back to the original position in $\frac{d_m - d_i}{2}$ more steps. Then, get to the final position v_i through the appropriate path within d_i steps.

Otherwise if N is even, the parity of the initial state stays invariant. Thus, we cannot reach one state from another one with different parities. If $\mathcal{U}$ and $\mathcal{V}$ have the same parity, we proceed in a similar way to that in the case N is odd to prove that they are mutually reachable. □

Then if N is odd, $\mathscr{M}$ consists on a single closed class of states, so it is irreducible and positive recurrent. It is trivial to verify aperiodicity and thus the chain is ergodic. However if N is even, there are 2^{w-1} closed classes of states, where each class consists of all configurations with the same parity. Let $\mathfrak{V}$ be any class of states and let $\mathcal{V} \in \mathfrak{V}$ be a configuration. For this particular configuration, we can partition the set of walkers $W = W_1 \cup W_2$ so that the ones in W_1 lie in odd positions of the cycle and the ones in W_2

lie in even positions. Let $\mathfrak{V}_1$ be the set of all states which lead to this same partition, and $\mathfrak{V}_2$ the set of those which lead to the complementary one. Notice that $\mathfrak{V} = \mathfrak{V}_1 \cup \mathfrak{V}_2$. Those states in $\mathfrak{V}_1$ are only reachable by an even number of steps from $\mathcal{V}$, and those in $\mathfrak{V}_2$ by an odd number of steps. Hence, if we restrict the Markov chain $\mathscr{M}$ to any class of states, it is irreducible, positive recurrent, but 2-periodic and hence not ergodic.

Observation 2.5.5. Note that for any fixed t, the distribution of $\mathcal{V}_t$ is just that of $\mathcal{V}$ in the static case. That is, the initial uniform distribution stays invariant, even though when N is even the chain is not ergodic and there is no limit distribution. Hence, by Theorem 2.5.2, if $\mu \to 0$ or ∞, then for any fixed t, $G(\mathcal{V}_t)$ is a.a.s. connected or a.a.s. disconnected, respectively.

In view of this observation, we assume $\mu = \Theta(1)$ for the remaining of the subsection, since we wish to study only the non-trivial dynamic situations. Under this assumption, from the proof of Theorem 2.5.2 and also recalling the restrictions on w, we have

$$d\varrho \sim \log w \to \infty, \qquad \text{and also} \qquad (1 - d/N)^w \sim e^{-d\varrho}. \tag{2.11}$$

We define H_t to be the random variable that counts the number of d-holes at time step t. Then from Subsection 2.5.1, H_t is asymptotically Poisson with expectation $\mu = \Theta(1)$. For the dynamic properties of $G(\mathcal{V}_t)$, we are interested in the probability that a new d-hole appears at a given time step. Moreover, we require knowledge of this probability conditional upon the number of d-holes already existing. If there is a d-hole from u to v at time step t and all walkers at u and v move in the same direction on the next step, a new d-hole may appear following one of the neighbours of u (provided no new walkers move in to destroy this). These two d-holes, though being different, are related, and the presence of the first makes the second more likely to occur than it would otherwise be. Similarly, the exact size of a d-hole following u may change in one step, making it technically a different d-hole, but again, related. In all these cases, the start vertex of the d-hole "moves" by at most 1; a d-hole at time step $t+1$ which does not follow u or a neighbour of u, is not related to a d-hole following u at time step t. To make this loose description rigorous, we need some definitions. Define a *d-hole line* to be a maximal sequence of pairs $(h_1, t_1), \ldots, (h_l, t_l)$ where h_i is a d-hole existing at time step t_i for $1 \le i \le l$, and such that $t_i = t_{i-1} + 1$ and the start vertex of h_i is adjacent to, or equal to, the start vertex of h_{i-1}, for $2 \le i \le l$. Fix two consecutive time steps t and $t+1$. If $t_1 = t+1$, we say that the line is *born* between t and $t+1$, if $t_l = t$ the line *dies* between t and $t+1$, and if $t = t_i$, $i \in \{1, \ldots, l-1\}$ we say that the line *survives* between t and $t+1$. Note that the time-reversal of the process has a d-hole line born at vertex u between $t+1$ and t iff the d-hole line dies at u between t and $t+1$.

We define random variables S_t, B_t and D_t to be the number of d-hole lines surviving, being born and dying, respectively, between t and $t+1$. Observe that

$$D_t + S_t = H_t \qquad \text{and} \qquad B_t + S_t = H_{t+1}. \tag{2.12}$$

We often omit t from the notation when it is understood, and simply write B, D and S. The proof of the following result is similar to that of Theorem 2.5.2, but rather more complicated.

Proposition 2.5.6. *For any fixed t, the random variables S_t, B_t and D_t are asymptotically jointly independent Poisson, with the expectations*

$$\mathbf{E}S_t \sim \begin{cases} \mu & \text{if } \varrho = o(1), \\ \mu - \lambda & \text{if } \varrho = \Theta(1), \\ 3\mu e^{-\varrho} & \text{if } \varrho = \omega(1), \end{cases} \quad \text{and} \quad \mathbf{E}B_t = \mathbf{E}D_t \sim \begin{cases} \frac{1}{2}\mu\varrho & \text{if } \varrho = o(1), \\ \lambda & \text{if } \varrho = \Theta(1), \\ \mu & \text{if } \varrho = \omega(1), \end{cases}$$

where $\lambda = \left(1 - \dfrac{3e^{-\varrho} - e^{-3\varrho/2}}{1 + e^{-\varrho/2}}\right)\mu$. *Here* $0 < \lambda < \mu$ *for* $\varrho = \Theta(1)$.

Proof. In C_N, let right denote 'clockwise' and left 'counter-clockwise'. Moreover, for a vertex $v \in V(C_N)$ and $i \geq 0$, let $v + i$ (respectively $v - i$) denote the vertex i positions to the right (resp. left) from v. All probabilities and events in this proof will involve two consecutive time steps t and $t+1$. We can describe these events from a static point of view, in terms of walkers' occupancy of certain *regions* (sets of arcs or vertices), as explained in Section 2.4. A summary of the sizes of the regions involved in these descriptions is given in Table 2.1.

There are three ways that a d-hole line can survive at vertex v during the interval of time $t, t+1$ according to the following descriptions (see also Figure 2.3):

s1 At time step t, there are no walkers between $v + 1$ and $v + d$. At least one walker at v moves right and all walkers at $v + d + 1$ and $v + d + 2$ (if there are any) move right.

s2 At time step t, there are no walkers between $v + 1$ and $v + d$. The walkers at v all move left and no walkers at $v - 1$ move right.

s3 At time step t, there are no walkers between $v + 1$ and $v + d$. The walkers at v all move left, at least one walker at $v - 1$ moves right and no walkers at $v + d + 1$ move left.

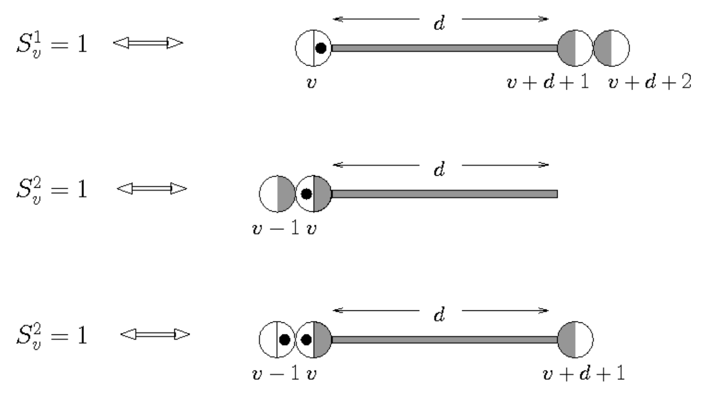


Figure 2.3: Survival of a d-hole line at vertex v

Similarly, there are four ways that a d-hole line can be born at v between time steps t and $t + 1$ according to the following descriptions (see also Figure 2.4):

b1 At time step t, there is a hole between $v + 1$ and $v + d$ of exact size $d - 2$. Then all walkers at v and $v + 1$ move left and all walkers at $v + d$ and $v + d + 1$ move right.

b2 At time step t, there is a hole between $v+1$ and $v+d+1$ of exact size $d-1$. Then all walkers at v and $v+1$ move left and all walkers at $v+d+1$ and $v+d+2$ move right.

b3 At time step t there is a hole between $v+1$ and $v+d+1$ of exact size $d-1$, and $v+d+2$ is occupied. Then all walkers at v and $v+1$ move left, all walkers at $v+d+1$ move right, and at least one walker at $v+d+2$ moves left.

b4 At time step t there is a hole between v and $v+d$ of exact size $d-1$, and $v-1$ is occupied. Then all walkers at v move left, all walkers at $v+d$ and $v+d+1$ move right, and at least one walker at $v-1$ moves right.

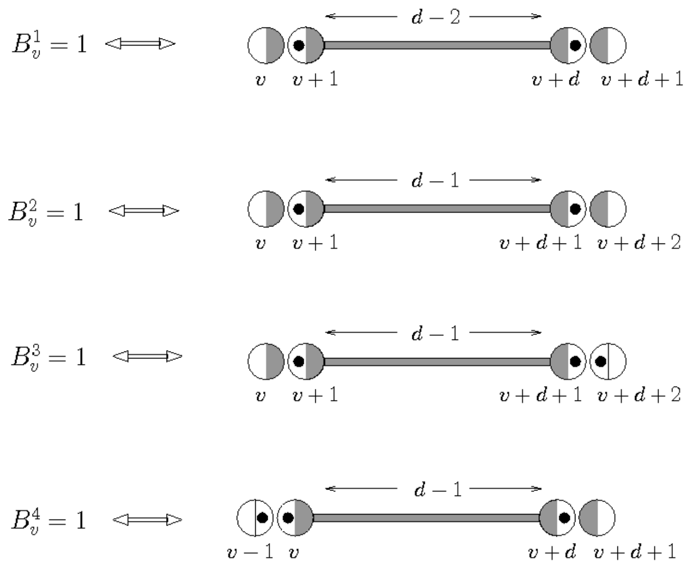


Figure 2.4: Birth of a d-hole line at vertex v

Finally, there are four ways that a d-hole line can die at v between time steps t and $t+1$ according to the following descriptions (see also Figure 2.5):

d1 At time step t, there is a hole between v and $v+d+1$ of exact size d. Then some walker at v moves right and some walker at $v+d+1$ moves left.

d2 At time step t, there is a hole between v and $v+d+2$ of exact size $d+1$. Then some walker at v moves right and some walker at $v+d+2$ moves left.

d3 At time step t there is a hole between v and $v+d+1$ of exact size d, and $v+d+2$ is occupied. Then some walker at v moves right, all walkers at $v+d+1$ move right, and some walker at $v+d+2$ moves left.

d4 At time step t there is a hole between v and $v+d+1$ of exact size d, and $v-1$ is occupied. Then some walker at $v-1$ moves right, all walkers at v move left, and some walker at $v+d+1$ moves left.

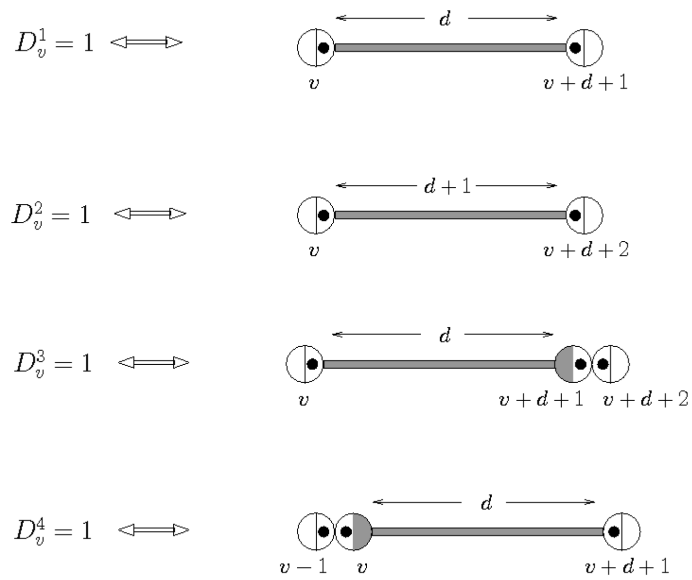


Figure 2.5: Destruction of a d-hole line at vertex v

Given any $v \in V(T_N)$ and for each $\alpha \in \{1,2,3\}$, let S_v^α be the indicator function of the event that a d-hole line survives at v due to movements of type sα. Similarly, for each $\alpha \in \{1,2,3,4\}$, let B_v^α (D_v^α) be the indicator function of the event that a d-hole line is born (dies) at v due to movements of type bα (dα). Hence, $S_v = \sum_{\alpha=1}^{3} S_v^\alpha$, $B_v = \sum_{\alpha=1}^{4} B_v^\alpha$ and $D_v = \sum_{\alpha=1}^{4} D_v^\alpha$ are the indicator variables for a survival, birth and death, respectively, at vertex v.

Fix any naturals ℓ_1, ℓ_2 and ℓ_3, and call $\ell = \ell_1 + \ell_2 + \ell_3$. Let $\boldsymbol{\mathcal{A}} = \big(\{1,\dots,3\}\big)^{\ell_1} \times \big(\{1,\dots,4\}\big)^{\ell_2+\ell_3}$, and let $\boldsymbol{\mathcal{T}}$ be the set of all ℓ-tuples of pairwise different vertices in $V(C_N)$. Given $\boldsymbol{\alpha} = (\alpha_i)_{i=1}^{\ell} \in \boldsymbol{\mathcal{A}}$, and $\boldsymbol{v} = (v_i)_{i=1}^{\ell} \in \boldsymbol{\mathcal{T}}$ let us define the event

$$\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}} = \Bigg(\bigwedge_{i=1}^{\ell_1} (S_{v_i}^{\alpha_i} = 1)\Bigg) \wedge \Bigg(\bigwedge_{i=\ell_1+1}^{\ell_1+\ell_2} (B_{v_i}^{\alpha_i} = 1)\Bigg) \wedge \Bigg(\bigwedge_{i=\ell_1+\ell_2+1}^{\ell} (D_{v_i}^{\alpha_i} = 1)\Bigg). \tag{2.13}$$

This allows us to express the joint factorial moments as

$$\mathbf{E}([S]_{\ell_1}[B]_{\ell_2}[D]_{\ell_3}) = \sum_{\boldsymbol{v}\in\boldsymbol{\mathcal{T}}} \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{A}}} \mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}). \tag{2.14}$$

In order to compute $\mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}})$, we partition the set of tuples $\boldsymbol{\mathcal{T}}$ into three disjoint classes: Let $\boldsymbol{\mathcal{T}}_2$ be the set of tuples $\boldsymbol{v} \in \boldsymbol{\mathcal{T}}$ such that all pairs of vertices in $\boldsymbol{v}$ are at distance greater than $d+3$; Let $\boldsymbol{\mathcal{T}}_1$ be the set of tuples $\boldsymbol{v} \in \boldsymbol{\mathcal{T}} \setminus \boldsymbol{\mathcal{T}}_2$ such that all pairs of vertices in $\boldsymbol{v}$ are at distance greater than $d-2$; Finally, define $\boldsymbol{\mathcal{T}}_0 = \boldsymbol{\mathcal{T}} \setminus (\boldsymbol{\mathcal{T}}_1 \cup \boldsymbol{\mathcal{T}}_2)$.

First observe that if $\boldsymbol{v} \in \boldsymbol{\mathcal{T}}_0$ then $\mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}) = 0$, since some pair of different vertices in $\boldsymbol{v}$ are at distance at most $d-2$ and this is not compatible with $\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}$. Now given any $\boldsymbol{v} \in \boldsymbol{\mathcal{T}}_2$, notice that the regions involved in the descriptions of the events $(S_{v_i}^{\alpha_i} = 1)$, $(B_{v_i}^{\alpha_i} = 1)$ and $(D_{v_i}^{\alpha_i} = 1)$ are disjoint for any choice of $\boldsymbol{\alpha}$. This allows us to compute $\mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}})$ by applying Lemma 2.4.2 to these regions, whose sizes are listed in Table 2.1. For each $j \in \{1,2,3\}$, let

	Size empty region	Size occupied regions
Survival s1	$d+1$	$1/2$
Survival s2	$d+1$	$1/2$
Survival s3	$d+1$	$2\times 1/2$
Birth b1	d	$2\times 1/2$
Birth b2	$d+1$	$2\times 1/2$
Birth b3	$d+1/2$	$3\times 1/2$
Birth b4	$d+1/2$	$3\times 1/2$
Death d1	d	$2\times 1/2$
Death d2	$d+1$	$2\times 1/2$
Death d3	$d+1/2$	$3\times 1/2$
Death d4	$d+1/2$	$3\times 1/2$

Table 2.1: Event descriptions according to their occupancy requirements

a_j be the number of entries α_i of $\boldsymbol{\alpha}$ with $1\le i\le \ell_1$ which are equal to j, i.e. the number of survivals of type sj. Similarly, for each $j\in\{1,2,3,4\}$, let b_j be the number of entries α_i of $\boldsymbol{\alpha}$ with $\ell_1+1\le i\le \ell$ which are equal to j, i.e. the number of births and deaths of types bj and dj. Observe that $\mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}})$ does not depend on the particular $\boldsymbol{v}\in\boldsymbol{\mathcal{T}}_2$ or on the order of the entries of $\boldsymbol{\alpha}$, but only on $\boldsymbol{a}=(a_1,a_2,a_3,)$ and $\boldsymbol{b}=(b_1,b_2,b_3,b_4)$. Hence we can denote this probability by $P_{\boldsymbol{a},\boldsymbol{b}}$, and it satisfies, by (2.11) and Lemma 2.4.2,

$$P_{\boldsymbol{a},\boldsymbol{b}} \sim e^{-\varrho(2\ell_1+2b_2+b_3+b_4)/2}\left(1-e^{-\varrho/2}\right)^{\ell+\ell_2+\ell_3+a_3+b_3+b_4} e^{-d\varrho\ell}. \tag{2.15}$$

From this and also by using

$$\sum_{\boldsymbol{\alpha}}\mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}) = \sum_{\substack{a_1+a_2+a_3=\ell_1\\ b_1+b_2+b_3+b_4=\ell_2+\ell_3}} \binom{\ell_1}{a_1,a_2,a_3}\binom{\ell_2+\ell_3}{b_1,b_2,b_3,b_4}P_{\boldsymbol{a},\boldsymbol{b}},$$

we obtain the contribution to $\mathbf{E}([S]_{\ell_1}[B]_{\ell_2}[D]_{\ell_3})$ due to tuples in $\boldsymbol{\mathcal{T}}_2$

$$\begin{aligned}\sum_{\boldsymbol{v}\in\boldsymbol{\mathcal{T}}_2}\sum_{\boldsymbol{\alpha}}\mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}) &\sim \left(N(1-e^{-\varrho})e^{-d\varrho}\right)^{\ell}\left(\frac{e^{-\varrho}(3-e^{-\varrho/2})}{1+e^{-\varrho/2}}\right)^{\ell_1}\\ &\qquad\left(\frac{(1-e^{-\varrho/2})(1+2e^{-\varrho/2}-e^{-\varrho})}{1+e^{-\varrho/2}}\right)^{\ell_2+\ell_3}\\ &\sim\begin{cases}\mu^{\ell_1}(\mu\varrho/2)^{\ell_2+\ell_3} & \text{if } \varrho=o(1),\\ (\mu-\lambda)^{\ell_1}\lambda^{\ell_2+\ell_3} & \text{if } \varrho=\Theta(1),\\ (3\mu e^{-\varrho})^{\ell_1}\mu^{\ell_2+\ell_3} & \text{if } \varrho=\omega(1),\end{cases}\end{aligned} \tag{2.16}$$

where $\lambda=\mu\left(1-\frac{e^{-\varrho}(3-e^{-\varrho/2})}{1+e^{-\varrho/2}}\right)$.

It only remains to bound the weight of tuples in $\boldsymbol{\mathcal{T}}_1$ in $\mathbf{E}([S]_{\ell_1}[B]_{\ell_2}[D]_{\ell_3})$. Fix some $\boldsymbol{v}\in\boldsymbol{\mathcal{T}}_1$ and observe that the regions involved in the description of $\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}$ need not be disjoint.

This allows that some walkers take part in the birth, death or survival of two different d-hole lines simultaneously (see Figure 2.6 for an example of this situation). Recall that all pairs of vertices in $\boldsymbol{v}$ are at distance greater that $d-2$, but some pair is at distance at most $d+3$. Given any $v_i \in \boldsymbol{v}$, we say that v_i is *restricted* if for some other $v_j \in \boldsymbol{v}$ with $j < i$ we have $\mathrm{dist}(v_i, v_j) \leq d+3$. Let $r > 0$ be the number of restricted vertices of the tuple $\boldsymbol{v}$.

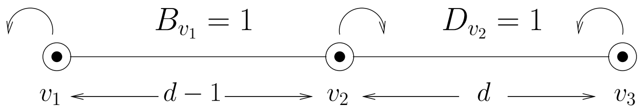


Figure 2.6: Walker taking part in a birth and a death simultaneously.

Suppose first that $\varrho = O(1)$. By looking at the regions involved in the description of $\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}$ assuming that r vertices are restricted, we observe that $\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}$ requires the following: Some region of size at least $\ell(d-2)$ is e.o.w., and moreover each arc in a set of at least $\ell + \ell_2 + \ell_3 - r$ is occupied. Hence, from Lemma 2.4.2,

$$\mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}) \leq \left(1 - \frac{\ell(d-2)}{N}\right)^w (1 - e^{-\varrho/2})^{\ell+\ell_2+\ell_3-r} = O\left(\frac{\varrho^{\ell_2+\ell_3-r}}{N^\ell}\right).$$

Multiplying this by the number $O(N^{\ell-r})$ of tuples in $\boldsymbol{\mathcal{T}}_1$ with exactly r restricted vertices, gives a contribution of $O\left(\frac{\varrho^{\ell_2+\ell_3}}{w^r}\right)$ to (2.14). Recalling that $\mu = \Theta(1)$, this is negligible compared to (2.16).

Otherwise suppose that $\varrho = \omega(1)$. Each survival in the definition of $\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}$ determines a region of size $d+1$ which must be e.o.w. Similarly, each birth and each death determines a region of size at least d which must be e.o.w. However, each of the regions corresponding to a restricted vertex may overlap by at most one arc (size $1/2$) with some other considered region. Summarising, $\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}$ requires that a region of size at least $\ell d + \ell_1 - r/2$ is e.o.w., and thus from Lemma 2.4.2,

$$\mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}) \leq \left(1 - \frac{\ell d + \ell_1 - r/2}{N}\right)^w = O\left(\frac{e^{-\ell_1\varrho + r\varrho/2}}{N^\ell}\right).$$

Multiplying this by the number $O(N^{\ell-r})$ of tuples in $\boldsymbol{\mathcal{T}}_1$ with exactly r restricted vertices, the weight in (2.14) due to these situations is

$$O\left(\frac{e^{-\ell_1\varrho + r\varrho/2}}{N^r}\right) = O\left(\frac{e^{-\ell_1\varrho}}{e^{r(d-1/2)\varrho}}\right),$$

which is negligible compared to (2.16).

We conclude that the main contribution to $\mathbf{E}([S]_{\ell_1}[B]_{\ell_2}[D]_{\ell_3})$ is due to tuples in $\boldsymbol{\mathcal{T}}_2$. As we are assuming μ bounded, it follows, from (2.16) and from Theorem 1.23 in [15], that S, B, D are jointly asymptotically independent Poisson, with the corresponding expectations. □

From (2.12) and Proposition 2.5.6, the following is immediate.

Corollary 2.5.7. *H_t and B_t are asymptotically independent, and so are D_t and H_{t+1}.*

It is natural to define the *lifespan* of a d-hole line as the number of time steps for which the line is alive. For any vertex v and time step t, consider the random variable $L_{v,t}$ defined as follows: If at time step $t+1$ there is a d-hole following vertex v, then $L_{v,t}$ is the number of time steps, possibly infinity, that the corresponding d-hole line stays alive starting from time step $t+1$; Otherwise, $L_{v,t}$ is defined to be 0. So if a birth takes place at vertex v precisely between time steps t and $t+1$, then $L_{v,t}$ corresponds to the lifespan of the d-hole line being born. Note that the random variables $L_{v,t}$ are identically distributed for all v and t.

Observe that we can always reach a state in which there are no d-holes. In view of (2.11), one way to do this is to force the walkers to move to positions in which they are almost equally spaced around the cycle. Then, by Lemma 2.5.4, for any initial state, the process will reach some state with no d-holes within finite expected time. Therefore the expected lifespan of any given d-hole line, given the configuration of walkers at its birth, is finite (it is simply a function of N, d and w). In view of this, we define the *average lifespan* of d-hole lines to be the expected time that a d-hole line will survive once born. Formally,

$$L_{\mathrm{av}} := \mathbf{E}\big(L_{v,t} \mid B_{v,t} = 1\big),$$

where $B_{v,t}$ is the indicator variable of having a birth at vertex v between time steps t and $t+1$. By symmetry, L_{av} is independent of v and t, and so is a function of N, d and w. The next result finds its size.

Theorem 2.5.8. *The average lifespan of d-hole lines satisfies*

$$L_{\mathrm{av}} = \frac{\mathbf{E}H}{\mathbf{E}B} \sim \begin{cases} 2\varrho^{-1} & \text{if } \varrho = o(1), \\ \frac{\mu}{\lambda} & \text{if } \varrho = \Theta(1), \\ 1 & \text{if } \varrho = \omega(1), \end{cases}$$

where λ is defined as in Proposition 2.5.6.

Proof. Observe that for any $t \in \mathbb{Z}$,

$$\sum_{v \in V(C_N)} L_{v,t-1} + \sum_{v \in V(C_N)} B_{v,t} L_{v,t} = H_{t-1} + \sum_{v \in V(C_N)} L_{v,t}, \tag{2.17}$$

where the distributions of H_t, $B_{v,t}$ and $L_{v,t}$ do not depend on v and t. Recall that a state with no d-holes can be reached within finite expected time, so $\mathbf{E}L_{v,t} < +\infty$. In view of this, we take expectations at both sides of (2.17) and get

$$N\mathbf{E}(B_{v,t} L_{v,t}) = \mathbf{E}H.$$

Now the remaining follows easily:

$$\mathbf{E}(L_{v,t} \mid B_{v,t} = 1) = \frac{\mathbf{E}(B_{v,t} L_{v,t})}{\mathbf{P}(B_{v,t} = 1)} = \frac{\mathbf{E}H}{N\mathbf{P}(B_{v,t} = 1)} = \frac{\mathbf{E}H}{\mathbf{E}B}.$$

The final formula for this expression comes immediately from Theorem 2.5.2 and Proposition 2.5.6. ∎

Intuitively, L_{av} measures the average lifespan of the 'typical' d-hole lines appearing in $\big(G(\mathcal{V}_t)\big)_{t\in\mathbb{Z}}$. However, there are other alternative measures of this intuitive notion which are also natural to consider. To give a feeling for the complexity of the question of how long a d-hole line lives, we introduce the following *train paradox.* A student wishes to measure the average length of a train in a station with two separate platforms A and B. Each morning, she chooses either platform A or B, at random. She waits for the first train to leave on that platform, and records its length. She finds after many days that the average length recorded is 9 cars. But she notices that, restricted to the days that the train is already at the platform when she arrives, the average length is only 8 cars.

Could it be that the shorter trains wait longer for her? No, because the trains stop at stations for equal times. Moreover, on any given platform they arrive regularly at equally spaced intervals, so the well known bus paradox does not directly apply.

Which is a better measure, the length of the first train to arrive, or the length of trains arriving at a prescribed time? The former, yielding the answer 9, might seem more natural at first. However, the explanation for the differing answers reveals the other to be meaningful, and perhaps even more so. The data above, in both versions of the paradox, arise if platform A has trains of average length 12 arriving every 10 minutes, and platform B has trains of average length 6 arriving at 5 minute intervals. In an extended time period, recording all the lengths of trains on all platforms will yield 8 as the average.

Returning to the walkers model, when N is even there are many different configurations of walkers that cannot arise from a given initial configuration $\mathcal{V}_0$. Recall that $\mathscr{M}$ is not ergodic, and that each configuration belongs to a closed class of mutually reachable configurations (precisely all those configurations with the same parity). The different classes of states correspond to different platforms in the train paradox. The quantity L_{av} in Theorem 2.5.8 is roughly equivalent to our traveller's measurement of length of trains restricted to those days that a train is just arriving. However, the train paradox shows that this is not the only reasonable measure of average length. Moreover, the situation is even more complicated, as the train paradox would be if on a given platform several trains could arrive simultaneously and also the inter-train time periods were variable . The average length of the first train to arrive would then be affected by any dependence between the length of a train and the time before the previous train. We wish to study the analogue of the average length of trains on a given platform: the average lifespan of d-hole lines given the initial state. If N is divisible by 2 and the initial configuration $\mathcal{V}_0$ is conditioned upon, the walkers process is "locked in" to a future in which the (conditional) average lifespan of d-hole lines may be different from L_{av}. However, we show that for almost all initial configurations this average is essentially asymptotically equal to L_{av}. For $T \in \mathbb{Z}^+$, we define the *average lifespan* of the d-hole lines born in $[0, T-1]$ to be

$$L_T = \frac{\sum\limits_{t=0}^{T-1} \sum\limits_{v\in V} B_{v,t} L_{v,t}}{\Big|\big\{(v,t) : B_{v,t} = 1\big\}\Big|},$$

where the denominator runs over all pairs $(v,t) \in V(C_N) \times \{0, \ldots, T-1\}$. If the denominator is zero (or the numerator is infinite, which happens with probability 0), the value is immaterial, and may be defined as 0. Note that L_T is a function of a given trajectory of the process.

We show that L_T converges in probability as $T \to \infty$ to a random variable which may depend on the class of the initial state, but nothing else. (Actually, the value is in general different for different classes.) Define

$$L^* = \frac{\mathbf{E}(H_0 \mid \mathfrak{V})}{\mathbf{E}(B_0 \mid \mathfrak{V})},$$

where $\mathfrak{V}$ is the random variable which accounts for the closed class in which the initial state $\mathcal{V}_0$ lies. The notation $f \sim g$ a.a.s. used in the following theorem denotes that for all $\epsilon > 0$, a.a.s. $|f/g - 1| < \epsilon$ (see for example [79]).

Theorem 2.5.9. *For the walkers model on C_N, L_T converges in probability as $T \to \infty$ (with N fixed) towards L^*. Furthermore, as $N \to \infty$, we have $L^* \sim L_{\text{av}}$ a.a.s.*

Proof. Let us define the truncated average lifespan of d-hole lines in $[0, T-1]$ to be

$$\overline{L_T} = \frac{\sum_{t=0}^{T-1} H_t}{H_0 + \sum_{t=0}^{T-2} B_t}$$

(defined by convention to be 0 if the denominator is 0). As we prove below, this is an approximation of L_T.

We first deal with the case that N is even. To prove the result we need to take into account the class of states containing the initial one, since different starting configurations of walkers may lead to different expected numbers of holes and births. Let $\mathfrak{V}$ be the random variable accounting for the closed class of states where the initial state $\mathcal{V}_0$ lies. We condition on the value of $\mathfrak{V}$. By Lemma 2.5.4, the hitting time between any two states in this class is finite. Consider the Doob Martingale $\Sigma_0, \dots, \Sigma_T$ defined by

$$\Sigma_i = \mathbf{E}\left(\sum_{t=0}^{T-1} H_t \;\middle|\; \mathfrak{V}, \mathcal{V}_0, \dots \mathcal{V}_{i-1} \right), \quad i = 0, \dots, T.$$

We have $\Sigma_0 = T\mathbf{E}(H_0 \mid \mathfrak{V})$ and $\Sigma_T = \sum_{t=1}^{T} H_t$. (We recall that this last expression is regarded in the probability space conditional upon the value of $\mathfrak{V}$.)

From the fact that the expected hitting time between any two states is finite, we deduce that the differences $|\Sigma_{i+1} - \Sigma_i|$ are bounded above by a constant independent of T. Hence as an immediate consequence of Azuma's inequality we get that, conditional upon the value of $\mathfrak{V}$,

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} H_t = \mathbf{E}(H_0 \mid \mathfrak{V}) \quad \text{in probability} \tag{2.18}$$

and by a similar argument

$$\lim_{T\to\infty} \frac{1}{T} \left(H_0 + \sum_{t=0}^{T-2} B_t \right) = \mathbf{E}(B_0 \mid \mathfrak{V}) \quad \text{in probability.} \tag{2.19}$$

Then, by taking the ratio of (2.18) and (2.19), we get that

$$\lim_{T\to\infty} \overline{L_T} = L^* \quad \text{in probability.} \tag{2.20}$$

The case that N is odd is easier. There is just one closed class of states and $\mathcal{M}(N, w, d)$ is ergodic. By the martingale argument as above, we get (noting Theorem 2.5.8)

$$\lim_{T\to\infty} \overline{L_T} = \frac{\mathbf{E}H_0}{\mathbf{E}B_0} = L_{\text{av}} \quad \text{in probability.} \tag{2.21}$$

Moreover, since H_t and B_t are at most N and the expected lifespan of any line is finite, we obtain

$$\lim_{T\to\infty} L_T - \overline{L_T} = 0 \quad \text{in probability.} \tag{2.22}$$

Thus the portions of lifespans omitted in $\overline{L_T}$ have finite expectation, which is insignificant since the denominator grows linearly with T as shown in (2.19).

In order to finish the proof, it suffices to show that as $N \to \infty$

$$L^* \sim L_{\text{av}} \quad \text{a.a.s.}$$

We note that the quantity $\mathbf{E}(H_0 \mid \mathfrak{V})/\mathbf{E}(B_0 \mid \mathfrak{V})$ may vary depending on the particular closed class of states $\mathfrak{V}$. Let us study this in more detail. Let $\mathcal{V} \in \mathfrak{V}$ be a configuration. As in Subsection 2.5.2, let us partition the set of walkers W into W_1 and W_2 according to the parity of their positions in the cycle. Let us define the *imbalance* of the configuration as $\Delta(\mathcal{V}) = |w_1 - w_2|$ where $w_i = |W_i|$. It makes sense to define $\Delta(\mathfrak{V}) = \Delta(\mathcal{V})$ since it does not depend on the choice of $\mathcal{V} \in \mathfrak{V}$.

We can compute the expectations of S, B and D conditional upon $(\mathcal{V}_0 \in \mathfrak{V})$ by proceeding the same way as in Proposition 2.5.6. The only difference is that w_1 walkers must go to $\frac{N}{2}$ of the vertices (say those with odd position) and w_2 must go to the other $\frac{N}{2}$. We omit details here since they are fairly tedious but completely analogous to the previous computations. We note that these expectations do not depend on the particular partition (W_1, W_2) but only on the imbalance $\Delta(\mathfrak{V})$. In all cases we get the following

$$\frac{\mathbf{E}(H_0 \mid \mathfrak{V})}{\mathbf{E}(B_0 \mid \mathfrak{V})} = \Theta\left(\frac{\mathbf{E}H_0}{\mathbf{E}B_0}\right). \tag{2.23}$$

In fact, for $\frac{\Delta(\mathfrak{V})}{N} = O(1)$ we have $\mathbf{E}(H_0 \mid \mathfrak{V}) = \Theta(\mathbf{E}H_0)$ and $\mathbf{E}(B_0 \mid \mathfrak{V}) = \Theta(\mathbf{E}B_0)$. For $\frac{\Delta(\mathfrak{V})}{N} \to \infty$, these two statements are no longer true, but the extra factors in numerator and denominator of (2.23) cancel out to within a factor of $\Theta(1)$.

However, not all imbalances are equally likely. In fact for any $\epsilon > 0$, we have

$$\mathbf{P}\left(\Delta(\mathcal{V}_0) \geq w^{\frac{1+\epsilon}{2}}\right) = \mathbf{P}\left(\left|w_1(\mathcal{V}_0) - \frac{w}{2}\right| \geq \frac{1}{2} w^{\frac{1+\epsilon}{2}}\right) \leq \frac{w}{w^{1+\epsilon}} = o(1). \tag{2.24}$$

Moreover, for a (typical) class $\mathfrak{V}$ such that $\Delta(\mathfrak{V}) < w^{\frac{1+\epsilon}{2}}$ and by the method explained above, we get

$$\frac{\mathbf{E}(H_0 \mid \mathfrak{V})}{\mathbf{E}(B_0 \mid \mathfrak{V})} \sim \frac{\mathbf{E}H_0}{\mathbf{E}B_0}. \tag{2.25}$$

From this last fact together with (2.23) and (2.24), the theorem follows. □

Before stating the main results in this subsection, we need some definitions and a technical result. Let $\mathcal{E}$ be an event in the static model $G(\mathcal{V})$. We denote by $\mathcal{E}_t$ the event that $\mathcal{E}$ holds at step time t. In the $\big(G(\mathcal{V}_t)\big)_{t\in\mathbb{Z}}$ model, we define $L_t(\mathcal{E})$ to be the number of consecutive steps that $\mathcal{E}$ holds starting at step t (possibly 0 if E_t does not hold, or infinity). Formally,

$$L_t(\mathcal{E}) = \sum_{k=0}^{\infty} 1[\mathcal{E}_t] \cdots 1[\mathcal{E}_{t+k}].$$

Note that the distribution of $L_t(\mathcal{E})$ does not depend on t, and we will often omit the t when it is understood or not relevant.

Lemma 2.5.10. *Consider any event $\mathcal{E}$ in the static model. If we have that $\mathbf{E}(L(\mathcal{E})) < +\infty$ (but possibly $\mathbf{E}(L(\mathcal{E})) \to +\infty$ as $N \to +\infty$), then conditional upon $\mathcal{E}_t$ but not $\mathcal{E}_{t-1}$ we have*

$$\mathbf{E}(L_t(\mathcal{E}) \mid \overline{\mathcal{E}_{t-1}} \wedge \mathcal{E}_t) = \frac{\mathbf{P}(\mathcal{E})}{\mathbf{P}(\overline{\mathcal{E}_{t-1}} \wedge \mathcal{E}_t)},$$

which does not depend on t.

Proof. We have that

$$L_{t-1} + 1[\overline{\mathcal{E}_{t-1}}]L_t = 1[\mathcal{E}_{t-1}] + L_t$$

and taking expectations and using the hypothesis that $\mathbf{E}(L(\mathcal{E})) < +\infty$ we get

$$\mathbf{E}(1[\overline{\mathcal{E}_{t-1}}]L_t(\mathcal{E})) = \mathbf{P}(\mathcal{E}), \quad \forall t \in \mathbb{Z}.$$

Using the fact that

$$\mathbf{E}(L_t(\mathcal{E}) \mid \overline{\mathcal{E}_{t-1}} \wedge \mathcal{E}_t) = \frac{\mathbf{E}(1[\overline{\mathcal{E}_{t-1}} \wedge \mathcal{E}_t]L_t(\mathcal{E}))}{\mathbf{P}(\overline{\mathcal{E}_{t-1}} \wedge \mathcal{E}_t)} = \frac{\mathbf{E}(1[\overline{\mathcal{E}_{t-1}}]L_t(\mathcal{E}))}{\mathbf{P}(\overline{\mathcal{E}_{t-1}} \wedge \mathcal{E}_t)},$$

the result follows. □

We turn now to connectivity issues, for which we use (2.4). For each $t \in \mathbb{Z}$, we define $\mathcal{C}_t$ (respectively $\mathcal{D}_t$) to be the event that $G(\mathcal{V}_t)$ is connected (respectively disconnected). The next lemma gives the probability that the connectedness of $G(\mathcal{V}_t)$ changes between step times t and $t+1$.

Lemma 2.5.11. *The probability that $G(\mathcal{V}_t)$ is connected and that $G(\mathcal{V}_{t+1})$ is disconnected is given by*

$$\mathbf{P}\big(\mathcal{C}_t \wedge \mathcal{D}_{t+1}\big) \sim \begin{cases} \frac{1}{2}\mu^2 e^{-\mu}\varrho & \text{if } \varrho = o(1), \\ e^{-\mu}\left(1+\mu-(1+\mu+\lambda+\lambda^2)e^{-\lambda}\right) & \text{if } \varrho = \Theta(1), \\ (1+\mu)e^{-\mu}(1-(1+\mu)e^{-\mu}) & \text{if } \varrho = \omega(1), \end{cases}$$

where λ is defined as in Proposition 2.5.6.

Proof. In view of (2.4), we have that $\mathbf{P}\big(\mathcal{C}_t \wedge \mathcal{D}_{t+1}\big) = \mathbf{P}\big(H_{t+1} \geq 2 \wedge H_t < 2\big)$. Then we can split this second probability according to the events $H_t = 0$ and $H_t = 1$. Noting that

$$\mathbf{P}\Big(H_{t+1} \geq 2 \wedge H_t = 0\Big) = \mathbf{P}\Big(H_t = 0 \wedge B_t \geq 2\Big),$$

and

$$\mathbf{P}\Big(H_{t+1} \geq 2 \wedge H_t = 1\Big) = \mathbf{P}\Big(S_t + B_t \geq 2 \wedge S_t + D_t = 1\Big)$$
$$= \mathbf{P}\Big(S_t = 1 \wedge B_t \geq 1 \wedge D_t = 0\Big) + \mathbf{P}\Big(S_t = 0 \wedge B_t \geq 2 \wedge D_t = 1\Big),$$

the result follows from Proposition 2.5.6 and Corollary 2.5.7. □

One of our main results concerns the expected time that the graph of walkers remains (dis)connected, after a point in time at which it becomes (dis)connected. Define a *disconnected period* to be a maximal sequence of consecutive time steps for which the graph of walkers $G(\mathcal{V}_t)$ is disconnected. Note that if a disconnected period starts at time step t, then $L_t(\mathcal{D})$ is the random variable counting the length of that disconnected period. By Lemma 2.5.4, from a disconnected state the graph will always reach some connected one within finite expected time (for example, one state in which all walkers occupy one of two adjacent sites). Thus the expected length of any disconnected period is finite (but depending on N). Formally we define the *average length* of a disconnected period starting at time t to be

$$L\mathcal{D}_{\text{av}} := \mathbf{E}(L_t(\mathcal{D}) \mid \mathcal{C}_{t-1} \wedge \mathcal{D}_t).$$

By interchanging the words 'disconnected' and 'connected' and also the events $\mathcal{D}$ and $\mathcal{C}$ in the above definitions, we can define *connected periods* and also $L\mathcal{C}_{\text{av}}$. By symmetry, $L\mathcal{D}_{\text{av}}$ and $L\mathcal{C}_{\text{av}}$ are independent of t, and so they are functions of N, d and w. The next result finds their size.

Theorem 2.5.12. *For the walkers model on the cycle C_N, the average length of a connected and a disconnected period of $\big(G(\mathcal{V}_t)\big)_{t\in\mathbb{Z}}$ satisfy respectively*

$$L\mathcal{C}_{\text{av}} \sim \begin{cases} 2\frac{1+\mu}{\mu^2}\varrho^{-1} & \text{if } \varrho = o(1), \\ \frac{1+\mu}{1+\mu-(1+\mu+\lambda+\lambda^2)e^{-\lambda}} & \text{if } \varrho = \Theta(1), \\ \frac{e^\mu}{e^\mu-(1+\mu)} & \text{if } \varrho = \omega(1) \end{cases} \qquad \textit{and}$$
$$L\mathcal{D}_{\text{av}} \sim \begin{cases} 2\frac{e^\mu-1-\mu}{\mu^2}\varrho^{-1} & \text{if } \varrho = o(1), \\ \frac{e^\mu-1-\mu}{1+\mu-(1+\mu+\lambda+\lambda^2)e^{-\lambda}} & \text{if } \varrho = \Theta(1), \\ \frac{e^\mu}{1+\mu} & \text{if } \varrho = \omega(1), \end{cases}$$

where λ is defined as in Proposition 2.5.6.

Proof. Recall that a connected state is reached within finite expected time starting from any given state, so $\mathbf{E}(L(\mathcal{D})) < +\infty$. Then in view of Lemma 2.5.10,

$$L\mathcal{D}_{\text{av}} = \mathbf{E}(L_t(\mathcal{D}) \mid \mathcal{C}_{t-1} \wedge \mathcal{D}_t) = \frac{\mathbf{P}(\mathcal{D})}{\mathbf{P}(\mathcal{C}_{t-1} \wedge \mathcal{D}_t)},$$

and the asymptotic value of $L\mathcal{D}_{\text{av}}$ follows from Corollary 2.6.7 and Lemma 2.5.11. The computation of $L\mathcal{C}_{\text{av}}$ is completely analogous. □

It is also interesting to ask how the average length of a (dis)connected period relates to the initial configuration of the walkers (see the train paradox discussed earlier in this subsection). As noted previously, if N is odd then $\mathscr{M}$ is ergodic and parity is immaterial. However, if N is even and the initial configuration is conditioned upon, the (conditional) average length of these periods may be different from $L\mathcal{C}_{\mathrm{av}}$ and $L\mathcal{D}_{\mathrm{av}}$. For $T \in \mathbb{Z}^+$, we define the *average disconnection time* of the graph of walkers in $[1, T]$ to be

$$L\mathcal{D}_T = \frac{\sum\limits_{t=1}^{T} L_t(\mathcal{D})}{\Big|\big\{t \in \{1, \ldots, T\} \,:\, L_t(\mathcal{D}) > 0\big\}\Big|}.$$

By changing $\mathcal{D}$ to $\mathcal{C}$ in the previous definition, we define $L\mathcal{C}_T$, the *average connection time* of the graph of walkers in $[1, T]$. We show that $L\mathcal{C}_T$ and $L\mathcal{D}_T$ converge in probability as $T \to \infty$ to a random variable which may depend on the class of the initial state, but nothing else. (Actually, the value is in general different for different classes.) Define

$$L\mathcal{C}^* = \frac{\mathbf{P}(\mathcal{C} \mid \mathfrak{V})}{\mathbf{P}(\mathcal{D}_{t-1} \wedge \mathcal{C}_t \mid \mathfrak{V})} \qquad \text{and} \qquad L\mathcal{D}^* = \frac{\mathbf{P}(\mathcal{D} \mid \mathfrak{V})}{\mathbf{P}(\mathcal{C}_{t-1} \wedge \mathcal{D}_t \mid \mathfrak{V})},$$

where $\mathfrak{V}$ is the random variable accounting for the closed class in which the initial state $\mathcal{V}_0$ lies. The following is an analogue of Theorem 2.5.9.

Theorem 2.5.13. *For the walkers model on C_N, $L\mathcal{D}_T$ converges in probability as $T \to \infty$ (with N fixed) towards $L\mathcal{D}^*$. Furthermore, as $N \to \infty$, we have $L\mathcal{D}^* \sim L\mathcal{D}_{\mathrm{av}}$ a.a.s. The same statements hold changing $\mathcal{D}$ to $\mathcal{C}$.*

Proof. We define the truncated average disconnection time of the graph of walkers in $[1, T]$ as

$$\overline{L\mathcal{D}_T} = \frac{\sum_{t=1}^{T} 1[\mathcal{D}_t]}{\sum_{t=1}^{T} 1[\mathcal{C}_{t-1} \wedge \mathcal{D}_t]}.$$

The same argument in the proof of Theorem 2.5.9, but replacing H_t with $1[\mathcal{D}_t]$ and B_t with $1[\mathcal{C}_{t-1} \wedge \mathcal{D}_t]$, yields the statement. The proof for $L\mathcal{C}^*$ is completely analogous. ☐

2.6 The Grid

In this section, we study the walkers model for $M_N = T_N$, the toroidal grid with $N = n^2$ vertices. We can refer to vertices by using coordinates in $\mathbb{Z}_n \times \mathbb{Z}_n$. For the grid we encounter significant new obstacles as compared to the cycle; see for instance the Geometric Lemma below.

For any $p \in [1, \infty]$ and any two vertices u and v in T_N, we define the distance $\mathrm{dist}_{\ell^p}(u, v)$ as the minimal ℓ^p distance between any two points u' and v' in the square grid such that the coordinates of u' are congruent to those of u taken modulo n, and similarly for v and v'. Let us fix any such metric for our study of the connectedness of $G(\mathcal{V})$, and write $\mathrm{dist}(u, v)$ for short. We use $\mathrm{dist}(\cdot, \cdot)$ to refer to this measure of distance, in distinguishing it from Euclidean distance. Note that

$$\frac{1}{2} \cdot \mathrm{dist}_{\ell^1}(u, v) \le \mathrm{dist}(u, v) \le \mathrm{dist}_{\ell^1}(u, v). \tag{2.26}$$

The number of vertices at distance at most d from any given one is $h = \Theta(d^2)$. The exact expression of h depends on the choice of the metric. Some examples are found in Table 2.2.

Observation 2.6.1. Assume that $d < 2n$ (otherwise the graph of walkers is always complete). For each $i, j < 4n/d$, let v_{ij} denote the vertex in T_N with coordinates $(\lfloor id/4 \rfloor, \lfloor jd/4 \rfloor)$. Let $\mathcal{S}_{ij}$ denote the set of grid points closer to v_{ij} than any of the other $v_{i'j'}$. Then there are $\Theta(N/d^2)$ disjoint sets $\mathcal{S}_{ij}$ each containing $\Theta(d^2)$ points. The probability that at least one of these S_{ij} is empty of walkers is at most

$$\Theta(N/d^2)(1 - \Theta(d^2/N))^w = O(\sqrt{w})e^{-\Omega(\sqrt{w})},$$

which goes to 0 if $d^2 = \Omega(N/\sqrt{w})$. Thus, a.a.s. each of these pieces is occupied by at least one walker, and $G(\mathcal{V})$ is connected.

In view of the observation, we assume for the rest of the section that $h = o(N)$, i.e. $d = o(n)$. If $d = \Omega(n)$, then $G(\mathcal{V})$ is a.a.s. connected.

Metric	h
ℓ^1	$h = 2d(d+1)$
ℓ^2	$h \sim \pi d^2$ if $d \to \infty$
ℓ^∞	$h = 4d(d+1)$

Table 2.2: Number of vertices at distance at most d from a given vertex.

We wish to study the connection and disconnection of $G(\mathcal{V})$ in a similar way to the cycle. For the grid, the notion of hole does not help, and we deal directly with components. Recall from the introduction that a simple component is one with just one vertex. These play a major role, and we shall prove that, for the interesting values of the parameters, a.a.s. there only exist simple components besides one giant one.

Let Γ be any given component. The *edges* of Γ are the straight edges joining occupied vertices in Γ of distance at most d. The associated *forbidden region* $\mathcal{A}_\Gamma$ is the set of vertices not in Γ, but at distance at most d from some vertex in Γ (i.e. those vertices which must be free of walkers for Γ to exist as a component). The *exterior* $\mathcal{E}_\Gamma$ of Γ is the set containing all those vertices not in $\Gamma \cup \mathcal{A}_\Gamma$. We partition $\mathcal{E}_\Gamma$ into *external regions* as follows: two vertices belong to the same external region when they can be joined by a continuous curve not intersecting any edge of Γ. Figure 2.7 shows a component with different external regions.

Recall that, in the terminology of planar maps, the *bounding cycle* of a face is a walk around the boundary of the face. We say that it is *positively oriented* if the walk keeps the face to the left. Given an external region $\mathcal{E}_\Gamma^i$, let Γ' be any connected subgraph of Γ that has no edges crossing and such that no vertices of Γ are contained in the face F of Γ' which contains $\mathcal{E}_\Gamma^i$. Such subgraphs always exist: for instance, take the spanning tree of Γ whose length (sum of lengths of edges) in terms of dist has been minimised, and, subject to this, has the shortest Euclidean length. We refer to the positively oriented bounding cycle of this face F as a *boundary walk* β in Γ with respect to $\mathcal{E}_\Gamma^i$. Such a walk is *maximal* if the face F does not properly contain a face of some other subgraph of Γ of the same type. In Figure 2.6, $(1, 5, 3, 7, 9, 11, 13, 15, 13, 12, 14, 12, 10, 8, 6, 2, 4, 1)$ is a non-maximal boundary walk, and $(1, 5, 3, 7, 9, 11, 13, 15, 13, 14, 12, 10, 8, 6, 2, 4, 1)$ is a maximal one. A maximal one

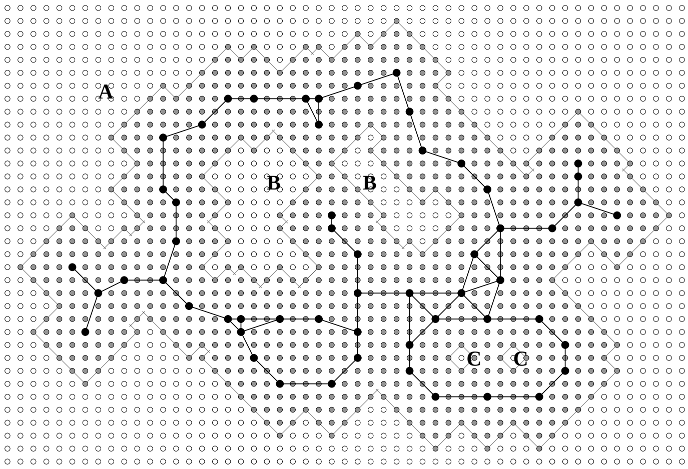


Figure 2.7: Component (black), forbidden region (gray), external regions A, B and C (white)

always exists, because any non-maximal one can be diverted around any face that prevents it from being maximal. Note each edge in a boundary walk, appears at most twice, once in

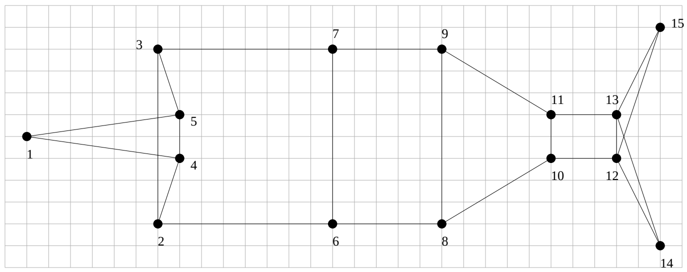


Figure 2.8: The boundary walk

each direction.

For $i < n$, let us call a *v-band* of width i to any subset of T_N defined by $\{a, \ldots, a + i - 1\} \times \mathbb{Z}_n$. Similarly, we define a *h-band* of height j. Define a *rectangle* of width i and height j to be the intersection of a v-band of width i and a h-band of height j. We can compare vertices in a rectangle according to their coordinates, and use statements such as v_1 *is more left than* v_2 or v_3 *is an uppermost vertex* in the rectangle.

We say that a component Γ with at least 2 vertices is *embeddable* if all of its vertices are contained in a rectangle of width $n - 2d$ and height $n - 2d$. In particular, this implies that Γ contains no non-separating cycle of the torus. For a given embeddable component Γ, we define its *origin* as the leftmost of the lower-most vertices of Γ. The *outside region* of an embeddable component is the only external region of the component having vertices outside any rectangle containing the component.

On the other hand, these components which are not embeddable wrap around the torus, and must be large. In fact, each of such components must contain at least $\Theta(n/d)$ vertices. Note that sometimes several non-embeddable components can coexist together. However, there are some non-embeddable components which are so spread around the torus that do not allow any room for other non-embeddable ones. We call these components *solitary*. Formally, a non-embeddable component is solitary if it is not the subgraph of some graph of walkers containing more than one non-embeddable components. Note that this property is simply determined by the subgraph of the toroidal grid remaining when the component and the vertices of distance at most d from it are deleted. By definition we can have at most one solitary component. We cannot disprove the existence of this solitary component, since with probability $1-o(1)$ there exists one giant component of this nature. For components which are not solitary (i.e., either embeddable or non-embeddable but able to coexist with other non-embeddable ones), we will give asymptotic bounds on the probability of their existence. We will show that, under certain conditions, they seldom occur, so we just find a few simple components and a giant one which is solitary.

Let X, Y and Z be respectively: the number of simple components; the number of embeddable components; and the number of non-embeddable components which are not solitary.

2.6.1 Static Properties

In this subsection, we study the connectedness of $G(\mathcal{V})$ in the static situation for the case $M_N = T_N$; in particular, we analyse the behaviour of X, Y and Z. To examine the connectedness of $G(\mathcal{V})$, we need to redefine the μ used for the cycle. For the remaining of Section 2.6, let $\mu = N\left(1-e^{-\varrho}\right)e^{-h\varrho}$. Hence

$$\mu \sim \begin{cases} we^{-h\varrho} & \text{if } \varrho = o(1), \\ N\left(1-e^{-\varrho}\right)e^{-h\varrho} & \text{if } \varrho = \Theta(1), \\ Ne^{-h\varrho} & \text{if } \varrho = \omega(1). \end{cases}$$

We first characterise the number of simple components in terms of μ.

Proposition 2.6.2. *The expected number of simple components of $G(\mathcal{V})$ for T_N satisfies*

$$\mathbf{E}X \sim N\left(1-e^{-\varrho}\right)\left(1-\frac{h}{N}\right)^{w}.$$

Furthermore,

(i). if $\mu \to 0$ then $\mathbf{E}X \to 0$ and there are no simple components a.a.s.,

(ii). if $\mu \to \infty$ then there exist simple components a.a.s. (and $G(\mathcal{V})$ is disconnected),

(iii). if $\mu = \Theta(1)$ then X is asymptotically Poisson with mean μ.

Proof. We repeat the proof of Theorem 2.5.2 in the present context. To compute $\mathbf{E}[X]_k$, we focus on the set $\boldsymbol{\mathcal{T}}_1 = \{(v_1,\dots,v_k) |\, \mathrm{dist}(v_i,v_j) > 2d \text{ if } i \neq j\}$, with size $|\boldsymbol{\mathcal{T}}_1| \sim N^k$. Applying Lemma 2.4.2, we obtain

$$\mathbf{E}[X]_k \sim \left[N\left(1-e^{-\varrho}\right)e^{-h\varrho+O(h^2w/N^2)}\right]^k. \tag{2.27}$$

Comparing with (2.7), the rest of the proof is as for Theorem 2.5.2. □

From part (ii) of the proposition, if $h\varrho = O(1)$ then $\mu \to \infty$ and $G(\mathcal{V})$ is disconnected a.a.s. In view of this, we may restrict to the condition $h\varrho \to \infty$ in the study of embeddable components and non-embeddable components which are not solitary.

Given a boundary walk $\beta = (v_1, \ldots, v_k)$ we define its length as the sum of the distances (using the chosen metric) between consecutive vertices in β.

$$\text{length}(\beta) = \sum_{1 \le i < k} \text{dist}(v_i, v_{i+1})$$

We shall write $\text{length}_{\ell^p}(\beta)$ when we want specify that we are measuring distances in ℓ^p. Similarly we define $\text{length}_v(\beta)$ (the vertical length) as the sum of the differences between y coordinates of consecutive vertices along the cycle, and $\text{length}_h(\beta)$ (the horizontal length) using x coordinates in the same way.

The next lemma relates the size of the forbidden region outside a boundary cycle of a component to the length of the cycle, and will play a key role in proving the main results.

Lemma 2.6.3 (Geometric Lemma). *Let Γ be a component in T_N with β one of its maximal boundary walks, and $l = \text{length}(\beta)$ its length. Assume that Γ has at least two occupied sites. Then the size of the forbidden region $\mathcal{A}_\beta$ outside β is bounded below by $|\mathcal{A}_\beta| \ge dl/J$, for some sufficiently large constant J. Moreover, if Γ is rectangular, and β is a maximal boundary walk with respect to the outside region, we have $|\mathcal{A}_\beta| \ge h + dl/J$.*

Proof. For convenience we take $J = 10^{10}$, though probably without large modifications the proof method will yield the result for $J = 1000$. Observe that in the 3×3 subgrid centred on the endpoint of an edge of β, there must be at least one vertex in the forbidden region outside β; otherwise the boundary walk can be re-routed to contradict its maximality. Each point of forbidden region can be counted in this way at most 8 times. Thus the forbidden region has size at least $l/(8d)$. This implies the first statement in the lemma provided that $8d^2 \le 10^{10}$.

Throughout this proof, all distances referred to are measured using $\text{dist}(\cdot)$. For the second statement in the lemma, we begin with the fact that for a rectangular component, there are "caps" of empty region of size at least $h/2 - d$ on top and bottom of the component, being the set of grid points within distance d, but above the leftmost point of Γ of greatest vertical coordinate (or below the leftmost of least vertical coordinate). Without loss of generality, we assume that the component has vertical height greater than 1 (otherwise, we may interchange "vertical" and "horizontal"). Then there are also intervals of empty region of length d projecting outwards from the left- and right-most vertices of the top level of Γ. These caps and intervals account for an forbidden region of size h, but use up some of the forbidden region counted in relation to ends of edges of β in the argument above. The number of end-vertices of edges of β that are involved in the region of size h described is at most $4d + 2$ and there is still at least one end-vertex of an edge of β not involved (the rightmost one in the bottom row of Γ, say — we cannot claim two vertices here because it might be the same as the leftmost one) which means the above argument is valid if the bound on the size of the forbidden region is reduced by a factor of $4d + 3$. That is, the size of the forbidden region is at least $h + l/8d(4d + 3)$. So the second statement in the lemma holds provided that $8d^2(4d + 3) \le 10^{10}$. Hence we may now assume that $d > 400$.

Choose $k = \lceil \frac{d}{100} \rceil$. Then $k \le (d + 99)/100 \le (d/100)(1 + 99/d) < d/80$. Assume without loss of generality that for the boundary walk β we have $\text{length}_v(\beta) \ge \text{length}(\beta)/2$.

We place intervals of length k (note that all the ℓ_p metrics measure these intervals the same) horizontally along all grid lines from a maximal boundary cycle towards the outside. Those starting from a vertex of β point towards the outside according to the previous edge of β. (We assume β is oriented in some direction.) We delete any intervals that touch the boundary cycle in two or more places. Then each remaining interval will touch k vertices of the forbidden region outside β, and each such vertex will be touched by at most two different intervals. (Two intervals coming from opposite directions may touch the same vertex.)

We need to bound the number of intervals which were deleted. Call an edge of Γ *short* if the distance between its endpoints is less than $d/4$, and *long* otherwise. Suppose that an interval (that is to be deleted) touches a short edge at a point E and a point F on another edge of β, with the part between E and F in the exterior of β. Suppose that F has distance at most $d-d/4-k$ from both end-vertices of its edge. Then by the triangle inequality each end-vertex (say P and Q) of that edge is of distance at most d from each end-vertex (R and T) of E's edge. Thus the quadrilateral $PQRT$ (or triangle, if two points coincide) has diameter at most d, and thus the walk β can be changed to make a smaller face "outside" (meaning the side which was minimised). This can clearly be done also if other parts of β enter this quadrilateral, contradicting the maximality of β. Thus, F has distance at least $d-d/4-k > 2d/3$ from one of the end-vertices of its edge. So F's edge is long, or, if F lies on more than one edge, they are all long.

We call a *middle* interval any interval originating from a long edge from point at least $1/8$ of the length of the edge from each end. Suppose that such a middle interval originating at a point E (called a *middle point*), on edge e of β, is deleted. Then it hits some other edge f at a point F of horizontal distance at most k from E. If f has an end-vertex of distance less than $d/8-k \geq d/8-d/80 \geq d/9$ from F, we get a contradiction as above. Since e is long, its end vertices have distance at least $d/32$ from E. Since e and f are straight, either point on e of distance $rd/32$ from E, for any $r>0$, has horizontal distance at most $(r+1)k$ from f or the extension of f. (Furthermore, note for later that every pair of points on e and f at the same vertical coordinate have distance at most $32k$ apart.) It follows that the end vertices P and R of e and f above the line EF have distance less than d, as do the ends Q and T below (or on) the line EF. If either $P=R$ or $Q=T$ we now obtain a contradiction as before by re-routing β. A similar contradiction arises if either PT or QR has length at most d. So, picturing e and f with the line EF as making a near-perfect but very thin "H", the distances along the H from P to T and from Q to R are both greater than d. Thus the sum of lengths of e and f is at least $2d-2k$ (since the edge EF is included twice in these distances) and so each of e and f has length at least $d-2k$, and the lengths of the arms of e and f above EF differ by at most k (as for the arms below EF). It follows that both PR and QT have length at most $34k$ (recalling the observation above about horizontal distance), which is at most $d/2$. Without loss of generality, PR is not an edge of β. This again contradicts the choice of β, either by shortening it at PR or, if another edge of β crosses PR, by joining P or R to an end of such an edge. We conclude that no middle intervals are deleted. Moreover, this shows that the point F in the previous paragraph cannot be a middle point. So every interval starting at a short edge that is deleted first hits a non-middle points of a long edge. Such a point can be hit by only one interval from a short edge unless it is a vertex of β, in which case it can be hit by two intervals. Thus if i_1 is the number of intervals starting at short edges that are deleted, j_1 is the number of long edges of β, and j_2 is the number of non-middle intervals originating at non-ends of long intervals,

$$i_1 \leq 4j_1 + j_2.$$

Moreover, since no middle intervals are deleted, the number i_2 of intervals originating on long edges that are deleted similarly satisfies

$$i_2 \leq 4j_1 + j_2.$$

If j_3 is the total number of intervals before any deletions occur, we clearly have

$$j_2 \leq j_3/4$$

and also since long edges have length at least $d/4$,

$$j_3 \geq \text{length}_v(\beta) \geq \text{length}(\beta)/2 \geq j_1 d/8. \tag{2.28}$$

Combining these gives

$$i_1 + i_2 \leq 8j_1 + 2j_2 \leq j_3 \left(\frac{64}{d} + \frac{1}{2} \right) \leq 3j_3/4$$

which shows that at least $j_3/4$ of all the intervals are not deleted. These intervals each cover $k \geq d/100$ vertices of the forbidden region, at most two covering any one such vertex, so the first part of the lemma follows for $d > 400$, using (2.28).

For the second claim of the lemma when $d > 400$, we may again add the caps of size $h - 2d$ but also two extra intervals of length d at the sides: assuming that β is oriented in the clockwise direction, the intervals of length k projecting from the left-most vertex in the bottom level of Γ, and from the right-most vertex in the top level, are not used. It is here that extra intervals of empty region of length d may be found. □

The next lemma will be used to show that non-embeddable components which are not solitary are rare.

Lemma 2.6.4. *Let Γ be a component which is not embeddable and not solitary either. Then Γ has a maximal boundary walk β with* length$(\beta) \geq n - o(n)$*.*

Proof. Let us divide T_N into $\lfloor \frac{n}{d+1} \rfloor$ v-bands $c_1, \ldots, c_{\lfloor \frac{n}{d+1} \rfloor}$ of width $\geq d+1$, and similarly into $\lfloor \frac{n}{d+1} \rfloor$ h-bands $r_1, \ldots, r_{\lfloor \frac{n}{d+1} \rfloor}$ of width $\geq d+1$.

Let Γ be a component which is not embeddable and not solitary either. Since Γ is connected, the v-bands (or h-bands) not containing vertices of Γ must be consecutive. If there were at least 2 consecutive v-bands and at least 2 consecutive h-bands without vertices of Γ, then Γ would be embeddable since Γ would be contained in the complementary of the v-bands and h-bands. Hence, at most one v-band and some consecutive h-bands (or at most one h-band and some consecutive columns) may be without vertices of Γ. From the fact that Γ is not solitary, let us assume that Γ coexists with another non-embeddable component Γ'.

Case 1. Let us suppose first that Γ has no vertices in more that one v-band or h-band (v-band without loss of generality). Let c_1, c_2 be two consecutive v-bands not containing vertices of Γ. Hence all h-bands, excepting at most one, contain vertices in Γ. For each such h-band r_i, choose a vertex v_i in $\Gamma \cap r_i$. We can also find some vertex w_i in $(c_1 \cup c_2) \cap r_i$ such that w_i is at distance $\geq d+1$ from any vertex in Γ. By this construction, all w_i belong to the same external region of the component. Let β be any maximal boundary walk of Γ with respect to this external region. Then, the straight line joining v_i and w_i must intersect an edge of β, part of the edge contained in r_i. Hence, β crosses all h-bands except at most 3 and $\text{length}(\beta) \geq n - 4d - 3$.

Case 2. On the other hand, let us suppose that Γ has vertices in all v-bands and h-bands except for at most one of each. Without loss of generality, the other component, Γ', has vertices in all v-bands except for at most one. Thus, there are at least $\lfloor \frac{n}{d+1} \rfloor - 2$ v-bands with some vertices of both components. For each such v-band c_i, let us take vertices $v_i \in \Gamma \cap c_i$, $w_i \in \Gamma' \cap c_i$ and join them by a straight line. Notice that all the w_i belong to the same external region of Γ, and let β be any maximal boundary walk with respect to this region. Then the line joining v_i and w_i must intersect an edge of β, part of the edge contained in c_i. Hence, β crosses all v-bands except at most 4 and $\text{length}(\beta) \geq n - 5d - 4$. □

The next technical result shows that simple components are predominant a.a.s. in T_N. The proof uses the Geometric Lemma.

Lemma 2.6.5. *If $h\varrho \to \infty$, then $\mathbf{E}Y = o(\mathbf{E}X)$ and $\mathbf{E}Z = o(\mathbf{E}X)$.*

Proof. Let us first estimate the expected number $\mathbf{E}Y$ of rectangular components with more than one vertex.

Notice from (2.26) that the ℓ^1-length of the edges of any boundary walk of a component are integers between 1 and $2d$.

Let $\mathcal{B}$ be the set of walks in T_N which are (for some configuration of the walkers) a maximal boundary walk of some embeddable component with respect to its outside region. For each $\beta \in \mathcal{B}$, choose a rooted spanning tree $T(\beta)$ of the graph induced by the edges of β. Note that given any such tree T of m vertices, we may recover β by joining certain pairs of vertices (with no edges crossing). The edges added are just diagonals added to a face of degree $2m-2$.

For each vertex $v \in V$, natural $m \geq 2$ and tuple $\boldsymbol{l} = (l_1, \ldots, l_{m-1})$ of naturals $1 \leq l_i \leq 2d$, let $\mathcal{B}_{v,m,\boldsymbol{l}}$ be the set of all $\beta \in \mathcal{B}$ such that $T(\beta)$ has m vertices, is rooted at v and has edges of ℓ^1-lengths $l_1, \ldots, l_{m-1}$. The number of such trees is at most $\prod_{j=1}^{m-1} 16 l_j$, where a factor 4^m comes from the number of rooted plane trees, and each factor $4l_j$ is the number of vertices of ℓ^1 distance l_j from a given vertex. Therefore,

$$|\mathcal{B}_{v,m,\boldsymbol{l}}| \leq \prod_{j=1}^{m-1} C l_j, \tag{2.29}$$

where C is constant, and also clearly

$$\mathbf{E}Y \leq \sum_{\substack{v \in V \\ m \geq 2 \\ 1 \leq l_1, \ldots, l_{m-1} \leq 2d}} \sum_{\beta \in \mathcal{B}_{v,m,\boldsymbol{l}}} \mathbf{P}(Y_\beta = 1), \tag{2.30}$$

where Y_β indicates the event of having some embeddable component with β being a maximal boundary walk of the component with respect to its outside region.

By Lemma 2.6.3, the size $|\mathcal{A}_\beta|$ of the forbidden region outside β is an integer bounded below by $h + dl/J$, where $l = \text{length}(\beta)$. For technical purposes we consider a subset of $\mathcal{A}_\beta$ of size $h + \lceil dl/(2J+1)\rceil$, representing an region free of walkers. By Lemma 2.4.3, and noting that $\alpha > 1/2$ and hence $p < 2\varrho$, we obtain an upper bound for the probability of this emptiness occurring and the m vertices in β being occupied. Since this is necessary for the event Y_β to occur, we have

$$\begin{aligned}\mathbf{P}(Y_\beta = 1) &= O(1-e^{-\varrho/\alpha})(2\varrho)^{m-1}\left(1-\frac{S}{N}\right)^w \\ &= O(1-e^{-\varrho})(2\varrho)^{m-1}\left(1-\frac{h}{N}\right)^w e^{-\lceil dl/(2J+1)\rceil\varrho} \qquad (2.31)\end{aligned}$$

since $\alpha \le 1$. Furthermore, let $l' = l_1 + \cdots + l_{m-1}$. Then since the spanning tree has length no more than the length of β, $\text{length}_{\ell^1}(\beta) > l'$. By (2.26), we have $l \ge l'/2$, and hence we get

$$\mathbf{P}(Y_\beta = 1) = O(1-e^{-\varrho})(2\varrho)^{m-1}\left(1-\frac{h}{N}\right)^w e^{-dl'\varrho/J'}, \qquad (2.32)$$

where $J' = 2(2J+1)$.

From (2.30), (2.29) and (2.32), we get

$$\mathbf{E}Y = O(1) \sum_{\substack{m\ge 2\\ 1\le l_1,\ldots,l_{m-1}\le 2d}} N\left(\prod_{j=1}^{m-1} Cl_j\right)\left(1-e^{-\varrho}\right)(2\varrho)^{m-1}\left(1-\frac{h}{N}\right)^w e^{-dl'\varrho/J'}$$

Therefore, using Proposition 2.6.2 for the asymptotic value of $\mathbf{E}X$,

$$\begin{aligned}\mathbf{E}Y/\mathbf{E}X &= O(1) \sum_{\substack{m\ge 2\\ 1\le l_1,\ldots,l_{m-1}\le 2d}} \left(\prod_{j=1}^{m-1} Cl_j\right)(2\varrho)^{m-1}e^{-dl'\varrho/J'} \\ &= O(1)\sum_{m\ge 2}\left(\frac{C'}{d}\sum_{k=1}^{2d} kd\varrho e^{-kd\varrho/J'}\right)^{m-1}. \qquad (2.33)\end{aligned}$$

In the case where $d\varrho \to \infty$, we have $\varrho e^{-c'd\varrho} = o(\max(1,\varrho)) = o(1)$ and hence $\mathbf{E}Y = o(\mathbf{E}X)$. In the case where $d\varrho = O(1)$, we use $\sum_{k\ge 1} kc^{-\epsilon kd\varrho} = O((d\varrho)^{-2})$ and (2.33) gives

$$\mathbf{E}Y/\mathbf{E}X = O(1)\sum_{m\ge 2}\left(\frac{C''}{d^2\varrho}\right)^{m-1} = O\left(\frac{1}{\log w}\right) = o(1) \qquad \text{as } d^2\varrho \to \infty.$$

To prove $\mathbf{E}Z = o(\mathbf{E}X)$, from Lemma 2.6.4, each component counted by Z has some maximal boundary walk β with $\text{length}(\beta) \ge n - o(n)$. If we apply Lemma 2.6.3 to this β, we have $|\mathcal{A}_\beta| \ge ld/J$, where $l = \text{length}(\beta)$. Using $d = o(n)$ (since $h = o(N)$), we have $|\mathcal{A}_\beta| \ge h + ld/2J$ for large N, and we then proceed similarly as for Y. □

Finally we derive the main result of this subsection.

Theorem 2.6.6.

(i). For $\mu \to \infty$, $G(\mathcal{V})$ is disconnected a.a.s.

(ii). For $\mu = \Theta(1)$, a.a.s. all but one components of $G(\mathcal{V})$ are simple, and the number X of simple components is asymptotically Poisson with expected value μ.

(iii). For $\mu \to 0$, $G(\mathcal{V})$ is connected a.a.s.

Proof. From Proposition 2.6.2, if $\mu \to \infty$, then $G(\mathcal{V})$ is disconnected a.a.s. In the other two cases, $\mu = O(1)$ and we must have $h\varrho \to \infty$. In this case we can apply Lemma 2.6.5, and get

$$\mathbf{P}(Y > 0) \leq \mathbf{E}Y = o(\mathbf{E}X) = o(1) \quad \text{and} \quad \mathbf{P}(Z > 0) \leq \mathbf{E}Z = o(\mathbf{E}X) = o(1).$$

Thus, a.a.s. we only have simple components and at most one solitary component. The rest of the theorem follows from the asymptotic distribution of X given in Proposition 2.6.2. □

The theorem immediately gives the following.

Corollary 2.6.7. $\mathbf{P}(G(\mathcal{V})$ *is connected*$) = e^{-\mu} + o(1)$.

2.6.2 Dynamic Properties

According to the model, from an initial random placement of the walkers, at each step, every walker moves from its current position to one of its neighbours, with probability 1/4 of going either way. This is a standard random walk on the grid for each walker. We wish to study the connectivity properties of $G(\mathcal{V}_t)$. The analysis of the dynamic case is quite similar to that of the cycle, so we state the major results, and point to the differing details in the proofs.

We define *states* (or *configurations*) and the graph of configurations in an analogous way to the cycle (see Subsection 2.5.2). In this case, there are $N^w = n^{2w}$ different configurations of walkers, each one represented by a vector $\mathcal{V} = (v_1, \ldots, v_w) \in (\mathbb{Z}_n \times \mathbb{Z}_n)^w$ where $v_i = (v_{ix}, v_{iy})$ indicates the label of the vertex being occupied by walker i. Given a configuration $\mathcal{V} = (v_1, \ldots, v_w)$, there exists an edge between $\mathcal{V}$ and all configurations $\mathcal{U} = (u_1, \ldots, u_w)$, such that $\mathrm{dist}(v_i, u_i) = 1$ for all $i \in \{1, \ldots, w\}$. Thus, any configuration has 4^w neighbours, and the relationship of being neighbours is symmetric. As in the case of the cycle, the dynamic process can be seen as a random walk on the graph of configurations, thus a Markov chain $\mathscr{M} = (\mathcal{V}_t)_{t \in \mathbb{Z}}$.

For N even, given any two configurations $\mathcal{U}$ and $\mathcal{V}$, we say that they have the *same parity* if for all i and j, $(u_{i,x} - u_{j,x}) + (u_{i,y} - u_{j,y}) \equiv (v_{i,x} - v_{j,x}) + (v_{i,y} - v_{j,y}) \mod 2$. With this definition of parity, Lemma 2.5.4 and its consequences also apply to the grid. Then, if N is odd $\mathscr{M}$ is ergodic, and if N is even there are 2^{w-1} closed classes of states, where each class consists on all configurations with the same parity. The Markov chain restricted to any of these classes of states is irreducible, positive recurrent, but 2-periodic so we don't have ergodicity.

Observation 2.6.8. Using the same argument as in Observation 2.5.5, for any fixed t, we can consider $\mathcal{V}_t$ as a static $\mathcal{V}$.

In view of this observation and Theorem 2.6.6, we assume $\mu = \Theta(1)$ for the remaining of the subsection. This covers the non-trivial dynamic situations where $G(\mathcal{V})$ is neither a.a.s. disconnected nor a.a.s. connected. Moreover in this subsection, we exclude the analysis for the case $d = 1$, for technical reasons. Hence, assume hereinafter that $d \geq 2$.

Let us first focus on the study of simple components. We define $X = X_t$ to be the random variable that counts the number of simple components at time step t. Given our assumptions about μ, for t in any fixed bounded time interval, X_t is asymptotically Poisson with expectation $\mu = \Theta(1)$, as studied in Proposition 2.6.2.

In analogy with d-hole lines in Subsection 2.5.2, we define a *simple component line* to be a maximal sequence of pairs $(v_1, t_1), \ldots, (v_l, t_l)$ where v_i is a simple component existing at time step t_i for $1 \leq i \leq l$, and such that $t_i = t_{i-1} + 1$ and the vertex v_i is adjacent to v_{i-1}, for $2 \leq i \leq l$. Birth, death and survival of lines, and the random variables B_t, D_t and S_t are defined analogously to the cycle case.

The following results will involve two consecutive time steps t and $t+1$. As discussed in Section 2.4, this can be modelled as an assignment of walkers into arcs in $A = A(T_N)$. We need some definitions: For each arc $e = (v_1, v_2) \in A$, define $\pi_1(e) = v_1$ and $\pi_2(e) = v_2$. Note that for any set $\mathcal{S}$ of vertices $\mathsf{Size}(\pi_1^{-1}(\mathcal{S})) = \mathsf{Size}(\pi_2^{-1}(\mathcal{S})) = \mathsf{Size}(\mathcal{S})$. Given any $v \in V$, let $\mathcal{H}_v$ be the set of vertices at distance between 1 and d from v. Observe that $\mathsf{Size}(\mathcal{H})_v = h$. In addition, define $\widehat{\mathcal{H}}_v = \pi_1^{-1}(\mathcal{H}_v)$ and $\widehat{\mathcal{H}}'_v = \pi_2^{-1}(\mathcal{H}_v)$, i.e. the sets of arcs with origin/destination in $\mathcal{H}_v$. Finally, for each arc $e = (v_1, v_2) \in A$, let $\widehat{\mathcal{B}}_e = \widehat{\mathcal{H}}'_{v_2} \setminus \widehat{\mathcal{H}}_{v_1}$ and $\widehat{\mathcal{B}}'_e = \widehat{\mathcal{H}}_{v_1} \setminus \widehat{\mathcal{H}}'_{v_2}$. By symmetry, we have that $\mathsf{Size}(\widehat{\mathcal{B}}_e) = \mathsf{Size}(\widehat{\mathcal{B}}'_e)$ and this quantity does not depend on the particular e, so let us denote it by b. In other words, b is the number of arcs whose origin is a vertex at distance strictly greater than d from v_1 and whose destination is at distance at most d from v_2. This new parameter turns out to play an important role in the characterisation of the dynamic properties of the graph of walkers. We have that $b = \Theta(d)$, but the exact expression of this b depends on the particular chosen metrics. Some examples are found in Table 2.3.

Metrics	b
ℓ^1	$b = 2d + 1/2$
ℓ^p $(p < \infty)$	$b \sim \left(1/\sqrt[p]{2} + 3/2\right) d$, if $d \to \infty$
ℓ^∞	$b = 3d + 1$

Table 2.3: Parameter b.

Here is a technical result

Lemma 2.6.9. *There exists $\epsilon > 0$ such that for any $v_1, v_2 \in V(T_N)$ with $\mathrm{dist}(v_1, v_2) > d-2$ we have $\mathsf{Size}(\widehat{\mathcal{H}}_{v_1} \cup \widehat{\mathcal{H}}_{v_2}) \geq (1+\epsilon)h$, $\mathsf{Size}(\widehat{\mathcal{H}}_{v_1} \cup \widehat{\mathcal{H}}'_{v_2}) \geq (1+\epsilon)h$ and $\mathsf{Size}(\widehat{\mathcal{H}}'_{v_1} \cup \widehat{\mathcal{H}}'_{v_2}) \geq (1+\epsilon)h$.*

Proof. We prove $\mathsf{Size}(\widehat{\mathcal{H}}_{v_1} \cup \widehat{\mathcal{H}}'_{v_2}) \geq (1+\epsilon)h$. The other two bounds are verified analogously. Suppose first that $d > R$, for some large enough but fixed R. Let $\mathcal{S}$ be set of vertices in $\mathcal{H}_{v_2}$ which are at distance at least $d+2$ from v_1. Observe that $\mathsf{Size}(\mathcal{S}) \geq h/4$ and that $\pi_2^{-1}(\mathcal{S})$ and $\widehat{\mathcal{H}}_{v_1}$ are disjoint. Then $\mathsf{Size}(\widehat{\mathcal{H}}_{v_1} \cup \widehat{\mathcal{H}}_{v_2}) \geq \mathsf{Size}(\widehat{\mathcal{H}}_{v_1} \cup \pi_2^{-1}(\mathcal{S})) = \mathsf{Size}(\widehat{\mathcal{H}}_{v_1}) + \mathsf{Size}(\mathcal{S}) \geq 5h/4$.

Otherwise suppose that $d \leq R$. Then $h \leq 10R^2$ for any ℓ^p distance we are considering. Recall that $d \geq 2$ so we can just guarantee that $\mathrm{dist}(v_1, v_2) > 0$. Let v_3 be the vertex in

$\mathcal{H}_{v_2}$ which is at greatest distance from v_1. At least one arc in $\pi_2^{-1}(v_3)$ must be in $\widehat{\mathcal{H}}'_{v_2} \setminus \widehat{\mathcal{H}}_{v_1}$. Then $\mathsf{Size}(\widehat{\mathcal{H}}_{v_1} \cup \widehat{\mathcal{H}}_{v_2}) \geq h+1 \geq (1+1/(10R^2))h$. □

To state the following result, we need one more definition: Given a collection of events $\mathcal{E}_1(N),\ldots,\mathcal{E}_k(N)$ and of random variables $W_1(N),\ldots,W_l(N)$ taking values in $\mathbb{N}$, with k and l fixed, we say that they are mutually asymptotically independent if for any $k', l', i_1, \ldots, i_{k'}, j_1, \ldots, j_{l'}, w_1, \ldots, w_{l'} \in \mathbb{N}$ such that $k' \leq k$, $l' \leq l$, $1 \leq i_1 < \cdots < i_{k'} \leq k$, $1 \leq j_1 < \cdots < j_{l'} \leq l$ we have that

$$\mathbf{P}\left(\bigwedge_{a=1}^{k'} \mathcal{E}_{i_a} \wedge \bigwedge_{b=1}^{l'} (W_{j_b} = w_b)\right) \sim \prod_{a=1}^{k'} \mathbf{P}(\mathcal{E}_{i_a}) \prod_{b=1}^{l'} \mathbf{P}(W_{j_b} = w_b). \tag{2.34}$$

We next have a result analogous to Proposition 2.5.6.

Proposition 2.6.10. *Assume $\mu = \Theta(1)$. Then for any two consecutive steps,*

$$\mathbf{E}S_t \sim \begin{cases} \mu & \text{if } d\varrho = o(1), \\ \mu e^{-b\varrho} & \text{if } d\varrho = \Theta(1), \\ 4\frac{1-e^{-\varrho/4}}{1-e^{-\varrho}} e^{-(b+3/4)\varrho}\mu & \text{if } d\varrho = \omega(1), \end{cases}$$

$$\mathbf{E}B_t = \mathbf{E}D_t \sim \begin{cases} b\varrho\mu & \text{if } d\varrho = o(1), \\ \mu\left(1 - e^{-b\varrho}\right) & \text{if } d\varrho = \Theta(1), \\ \mu & \text{if } d\varrho = \omega(1). \end{cases}$$

Moreover we have that

(i). If $d\varrho = o(1)$, then $\mathbf{P}(B > 0) \sim \mathbf{E}B$; $\mathbf{P}(D > 0) \sim \mathbf{E}D$; S is asymptotically Poisson; and $(B > 0)$, $(D > 0)$ and S are asymptotically mutually independent.

(ii). If $d\varrho = \Theta(1)$, then B, D and S are asymptotically mutually independent Poisson.

(iii). If $d\varrho = \omega(1)$, then B and D are asymptotically Poisson; $\mathbf{P}(S > 0) \sim \mathbf{E}S$; and B, D and $(S > 0)$ are asymptotically mutually independent.

Proof. The central ingredient in the proof is the computation of the joint factorial moments $\mathbf{E}([S]_{\ell_1}[B]_{\ell_2}[D]_{\ell_3})$ of these variables. In particular we find the asymptotic values of $\mathbf{E}S$, $\mathbf{E}B$ and $\mathbf{E}D$. Moreover, In the case $d\varrho = \Theta(1)$, we show that for any fixed naturals ℓ_1, ℓ_2 and ℓ_3 we have

$$\mathbf{E}([S]_{\ell_1}[B]_{\ell_2}[D]_{\ell_3}) \sim (\mathbf{E}S)^{\ell_1}(\mathbf{E}B)^{\ell_2}(\mathbf{E}D)^{\ell_3}. \tag{2.35}$$

Then, the result follows from Theorem 1.23 in [15]. The other cases are more delicate since (2.35) does not always hold for extreme values of the parameters, and we obtain a weaker result. In the case $d\varrho = o(1)$, we compute the moments for any natural ℓ_1 but only for $\ell_2, \ell_3 \in \{0,1,2\}$ and obtain

$$\begin{aligned} &\mathbf{E}([S]_{\ell_1}[B]_{\ell_2}[D]_{\ell_3}) \sim (\mathbf{E}S)^{\ell_1}(\mathbf{E}B)^{\ell_2}(\mathbf{E}D)^{\ell_3}, \quad \text{if } \ell_2, \ell_3 < 2, \\ &\mathbf{E}([S]_{\ell_1}[B]_2[D]_{\ell_3}) = o(\mathbf{E}([S]_{\ell_1}B\,[D]_{\ell_3})), \\ &\mathbf{E}([S]_{\ell_1}[B]_{\ell_2}[D]_2) = o(\mathbf{E}([S]_{\ell_1}[B]_2 D)). \end{aligned} \tag{2.36}$$

From this and by using upper and lower bounds given in [15], Section 1.4, applied to several variables, we deduce that S, $(B > 0)$ and $(D > 0)$ satisfy (2.34) and also

$$\mathbf{P}(S = k) \sim e^{-\mathbf{E}S} \frac{(\mathbf{E}S)^k}{k!} \quad \forall k \in \mathbb{N}, \quad \mathbf{P}(B > 0) \sim \mathbf{E}B \quad \text{and} \quad \mathbf{P}(D > 0) \sim \mathbf{E}D.$$

Similarly, in the case $d\varrho = \omega(1)$, we compute the moments for any naturals ℓ_2 and ℓ_3 but only for $\ell_1 \in \{0, 1, 2\}$ and obtain

$$\begin{aligned} &\mathbf{E}([S]_{\ell_1}[B]_{\ell_2}[D]_{\ell_3}) \sim (\mathbf{E}S)^{\ell_1}(\mathbf{E}B)^{\ell_2}(\mathbf{E}D)^{\ell_3}, \quad \text{if } \ell_1 < 2, \\ &\mathbf{E}([S]_2[B]_{\ell_2}[D]_{\ell_3}) = o(\mathbf{E}(S\,[B]_{\ell_2}[D]_{\ell_3})). \end{aligned} \tag{2.37}$$

From this and by using once more upper and lower bounds given in Section 1.4 of [15], we conclude that $(S > 0)$, B and D satisfy (2.34) and also

$$\mathbf{P}(S > 0) \sim \mathbf{E}S, \quad \mathbf{P}(B = k) \sim e^{-\mathbf{E}B} \frac{(\mathbf{E}B)^k}{k!} \quad \forall k \in \mathbb{N}$$
$$\text{and} \quad \mathbf{P}(D = k) \sim e^{-\mathbf{E}D} \frac{(\mathbf{E}D)^k}{k!} \quad \forall k \in \mathbb{N}.$$

In order to compute the moments, we first describe the survivals, births and deaths of simple components from a static point of view, in terms of occupancy of some regions (sets) of arcs in $A(T_N)$ by walkers. A summary of these descriptions is given in Table 2.4. Recall that we are assuming that $d \geq 2$. The case $d = 1$ is slightly different since one simple component can split into four simple components in one step. This case is not covered here.

Let us first deal with survivals. A simple component on vertex v survives between time steps t and $t+1$ iff exactly one arc $e = (v, v')$ in $\pi_1^{-1}(v)$ is occupied and $\widehat{\mathcal{H}}_v \cup \widehat{\mathcal{H}}'_{v'}$ is e.o.w. (as shown in Figure 2.9). Therefore, there are four ways to achieve this, one for each choice of e.

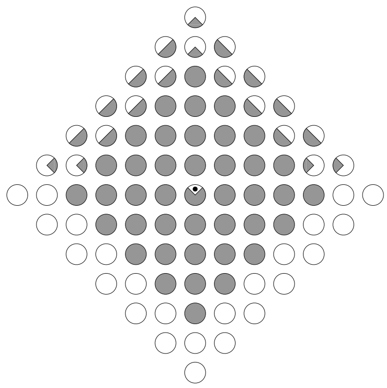


Figure 2.9: Survival of a simple component line at vertex v

There are 15 ways that a simple component line can be born at v between time steps t and $t+1$. We classify these events in four main classes, as shown in Figure 2.10.

b1 There are four births which correspond to this class: Exactly one arc $e' = (u, v)$ in $\pi_2^{-1}(v)$ is occupied, $\widehat{\mathcal{B}}'_{e'}$ is also occupied, and $\widehat{\mathcal{H}}'_v$ is e.o.w.

b2 There are six births which correspond to this class: Exactly two arcs in $\pi_2^{-1}(v)$ are occupied, and $\widehat{\mathcal{H}}'_v$ is e.o.w.

b3 There are four births which correspond to this class: Exactly three arcs in $\pi_2^{-1}(v)$ are occupied, and $\widehat{\mathcal{H}}'_v$ is e.o.w.

b2 There is one birth which corresponds to this class: All arcs in $\pi_2^{-1}(v)$ are occupied, and $\widehat{\mathcal{H}}'_v$ is e.o.w.

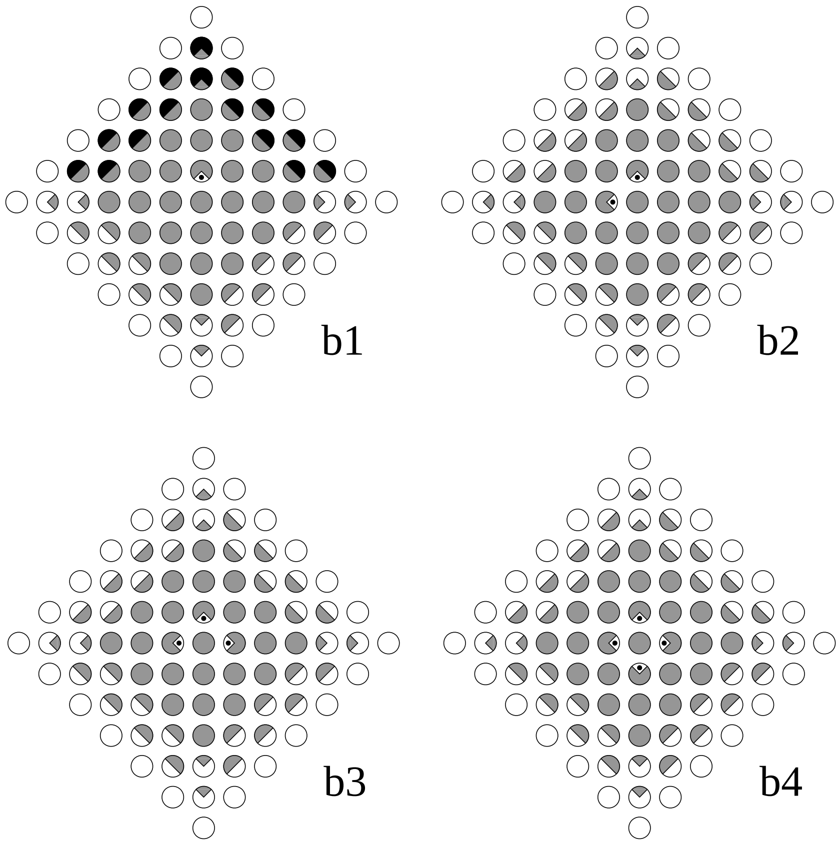


Figure 2.10: Birth of a simple component line at vertex v

Similarly, there are 15 events (classified in four classes) leading to the destruction of a simple component line at v, according to the following descriptions and as shown in Figure 2.11:

d1 There are four destructions which correspond to this class: Exactly one arc $e = (v, u)$ in $\pi_1^{-1}(v)$ is occupied, $\widehat{\mathcal{B}}_e$ is also occupied, and $\widehat{\mathcal{H}}_v$ is e.o.w.

d2 There are six destructions which correspond to this class: Exactly two arcs in $\pi_1^{-1}(v)$ are occupied, and $\widehat{\mathcal{H}}_v$ is e.o.w.

d3 There are four destructions which correspond to this class: Exactly three arcs in $\pi_1^{-1}(v)$ are occupied, and $\widehat{\mathcal{H}}_v$ is e.o.w.

d2 There is one destruction which corresponds to this class: All four arcs in $\pi_1^{-1}(v)$ are occupied, and $\widehat{\mathcal{H}}_v$ is e.o.w.

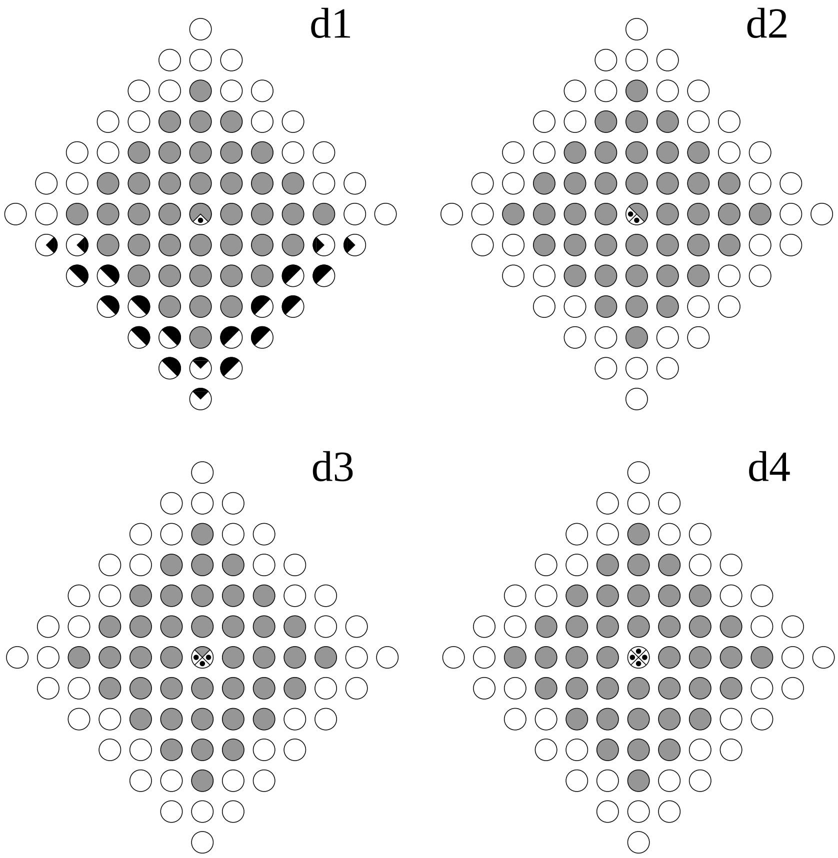


Figure 2.11: Destruction of a simple component line at vertex v

Given any $v \in V(T_N)$, for each of the four possible ways that a simple component line can survive at vertex v, we define the corresponding indicator random variable S_v^α, $\alpha \in \{1, \ldots, 4\}$. Similarly, for each of the 15 possible ways that a simple component line can be born (die) at vertex v, we define the corresponding indicator random variable B_v^α (D_v^α), $\alpha \in \{1, \ldots, 15\}$. Hence, $S_v = \sum_{\alpha=1}^4 S_v^\alpha$, $B_v = \sum_{\alpha=1}^{15} B_v^\alpha$ and $D_v = \sum_{\alpha=1}^{15} D_v^\alpha$ are the indicator variables for a survival, birth and death, respectively, at vertex v.

Fix any naturals ℓ_1, ℓ_2 and ℓ_3, and call $\ell = \ell_1 + \ell_2 + \ell_3$. Let $\boldsymbol{\mathcal{A}} = \big(\{1, \ldots, 4\}\big)^{\ell_1} \times \big(\{1, \ldots, 15\}\big)^{\ell_2+\ell_3}$, and let $\boldsymbol{\mathcal{T}}$ be the set of all ℓ-tuples of different vertices in $V(T_N)$. Given $\boldsymbol{\alpha} = (\alpha_i)_{i=1}^{\ell} \in \boldsymbol{\mathcal{A}}$, and $\boldsymbol{v} = (v_i)_{i=1}^{\ell} \in \boldsymbol{\mathcal{T}}$ let us define the event

$$\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}} = \Bigg(\bigwedge_{i=1}^{\ell_1} (S_{v_i}^{\alpha_i} = 1)\Bigg) \wedge \Bigg(\bigwedge_{i=\ell_1+1}^{\ell_1+\ell_2} (B_{v_i}^{\alpha_i} = 1)\Bigg) \wedge \Bigg(\bigwedge_{i=\ell_1+\ell_2+1}^{\ell} (D_{v_i}^{\alpha_i} = 1)\Bigg). \tag{2.38}$$

This allows us to express the joint factorial moments as

$$\mathbf{E}([S]_{\ell_1}[B]_{\ell_2}[D]_{\ell_3}) = \sum_{\boldsymbol{v}\in\boldsymbol{\mathcal{T}}} \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{A}}} \mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}). \tag{2.39}$$

	Ways	Size empty region	Sizes non-empty regions
Survival	4	$h+b+3/4$	$1/4$
Birth b1	4	$h+3/4$	$1/4$, b
Birth b2	6	$h+1/2$	$2\times 1/4$
Birth b3	4	$h+1/4$	$3\times 1/4$
Birth b4	1	h	$4\times 1/4$
Death d1	4	$h+3/4$	$1/4$, b
Death d2	6	$h+1/2$	$2\times 1/4$
Death d3	4	$h+1/4$	$3\times 1/4$
Death d4	1	h	$4\times 1/4$

Table 2.4: Event descriptions according to their occupancy requirements

In order to compute $\mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}})$, we partition the set of tuples $\boldsymbol{\mathcal{T}}$ into three disjoint classes: Let $\boldsymbol{\mathcal{T}}_2$ be the set of tuples $\boldsymbol{v}\in\boldsymbol{\mathcal{T}}$ such that all pairs of vertices in $\boldsymbol{v}$ are at distance greater than $2d+4$; Let $\boldsymbol{\mathcal{T}}_1$ be the set of tuples $\boldsymbol{v}\in\boldsymbol{\mathcal{T}}\setminus\boldsymbol{\mathcal{T}}_2$ such that all pairs of vertices in $\boldsymbol{v}$ are at distance greater than $d-2$; Finally, define $\boldsymbol{\mathcal{T}}_0=\boldsymbol{\mathcal{T}}\setminus(\boldsymbol{\mathcal{T}}_1\cup\boldsymbol{\mathcal{T}}_2)$.

First observe that if $\boldsymbol{v}\in\boldsymbol{\mathcal{T}}_0$ then $\mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}})=0$, since some pair of different vertices in $\boldsymbol{v}$ are at distance at most $d-2$ and this is not compatible with $\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}$. Now given any $\boldsymbol{v}\in\boldsymbol{\mathcal{T}}_2$, notice that the regions involved in the descriptions of the events $(S_{v_i}^{\alpha_i}=1)$, $(B_{v_i}^{\alpha_i}=1)$ and $(D_{v_i}^{\alpha_i}=1)$ are disjoint for any choice of $\boldsymbol{\alpha}$. This allows us to compute $\mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}})$ by applying Lemma 2.4.2 to these regions, whose sizes are listed in Table 2.4. For each $j\in\{1,\dots,4\}$, let a_j be the number of entries α_i of $\boldsymbol{\alpha}$ with $\ell_1+1\le i\le\ell$ which correspond to births of class bj or to deaths of class dj. Observe that $\mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}})$ does not depend on the particular $\boldsymbol{v}\in\boldsymbol{\mathcal{T}}_2$ or on the order of the entries of $\boldsymbol{\alpha}$, but only on $\boldsymbol{a}=(a_1,a_2,a_3,a_4)$. Hence we can denote this probability by $P_{\boldsymbol{a}}$, and it satisfies

$$\begin{aligned}P_{\boldsymbol{a}}\sim e^{-h\varrho\ell}&\left[\left(1-e^{-\varrho/4}\right)e^{-(b+3/4)\varrho}\right]^{\ell_1}\\&\left[\left(1-e^{-\varrho/4}\right)\left(1-e^{-b\varrho}\right)e^{-3\varrho/4}\right]^{a_1}\left[\left(1-e^{-\varrho/4}\right)^2e^{-\varrho/2}\right]^{a_2}\\&\left[\left(1-e^{-\varrho/4}\right)^3e^{-\varrho/4}\right]^{a_3}\left[\left(1-e^{-\varrho/4}\right)^4\right]^{a_4}.\end{aligned}\tag{2.40}$$

From this and also by using

$$\sum_{\boldsymbol{\alpha}}\mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}})=\sum_{a_1+a_2+a_3+a_4=\ell_2+\ell_3}\binom{\ell_2+\ell_3}{a_1,a_2,a_3,a_4}4^{q+a_1+a_3}6^{a_2}P_{\boldsymbol{a}},$$

we obtain the contribution to $\mathbf{E}([S]_{\ell_1}[B]_{\ell_2}[D]_{\ell_3})$ due to tuples in $\boldsymbol{\mathcal{T}}_2$

$$\begin{aligned}\sum_{\boldsymbol{v}\in\boldsymbol{\mathcal{T}}_2}\sum_{\boldsymbol{\alpha}}\mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}})\sim\left(\left(1-e^{-\varrho/4}\right)Ne^{-h\varrho}\right)^{\ell}&\left[4e^{-(b+3/4)\varrho}\right]^{\ell_1}\\&\left[4\left(1-e^{-b\varrho}\right)e^{-3\varrho/4}+6\left(1-e^{-\varrho/4}\right)e^{-\varrho/2}+\right.\end{aligned}$$

$$4\left(1-e^{-\varrho/4}\right)^2 e^{-\varrho/4} + \left(1-e^{-\varrho/4}\right)^3\Bigg]^{\ell_2+\ell_3}. \qquad (2.41)$$

Recall that the assumption $\mu = \Theta(1)$ implies $h\varrho \to \infty$. Thus, in the cases when $\varrho = O(1)$ we must have $d \to \infty$. Taking this into account, we get an equivalent but simpler asymptotic expression to that in (2.41) by considering separately the cases in which ϱ and $d\varrho$ are $o(1)$, $\Theta(1)$ or $\omega(1)$.

$$\sum_{\boldsymbol{v}\in\boldsymbol{\mathcal{T}}_2}\sum_{\boldsymbol{\alpha}} \mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}) \sim \left[4\frac{1-e^{-\varrho/4}}{1-e^{-\varrho}}e^{-(b+3/4)\varrho}\mu\right]^{\ell_1}\left[\left(1-e^{-b\varrho}\right)\mu\right]^{\ell_2+\ell_3}. \qquad (2.42)$$

It just remains to bound the weight in $\mathbf{E}([S]_{\ell_1}[B]_{\ell_2}[D]_{\ell_3})$ due to tuples in $\boldsymbol{\mathcal{T}}_1$. This contribution is negligible in some cases, but unfortunately it may be larger than (2.42) for some extreme values of the parameters. We analyse separate cases according to the asymptotic behaviour of $d\varrho$.

Case 1 ($d\varrho = \Theta(1)$). Since $\varrho = o(1)$, from (2.42) we can write

$$\sum_{\boldsymbol{v}\in\boldsymbol{\mathcal{T}}_2}\sum_{\boldsymbol{\alpha}} \mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}) \sim \left[e^{-b\varrho}\mu\right]^{\ell_1}\left[\left(1-e^{-b\varrho}\right)\mu\right]^{\ell_2+\ell_3}. \qquad (2.43)$$

Let us fix a tuple $\boldsymbol{v} \in \boldsymbol{\mathcal{T}}_1$ and also $\boldsymbol{\alpha} \in \boldsymbol{\mathcal{A}}$. For each i, $1 \le i \le \ell_1$, let $e_i = (v_i, v_i')$ be the arc involved in the description of $S_{v_i}^{\alpha_i}$. Then consider the set of arcs

$$\widehat{\mathcal{H}} = \left(\bigcup_{i=1}^{\ell_1}(\widehat{\mathcal{H}}_{v_i} \cup \widehat{\mathcal{H}}'_{v_i'})\right) \cup \left(\bigcup_{i=\ell_1+1}^{\ell_1+\ell_2}\widehat{\mathcal{H}}'_{v_i}\right) \wedge \left(\bigwedge_{i=\ell_1+\ell_2+1}^{\ell}\widehat{\mathcal{H}}_{v_i}\right).$$

Note that $\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}$ implies that $\widehat{\mathcal{H}}$ is e.o.w. Unfortunately since $\boldsymbol{v} \in \boldsymbol{\mathcal{T}}_1$ the sets of arcs involved in the definition of $\widehat{\mathcal{H}}$ may not be disjoint. In order to bound $\mathsf{Size}(\widehat{\mathcal{H}})$ we need one definition: We say that a given vertex v_i of $\boldsymbol{v}$ is restricted if there is some other v_j of $\boldsymbol{v}$ with $j < i$ and such that $\mathrm{dist}(v_i, v_j) \le 2d + 4$. Let r be the number of restricted vertices of the tuple $\boldsymbol{v}$, and observe that $r > 0$ since $\boldsymbol{v} \in \boldsymbol{\mathcal{T}}_1$. Then $\mathsf{Size}(\widehat{\mathcal{H}}) \ge (\ell - r + \epsilon)h$, since each unrestricted vertex contributes h to $\widehat{\mathcal{H}}$ and the first restricted one gives the term ϵh (see Lemma 2.6.9). Therefore

$$\mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}) \le (1-\mathsf{Size}(\widehat{\mathcal{H}}))^w\left(1-e^{-\varrho/4}\right)^\ell = O(e^{-(\ell-r+\epsilon)h\varrho}\varrho^\ell) = O\left(\frac{w^{r-\epsilon}}{N^\ell}\right).$$

But the number of choices for $\boldsymbol{v}$ with r restricted vertices is $O(N^{\ell-r}h^r)$, and this contributes to $\mathbf{E}([S]_{\ell_1}[B]_{\ell_2}[D]_{\ell_3})$ only $O((h\varrho)^r/w^\epsilon) = O(\log^r w/w^\epsilon)$, which is negligible compared to (2.43).

Case 2 ($d\varrho = o(1)$). From (2.42) we can write

$$\sum_{\boldsymbol{v}\in\boldsymbol{\mathcal{T}}_2}\sum_{\boldsymbol{\alpha}} \mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}) \sim \mu^{\ell_1}(b\varrho\mu)^{\ell_2+\ell_3}. \qquad (2.44)$$

Let us fix a tuple $\boldsymbol{v} \in \boldsymbol{\mathcal{T}}_1$ and also $\boldsymbol{\alpha} \in \boldsymbol{\mathcal{A}}$. Define $\widehat{\mathcal{H}}$ and *restricted* vertices as in the case $d\varrho = \Theta(1)$, and let $r > 0$ be the number of restricted vertices in $\boldsymbol{v}$. Let $\boldsymbol{v}_1$ be the set of vertices in $\boldsymbol{v}$ which correspond to a birth or death of class b1 or d1 in the description of $\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}$, and call $\boldsymbol{v}_2$ to the set of the remaining vertices in $\boldsymbol{v}$ not in $\boldsymbol{v}_1$. Recall that $|\boldsymbol{v}_1| = a_1 \leq \ell_2+\ell_3$. For each vertex $v_i \in \boldsymbol{v}_1$ which must undergo a birth according to $\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}$, let $e_i = (u_i, v_i)$ be the arc involved in the description of $B_{v_i}^{\alpha_i}$ and call $\widehat{\mathcal{B}}_{v_i}^* = \widehat{\mathcal{B}}'_{e_i}$. Similarly for each vertex $v_i \in \boldsymbol{v}_1$ which corresponds to a death, let $e_i = (v_i, u_i)$ be the arc involved in the description of $D_{v_i}^{\alpha_i}$ and call $\widehat{\mathcal{B}}_{v_i}^* = \widehat{\mathcal{B}}_{e_i}$. Note that $\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}$ requires that $\widehat{\mathcal{B}}_{v_i}^*$ and e_i are occupied for all $v_i \in \boldsymbol{v}_1$. In order to describe the different ways to achieve this, we need some definitions: Given some vertices $v_{i_1}, \dots, v_{i_k}$ in $\boldsymbol{v}_1$, we say that they *interfere* if $\bigwedge_{j=1}^{k} \widehat{\mathcal{B}}_{v_{i_j}}^* \neq \emptyset$. Obviously, they must be *restricted* except possibly for the one with smallest index. If $e_j \in \widehat{\mathcal{B}}_{v_i}^*$ for some pair of vertices $v_i, v_j \in \boldsymbol{v}_1$ we say that v_i and v_j *collaborate*. Notice that for v_i, v_j to collaborate, either $\ell_1 + 1 \leq i, j \leq \ell_1 + \ell_2$ or $\ell_1 + \ell_2 + 1 \leq i, j \leq \ell$. Moreover the one with highest index is restricted. We first suppose that for our choice of $\boldsymbol{v}$ and $\boldsymbol{\alpha}$ there are no vertices in $\boldsymbol{v}_1$ which collaborate, and study only how $\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}$ is affected by interferences of vertices. Let $\mathscr{P} = \mathscr{P}(\boldsymbol{\alpha}, \boldsymbol{v})$ be the set of all partitions of $\boldsymbol{v}_1$ into disjoint sets (called *blocks*) with the following property: If $\mathcal{P} \in \mathscr{P}$ then the vertices of each block of $\mathcal{P}$ interfere. Given $\mathcal{P} = \{P_1, \dots, P_k\} \in \mathscr{P}$ and $f_1, \dots, f_k$ different arcs in $A(T_N)$ such that $f_j \in \bigwedge_{v_i \in P_j} \widehat{\mathcal{B}}_{v_i}^*$, we define the following event $\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v},\mathcal{P},\boldsymbol{f}}$: All arcs in $\boldsymbol{f} = (f_1, \dots, f_k)$ are occupied, and moreover $\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}$ holds. Observe that $\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}$ implies that $\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v},\mathcal{P},\boldsymbol{f}}$ holds for some $\mathcal{P}$ and some $\boldsymbol{f}$, so we wish to compute $\mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v},\mathcal{P},\boldsymbol{f}})$. Let us call *leader* to the vertex in each block P_j with smallest index, and let r' be the number of restricted vertices which are either leaders of some block or belong to $\boldsymbol{v} \setminus \boldsymbol{v}_1$. Then there are $r = a_1 - k + r' > 0$ restricted vertices and, by the same argument as in the case $d\varrho = \Theta(1)$, we deduce that $\mathsf{Size}(\widehat{\mathcal{H}}) \geq (\ell - r + \epsilon)h$ (see Lemma 2.6.9). Therefore,

$$\mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v},\mathcal{P},\boldsymbol{f}}) = O(e^{-(\ell-r+\epsilon)h\varrho}\varrho^{\ell+k+a_2+2a_3+3a_4}) = O\left(\frac{w^{r-\epsilon}}{N^\ell}\varrho^{k+a_2+2a_3+3a_4}\right).$$

Observe that if the values of $\boldsymbol{a} = (a_1, a_2, a_3, a_4)$, r and k are given, the number of possible choices for $\boldsymbol{\alpha}, \boldsymbol{v}, \mathcal{P}, \boldsymbol{f}$ is $O(N^{\ell-r}h^{r'}b^{a_1})$. Hence the contribution to $\mathbf{E}([S]_{\ell_1}[B]_{\ell_2}[D]_{\ell_3})$ due to tuples $\boldsymbol{v} \in \boldsymbol{\mathcal{T}}_1$ in which no pair of vertices collaborates and with these particular $\boldsymbol{a}$, r and k is

$$O(N^{\ell-r}h^{r'}b^{a_1})O\left(\frac{w^{r-\epsilon}}{N^\ell}\varrho^{k+a_2+2a_3+3a_4}\right) = O\left(\frac{\log^{r'} w}{w^\epsilon}(b\varrho)^{a_1}\varrho^{a_2+2a_3+3a_4}\right),$$

negligible compared to (2.44). In particular, if $\ell_2, \ell_3 < 2$, then no pair of vertices in $\boldsymbol{v}_1$ can collaborate and the first line in (2.36) follows from (2.39) and (2.44).

Finally we must deal with tuples $\boldsymbol{v} \in \boldsymbol{\mathcal{T}}_1$ in which some pairs of vertices collaborate. Unfortunately, the weight in (2.39) due to these terms is larger than (2.42) when $d\varrho$ tends to 0 fast. Hence we restrict ℓ_2 and ℓ_3 to be at most 2 and prove only (2.36). Suppose that $\ell_2 = 2$ and that $\boldsymbol{v} \in \boldsymbol{\mathcal{T}}_1$ is such that v_{ℓ_1+1} and v_{ℓ_1+2} is the only pair which collaborates. Then at least v_{ℓ_1+2} must be restricted. By repeating the same argument as above, but considering partitions of $\boldsymbol{v}_1 \setminus \{v_{\ell_1+1}, v_{\ell_1+2}\}$ rather than $\boldsymbol{v}_1$ and taking into account that the number of choices for v_i and v_j is $O(bN)$, we deduce that the contribution of these tuples to (2.39) is

$$O\left(\frac{(b\varrho)^{1+\ell_3}}{w^{\epsilon'}}\right) = o\left((b\varrho)^{1+\ell_3}\right).$$

Similarly, if $\ell_3 = 2$ the weight in (2.39) due to tuples $\boldsymbol{v} \in \boldsymbol{\mathcal{T}}_1$ in which $v_{\ell_1+\ell_2+1}$ and v_ℓ is the only pair which collaborates is $o\big((b\varrho)^{\ell_2+1}\big)$. Finally if $\ell_2 = \ell_3 = 2$, the tuples $\boldsymbol{v} \in \boldsymbol{\mathcal{T}}_1$ in which both pairs $v_{\ell_1+1}, v_{\ell_1+2}$ and $v_{\ell_1+3}, v_{\ell_1+4}$ collaborate contribute $o\big((b\varrho)^2\big)$ to (2.39). In view of the above and also recalling the contribution to (2.39) due to tuples in $\boldsymbol{\mathcal{T}}_2$ and tuples in $\boldsymbol{\mathcal{T}}_1$ with no collaborations, we obtain (2.36) as announced.

Case 3 ($d\varrho = \omega(1)$). From (2.42) we can write

$$\sum_{\boldsymbol{v}\in\boldsymbol{\mathcal{T}}_2} \sum_{\boldsymbol{\alpha}} \mathbf{P}(\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}) \sim \left[4\frac{1-e^{-\varrho/4}}{1-e^{-\varrho}} e^{-(b+3/4)\varrho}\mu\right]^{\ell_1} \mu^{\ell_2+\ell_3}. \tag{2.45}$$

Let us fix a tuple $\boldsymbol{v} \in \boldsymbol{\mathcal{T}}_1$ and also $\boldsymbol{\alpha} \in \boldsymbol{\mathcal{A}}$. Define $\widehat{\mathcal{H}}$ and *restricted* vertices as in the case $d\varrho = \Theta(1)$, and let $r > 0$ be the number of restricted vertices in $\boldsymbol{v}$.

Suppose first that $\ell_1 \le 1$. In this case, the only possible vertex in $\boldsymbol{v}$ which involves a survival in the description of $\mathcal{E}_{\boldsymbol{\alpha},\boldsymbol{v}}$ cannot be restricted by definition, since it has the lower index. Then $\mathsf{Size}(\widehat{\mathcal{H}}) \ge (\ell - r + \epsilon)h + \ell_1(b+3/4)$, since the $\ell - r$ unrestricted vertex contribute $(\ell - r)h + \ell_1(b+3/4)$ to $\widehat{\mathcal{H}}$ and the first restricted one gives the term ϵh (see Lemma 2.6.9). Therefore, proceeding as in the case $d\varrho = \Theta(1)$ but keeping an extra $e^{-\ell_1(b+3/4)\varrho}$ factor in the computations, we conclude that the contribution to (2.39) due to tuples in $\boldsymbol{\mathcal{T}}_1$ with r restricted vertices is $O(e^{-\ell_1(b+3/4)\varrho} \log^r w/w^\epsilon)$. This is negligible compared to (2.45). Hence the first line in (2.37) follows from (2.39) and (2.45).

Otherwise consider the case $\ell_1 = 2$. Then we can only guarantee that $\mathsf{Size}(\widehat{\mathcal{H}}) \ge (\ell - r + \epsilon)h + (b+3/4)$, since v_1 is not restricted but cannot say anything about v_2. Repeating again the same argument as in the case $d\varrho = \Theta(1)$, we conclude that the contribution to (2.39) in this case is $O(e^{-(b+3/4)\varrho} \log^r w/w^\epsilon)$. This is $o\big(\mathbf{E}(S\,[B]_{\ell_2}[D]_{\ell_3})\big)$, and the second line in (2.37) follows. □

Lemma 2.6.11. *Assume that* $\mu = \Theta(1)$ *and* $d\varrho = o(1)$*. Then,*

- $\mathbf{P}\big((Y_t + Z_t = 0) \wedge (Y_{t+1} + Z_{2,t+1} > 0)\big) = o(d\varrho)$,
- $\mathbf{P}\big((Y_t + Z_t > 0) \wedge (B_t > 0)\big) = o(d\varrho)$.

Proof. Rather than the first statement we prove its equivalent in the time-reversed process: $\mathbf{P}\big((Y_t + Z_t > 0) \wedge (Y_{t+1} + Z_{t+1} = 0)\big) = o(d\varrho)$.

Let $\mathcal{B}$ be the set of walks in T_N which are (for some configuration of the walkers): a maximal boundary walk of some embeddable component with respect to its outside region; or a boundary walk of length at least $n - o(n)$ of some non-embeddable component which is not solitary (possible by Lemma 2.6.4). For each $\beta \in \mathcal{B}$ let $\mathcal{A}_\beta$ be the forbidden region outside β. Fix any arbitrary vertices u and v such that $\operatorname{dist}(u,v) \in \{d-1, d, d+1, d+2\}$. Note that if u and v are occupied at time t this is a necessary condition for a creation or destruction of an edge involving those vertices.

Let $\mathcal{E}_1$ be the following event: at time t there is some walk $\beta \in \mathcal{B}$ containing u and v with all the vertices occupied, and moreover $\mathcal{A}_\beta$ is e.o.w. To compute $\mathbf{P}(\mathcal{E}_1)$ we extend the argument in the proof of Lemma 2.6.5 but taking into account that the position of u and v is fixed. We obtain

$$\mathbf{P}(\mathcal{E}_1) = O(1)\frac{\varrho}{N} e^{-d^2\varrho/(2R')} \sum_{m\ge 2} \left(\frac{C''}{d^2\varrho}\right)^{m-2} = O\left(\frac{\varrho}{N w^\epsilon}\right),$$

for some $\epsilon > 0$.

Let $\mathcal{E}_2$ be the following event: at time t there is some walk $\beta \in \mathcal{B}$ containing u but not v with all the vertices occupied, v is also occupied, and moreover $\mathcal{A}_\beta$ is e.o.w. We compute $\mathbf{P}(\mathcal{E}_2)$ again by extending the argument in the proof of Lemma 2.6.5 but taking into account that the position of u is fixed. We obtain

$$\mathbf{P}(\mathcal{E}_1) = O(1)\frac{\varrho}{N}\sum_{m\geq 2}\left(\frac{C''}{d^2\varrho}\right)^{m-1} = O\left(\frac{\varrho}{N\log w}\right).$$

Let R be the constant in the statement of Lemma 2.6.3. Let $\mathcal{E}_3$ be the following event: at time t there is some walk $\beta \in \mathcal{B}$ of length at least $3Rh/d$ not containing u and v with all the vertices occupied, u and v are also occupied, and moreover $\mathcal{A}_\beta$ is e.o.w. In view of Lemma 2.6.3, $\mathcal{A}_\beta \geq 2h + dl/3R$. Hence by repeating the argument in the proof of Lemma 2.6.5 but keeping an extra $e^{-h\varrho}$ factor, we deduce that

$$\mathbf{P}(\mathcal{E}_3) = O(1)\varrho^2 e^{-h\varrho}\sum_{m\geq 2}\left(\frac{C''}{d^2\varrho}\right)^{m-1} = O\left(\frac{\varrho}{N\log w}\right).$$

Let $\mathcal{E}_4$ be the following event: at time t there is some walk $\beta \in \mathcal{B}$ of length at most $3Rh/d$ with all the vertices occupied, u and v are at distance greater than $2d+4$ from the minimal rectangle containing β, u and v are also occupied, u is isolated, and moreover $\mathcal{A}_\beta$ is e.o.w. Observe that the conditions on β are asymptotically independent from those on u and v since when describing them in terms of occupancy of regions of vertices, the regions involved are pairwise disjoint. Hence

$$\mathbf{P}(\mathcal{E}_4) = O(1)\varrho^2 e^{-h\varrho}\sum_{m\geq 2}\left(\frac{C''}{d^2\varrho}\right)^{m-1} = O\left(\frac{\varrho}{N\log w}\right).$$

Let $\mathcal{E}_5$ be the following event: at time t there is some walk $\beta \in \mathcal{B}$ of length at most $3Rh/d$ of size at least 3 not containing u and v with all the vertices occupied, u and v are at inside or at distance at most $2d+4$ from the minimal rectangle containing β, u and v are also occupied, and moreover $\mathcal{A}_\beta$ is e.o.w. Note that if length$(\beta) \leq 3Rh/d$, this minimal rectangle has size $\Theta(d^2)$ and therefore

$$\mathbf{P}(\mathcal{E}_5) = O(1)\varrho^2\frac{d^2}{N}\sum_{m\geq 3}\left(\frac{C''}{d^2\varrho}\right)^{m-1} = O\left(\frac{\varrho}{N\log w}\right).$$

Let $\mathcal{E}_6$ be the following event: at time t there is some walk $\beta \in \mathcal{B}$ of size 2 not containing u and v at distance at most $2d+4$ from u with all the vertices occupied, u and v are also occupied, u is isolated, and moreover $\mathcal{A}_\beta$ is e.o.w. Observe that at least ϵh vertices in $\mathcal{H}_u$ do not intersect $\mathcal{A}_\beta$, for some $\epsilon > 0$ and must be e.o.w. Then We have

$$\mathbf{P}(\mathcal{E}_6) = O(1)\varrho^2 e^{-\epsilon h\varrho}\frac{d^2}{N}\left(\frac{C''}{d^2\varrho}\right) = O\left(\frac{\varrho}{Nw^\epsilon}\right).$$

Now observe that $(Y_t + Z_t > 0) \wedge (Y_{t+1} + Z_{t+1} = 0)$ implies that $\mathcal{E}_1$, $\mathcal{E}_2$, $\mathcal{E}_3$ or $\mathcal{E}_4$ holds for some u and v. Similarly $(Y_t + Z_t > 0) \wedge (B_t > 0)$ implies that $\mathcal{E}_1$, $\mathcal{E}_2$, $\mathcal{E}_3$, $\mathcal{E}_4$, $\mathcal{E}_5$ or $\mathcal{E}_6$ holds for some u and v. Hence both events have probability $O(Nd)o(\varrho/N) = o(d\varrho)$. □

From Proposition 2.6.10, we can easily derive important consequences analogous to those of the cycle, always under the assumption stated after Observation 2.6.8. The first one gives us the probability that $G(\mathcal{V}_t)$ is connected but $G(\mathcal{V}_{t+1})$ is disconnected. The proof is not so easy as that of Lemma 2.5.11. Let $\mathcal{C}_t$ and $\mathcal{D}_t$ the events that $G(\mathcal{V}_t)$ is respectively connected and disconnected.

Lemma 2.6.12. *Assume that* $\mu = \Theta(1)$. *Then,*

$$\mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1}) \sim e^{-\mu}(1 - e^{-\mathbf{E}B}), \qquad \mathbf{P}(\mathcal{D}_t \wedge \mathcal{C}_{t+1}) \sim e^{-\mu}(1 - e^{-\mathbf{E}B})$$
$$\mathbf{P}(\mathcal{C}_t \wedge \mathcal{C}_{t+1}) \sim e^{-\mu}e^{-\mathbf{E}B}, \qquad \mathbf{P}(\mathcal{D}_t \wedge \mathcal{D}_{t+1}) \sim 1 - 2e^{-\mu} + e^{-\mu}e^{-\mathbf{E}B}$$

Proof. First observe that $X_t = S_t + D_t$ and $X_{t+1} = S_t + B_t$. Therefore we have

$$\mathbf{P}(X_t = 0 \wedge X_{t+1} > 0) = \mathbf{P}(S_t = 0 \wedge D_t = 0 \wedge B_t > 0),$$

and by Proposition 2.6.10 we get

$$\mathbf{P}(X_t = 0 \wedge X_{t+1} > 0) \sim e^{-\mathbf{E}S - \mathbf{E}D}(1 - e^{-\mathbf{E}B}) \sim e^{-\mu}(1 - e^{-\mathbf{E}B}). \tag{2.46}$$

We want to connect this probability with $\mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1})$. In fact, by partitioning $(X_t = 0 \wedge X_{t+1} > 0)$ and $(\mathcal{C}_t \wedge \mathcal{D}_{t+1})$ into disjoint events, we obtain

$$\mathbf{P}(X_t = 0 \wedge X_{t+1} > 0) = \mathbf{P}(\mathcal{C}_t \wedge X_{t+1} > 0) + \mathbf{P}(\mathcal{D}_t \wedge X_t = 0 \wedge X_{t+1} > 0),$$
$$\mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1}) = \mathbf{P}(\mathcal{C}_t \wedge X_{t+1} > 0) + \mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1} \wedge X_{t+1} = 0),$$

and thus we can write

$$\mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1}) = \mathbf{P}(X_t = 0 \wedge X_{t+1} > 0) + P_1 - P_2, \tag{2.47}$$

where $P_1 = \mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1} \wedge X_{t+1} = 0)$ and $P_2 = \mathbf{P}(\mathcal{D}_t \wedge X_t = 0 \wedge X_{t+1} > 0)$.

Now suppose that $d\varrho = o(1)$. In that case, $\mathbf{P}(X_t = 0 \wedge X_{t+1} > 0) = \Theta(d\varrho)$ (see (2.46) and Proposition 2.6.10). Also observe that $\mathcal{D} \wedge (X = 0)$ implies that $Y + Z_2 > 0$. In fact, we must have at least two components of size greater than 1, so at least one of these must contribute to Y or Z_2. Then, we have that $P_1 \leq \mathbf{P}(Y_t + Z_t = 0 \wedge Y_{t+1} + Z_{2,t+1} > 0)$ and $P_2 \leq \mathbf{P}(Y_t + Z_t > 0 \wedge B_t > 0)$, and from Lemma 2.6.11 we get

$$P_1, P_2 = o\big(\mathbf{P}(X_t = 0 \wedge X_{t+1} > 0)\big). \tag{2.48}$$

Otherwise if $d\varrho = \Omega(1)$, then $\mathbf{P}(X_t = 0 \wedge X_{t+1} > 0) = \Theta(1)$. In this case, we simply use the fact that $P_1 \leq \mathbf{P}(Y_{t+1} + Z_{2,t+1} > 0) = o(1)$ and $P_2 \leq \mathbf{P}(Y_t + Z_t > 0) = o(1)$ (see Theorem 2.6.6 and Observation 2.6.8), and deduce that (2.48) also holds.

Finally, the asymptotic expression of $\mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1})$ is obtained from (2.46), (2.47) and (2.48). Moreover, by considering the time-reversed process, we deduce that $\mathbf{P}(\mathcal{D}_t \wedge \mathcal{C}_{t+1}) = \mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1})$. The remaining probabilities in the statement are computed from Corollary 2.6.7 and Observation 2.6.8, and using the fact that

$$\mathbf{P}(\mathcal{C}_t \wedge \mathcal{C}_{t+1}) = \mathbf{P}(\mathcal{C}_t) - \mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1}),$$
$$\mathbf{P}(\mathcal{D}_t \wedge \mathcal{D}_{t+1}) = \mathbf{P}(\mathcal{D}_t) - \mathbf{P}(\mathcal{D}_t \wedge \mathcal{C}_{t+1}).$$

□

In a similar way to the cycle case, we define the *lifespan* of a simple component line as the number of time steps for which the line is alive. For any vertex v and time step t, consider the random variable $L_{v,t}$ defined as follows: If at time step $t+1$ there is a simple component at v, then $L_{v,t}$ is the number of time steps (possibly infinity) that the corresponding simple component line stays alive starting from time step $t+1$; Otherwise, $L_{v,t}$ is defined to be 0. So if a birth takes place at vertex v precisely between time steps t and $t+1$, then $L_{v,t}$ corresponds to the lifespan of the simple component line being born.

Define L_{av} (the average lifespan of simple component lines), L_T and L^* as in Subsection 2.5.2 but in terms of the new definition of $L_{v,t}$. See also the train paradox discussed in that subsection. The next result characterises the average lifespan of simple component lines and how it relates to the initial configuration of walkers.

Theorem 2.6.13. *For the walkers model on T_N,*

$$L_{\mathrm{av}} \sim \begin{cases} \frac{1}{b\varrho} & \text{if } d\varrho = o(1), \\ \frac{\mu}{\mu\left(1-e^{-b\varrho}\right)} & \text{if } d\varrho = \Theta(1), \\ 1 & \text{if } d\varrho = \omega(1). \end{cases}$$

Furthermore, L_T converges in probability for T growing large (N fixed) towards L^*, where $L^* \sim L_{\mathrm{av}}$ a.a.s.

Proof. The argument used for Theorem 2.5.8 can be adapted by replacing H with X, and we get

$$L_{\mathrm{av}} = \frac{\mathbf{E}X}{\mathbf{E}B}.$$

Then the first part of the theorem follows from Theorem 2.6.6 and Proposition 2.6.10. The proof of the second part is analogous to that of Theorem 2.5.9, also changing H for X. □

Our final result provides the expected time that the graph of walkers remains connected or disconnected, after the point in time that it becomes so. Define (dis)connected periods, $L\mathcal{C}_{\mathrm{av}}$, $L\mathcal{C}_T$, $L\mathcal{C}^*$, $L\mathcal{D}_{\mathrm{av}}$, $L\mathcal{D}_T$ and $L\mathcal{D}^*$ as in Subsection 2.5.2.

Theorem 2.6.14. *For the walkers model on T_N, the average length of a connected and a disconnected period of $\left(G(\mathcal{V}_t)\right)_{t\in\mathbb{Z}}$ satisfy respectively*

$$L\mathcal{C}_{\mathrm{av}} \sim \begin{cases} \frac{1}{\mu b\varrho} & \text{if } d\varrho = o(1), \\ \frac{1}{1-e^{-\mu\left(1-e^{-b\varrho}\right)}} & \text{if } d\varrho = \Theta(1), \\ \frac{1}{1-e^{-\mu}} & \text{if } d\varrho = \omega(1) \end{cases} \qquad \text{and}$$

$$L\mathcal{D}_{\mathrm{av}} \sim \begin{cases} \frac{e^{\mu}-1}{\mu b\varrho} & \text{if } d\varrho = o(1), \\ \frac{e^{\mu}-1}{1-e^{-\mu\left(1-e^{-b\varrho}\right)}} & \text{if } d\varrho = \Theta(1), \\ e^{\mu} & \text{if } d\varrho = \omega(1). \end{cases}$$

Furthermore, $L\mathcal{C}_T$ ($L\mathcal{D}_T$) converges in probability for T growing large (N fixed) towards $L\mathcal{C}^*$ ($L\mathcal{D}^*$), where $L\mathcal{C}^* \sim L\mathcal{C}_{\mathrm{av}}$ a.a.s. and $L\mathcal{D}^* \sim L\mathcal{D}_{\mathrm{av}}$ a.a.s.

Proof. We repeat the same arguments as those in the proofs of Theorems 2.5.12 and 2.5.13. We require the asymptotic values of $\mathbf{P}(\mathcal{C})$, $\mathbf{P}(\mathcal{D})$, $\mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1})$ and $\mathbf{P}(\mathcal{D}_t \wedge \mathcal{C}_{t+1})$, which are obtained from Corollary 2.6.7 and Lemma 2.6.12 □

2.7 Empirical Analysis for the Grid

In the present section we empirically validate the asymptotics results previously obtained in this chapter, for grids of reasonable size. In particular, for the static case we, deal with grids of size $N = 1000 \times 1000$, $N = 3000 \times 3000$ and $N = 10000 \times 10000$. For the dynamic case, the size is $N = 1000 \times 1000$. The experiments show that, in most cases, the behaviour of the model is not far from the theoretical predictions in the limit. This study is restricted to the ℓ^1-normed distance case, and thus we have $h = 2d(d+1)$.

2.7.1 Static Properties

The results in Subsection 2.6.1, provide a sharp characterisation of the connectedness of $G(\mathcal{V})$ in terms of $\mu = N\,(1-e^{-\varrho})\,e^{-h\varrho}$, and show the existence of a *phase transition* when $\mu = \Theta(1)$. At this point, there is a Poisson number of simple components, and the remaining walkers belong to one single giant component. From Theorem 2.6.6, the relationship between w and h (or d) at the connectivity threshold can be easily computed. Observe that if we set $\mu = \Theta(1)$, then having a large amount of walkers at the threshold requires having a small h and vice versa. For instance, some usual situations can be summarised in the following result.

Proposition 2.7.1. *In the case that $\mu = \Theta(1)$, we have that*

(i). $h = \Theta(1)$ *iff* $w = \Theta(N \log N)$,

(ii). $h = \Theta(\log N)$ *iff* $w = \Theta(N)$,

(iii). $h = \Theta(N^c)$ *iff* $w = \Theta(N^{1-c} \log N)$, *for* $0 < c \leq 1$,

(iv). $h = \Theta(\frac{N}{\log N})$ *iff* $w = \Theta(\log N \log\log N)$.

Proof. If we apply logarithms to $\mu = N\,(1-e^{-\varrho})\,e^{-h\varrho}$, we obtain that $\log N(1-e^{-\varrho}) = h\varrho + \Theta(1)$. Then the proof is immediate from elementary computations, taking into account the initial restrictions imposed to w and h. □

Now, we test experimentally the asymptotic relations in Proposition 2.7.1 for some interesting values of N which may arise in real life. We deal with grids of sizes $N = 1000^2$, 3000^2 and 10000^2, and study each case for d ranging from a constant to a function growing large slightly slower than n. For each pair N, d, we choose the amount of walkers w that makes $\mu = \log 2$. (Since w must be an integer, we choose the closest one.) A summary of these parameters can be found in Table 2.5.

Note that we are demanding $\mu = \log 2$ because the condition $\mu = \Theta(1)$ is purely asymptotic and makes no sense for fixed values of N. The reason for the choice of this particular value is that, according to the theoretical results, the number of simple components when $\mu = \log 2$ should be roughly Poisson with expectation $\log 2$. This makes the probability of $G(\mathcal{V})$ being connected (or disconnected) be around 1/2.

For each triple of parameters N, w and d described in Table 2.5, we experimentally place w walkers u.a.r. on a grid of size N, check whether $G(\mathcal{V})$ is connected or not, and count the number of occupied vertices, the number of components, the size of the biggest component and the average size of the remaining ones. This experiment is independently

	$N = 1000 \times 1000$	$N = 3000 \times 3000$	$N = 10000 \times 10000$
d constant	$d = 3$ $w = 555377$	$d = 3$ $w = 5866110$	$d = 3$ $w = 75639720$
$d = \log n$	$d = 7$ $w = 106128$	$d = 8$ $w = 875018$	$d = 9$ $w = 9079434$
$d = n^{1/3}$	$d = 10$ $w = 50804$	$d = 14$ $w = 275985$	$d = 22$ $w = 1436466$
$d = n^{1/2}$	$d = 32$ $w = 4113$	$d = 55$ $w = 14538$	$d = 100$ $w = 55931$
$d = n^{2/3}$	$d = 100$ $w = 301$	$d = 208$ $w = 719$	$d = 464$ $w = 1825$
$d = n/\log n$	$d = 145$ $w = 122$	$d = 375$ $w = 177$	$d = 1086$ $w = 249$

Table 2.5: Parameters at the phase transition ($\mu = \log 2$).

repeated 100 times and we take averages of the observed magnitudes. Then we compare the obtained data with what we would expect to get according to the theoretical results, which is listed in Table 2.6.

Occupied vertices	$N(1 - e^{-\varrho})$
Probability that $G(\mathcal{V})$ is connected	$e^{-\mu}$
Number of components	$1 + \mu$
Size of the biggest component	$N(1 - e^{-\varrho}) - \mu$
Average size of the other components	1

Table 2.6: Asymptotic expected values for N growing large.

For each particular run of our experiments, our algorithm must assign at random grid coordinates (i, j) to each walker. It is convenient to store this data in a Hashing table of size w instead of using a $n \times n$ table in order to optimise space resources. By doing this we don't loose much time efficiency, since the cost of checking whether a given vertex is occupied remains constant in expectation. We use then a Depth-First-Search to find all components. The whole algorithm takes expected time $\Theta(wh)$, since for each walker we examine all the grid positions within distance d, and requires space $\Theta(w)$. Moreover since $\mu = \Theta(1)$, we have $wh \sim N \log w$, and then the time complexity is roughly proportional to N apart from logarithmic factors.

Tables 2.7, 2.8 and 2.9 contrast the averages of the experimental results with the asymptotic expected values (see Table 2.6) for the selected parameters.

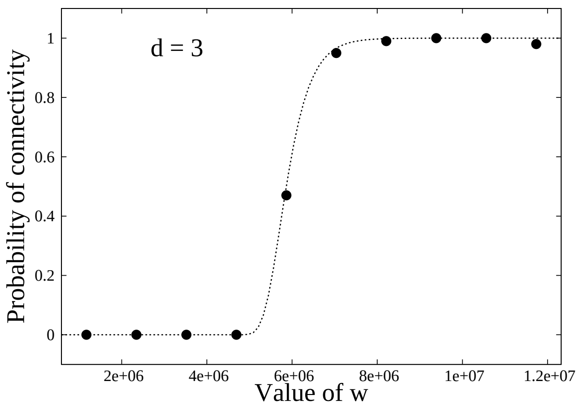
What we described so far accounts for the situation at the *phase transition*. However, we also want to verify experimentally that there is indeed a phase transition. We consider only the case $N = 3000 \times 3000$ and deal with the same types of d as before: $d =$ constant, $d = \log n$, $d = n^{1/3}$, $d = n^{1/2}$, $d = n^{2/3}$ and $d = n/\log n$. For each d, we consider 10 equidistant values for w, ranging between $w_0/5$ and $2w_0$, where w_0 is the amount of walkers needed to have $\mu = \log 2$ (see Table 2.5). (As before, all these quantities are rounded to the nearest integer.) For each triple of parameters N, w and d, we sample again 100 independent random instances of $G(\mathcal{V})$ and check whether they are connected. The probability of connectivity can be estimated from the ratio between *connected* outputs and the total number of trials.

Since we are just concerned with connectivity, we can slightly modify our previous algorithm to improve time performance. Given a random arrangement of walkers in the grid T_N stored as before in a Hashing table, we first examine the existence of simple components. We run along the table and, for each unmarked walker, we look for another walker within distance d and mark both as “not in a simple component”. If we detect a simple component, we stop and output *disconnected*. Otherwise, we perform as before a Depth-First-Search to find all components. In the worst case, the algorithm has the same complexity as the previous one, but if $G(\mathcal{V})$ has some simple components, we may be lucky and have a quick output. This proves quite useful for our particular kind of graphs since simple components are very common.

The plots in Figures 2.12, 2.13 and 2.14 show for each grid of size N and distance d, the evolution of the probability that $G(\mathcal{V})$ is connected as we increase the amount of walkers. The dots correspond to the experimental values we obtained. In contrast, the curves show the theoretical value of this probability according to Subsection 2.6.1. This is asymptotically $e^{-\mu}$, where the expression of μ is given in Theorem 2.6.6.

We used for the tests the joint effort of 10 computers in the MALLBA cluster at LSI with the following power:

- Processor: AMD K6(tm) 3D processor (450 MHz)
- Main memory: 256 Mb



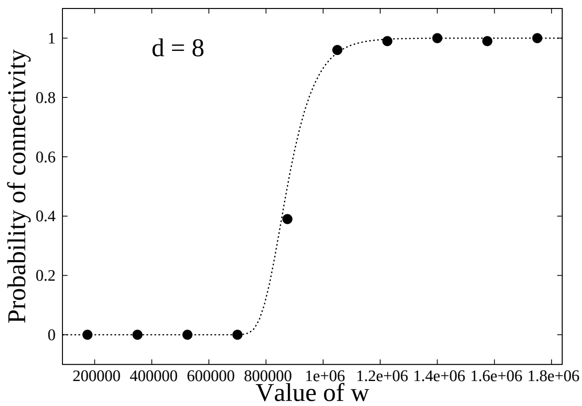


Figure 2.12: Threshold of connectivity.

Conclusions for the Static Case Experiments

Our experimental results show that the qualitative behaviour of the walkers model sticks reasonably well to the theoretical predictions. In fact, we observe a clear threshold phenomenon on the connectivity property even though in some cases the observed critical point

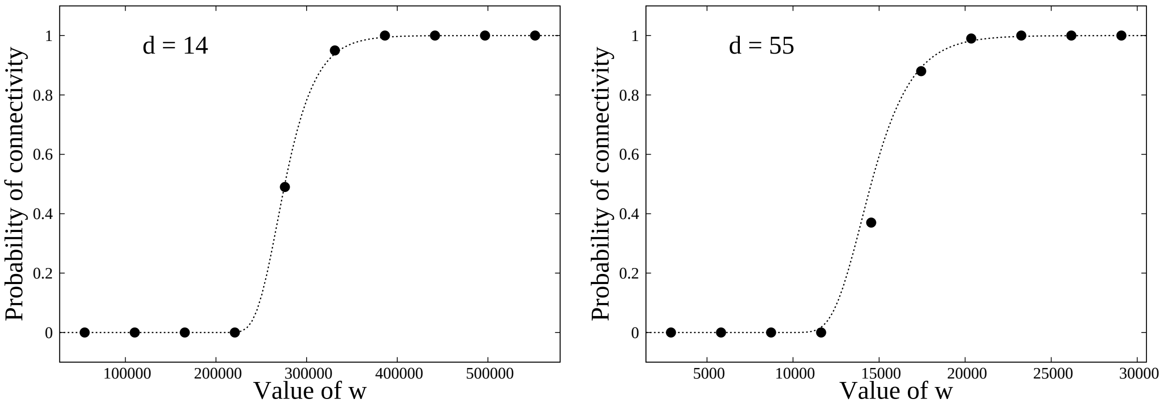


Figure 2.13: Threshold of connectivity.

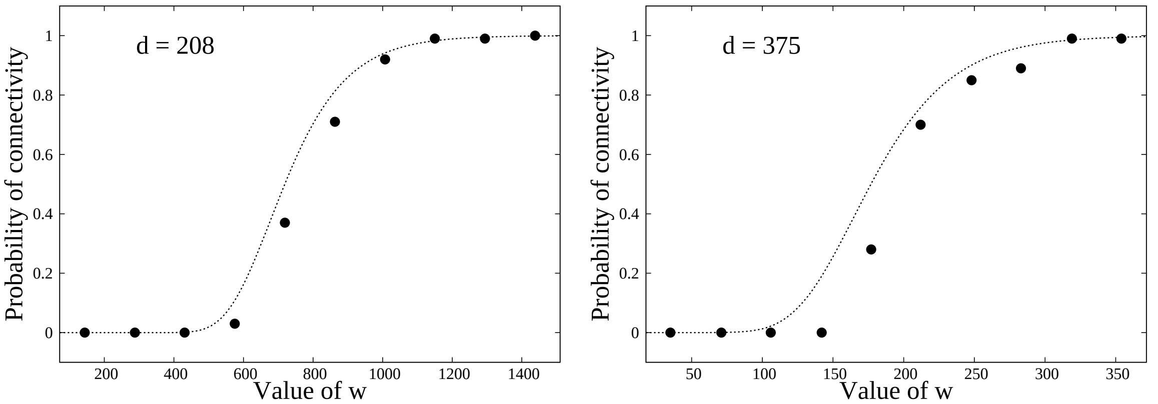


Figure 2.14: Threshold of connectivity.

is slightly displaced from its theoretical location. Furthermore, at the phase transition, we observed that there is indeed one giant component consisting of the vast majority of walkers, and a few small components (not far from being simple in most cases).

From a quantitative point of view, the accuracy of our predictions is dramatically better for small d. This is probably due to the fact that, at the phase transition, a smaller distance d requires a bigger amount of walkers w, and we recall that the asymptotic results in Subsection 2.6.1 require $w \to \infty$ as a regularity condition. For instance, in our last case where $d = n/\log n$, the corresponding w is essentially logarithmic on N. Then, we may need to consider exponentially huge grid sizes in order to have a big amount of walkers and get reliable predictions.

Strangely enough, for the cases we considered, the accuracy of the predictions does not seem to improve significantly as we increase the grid size N from 10^6 to 10^8. Possibly the improvement is too small to be detected within the precision of our experiments. We could always perform more trials for each test, or we could even consider much bigger values of N, but this last is beyond our current computational means.

2.7.2 Dynamic Properties

In this subsection we experimentally check the validity of the asymptotic results in Theorem 2.6.14 in Subsection 2.6.2 for grids of size $N = 1000 \times 1000$. We consider the parameters d and w listed in Table 2.5.

For each triple N, w and d, we experimentally place the walkers on the grid u.a.r. as in the static situation, but then we perform T dynamic steps, for some big enough chosen T. We examine the connectivity of $G(\mathcal{V}_t)$ at each time step, measure the length of the different periods we encounter in which $G(\mathcal{V}_t)$ is (dis)connected, and then take the average. This accounts for the average connection time $L\mathcal{C}_T$ and the average disconnection time $L\mathcal{D}_T$ of $\big(G(\mathcal{V}_t)\big)_{t\in\mathbb{Z}}$ between times 1 and T. We repeat this experiment independently 50 times and take the average of the averages. We choose T to be about 500 times the final value we expect to get. In our cases, this ranges from 1000 to 20000 depending on the parameters N, w and d (see last column in Table 2.10).

Our algorithm is an easy extension of the one used for the static situation. For each walker we choose its grid coordinates at random and store them in a Hashing table of size w. To perform a dynamic transition, we just need to run along the walkers and move each one to any of the 4 neighbouring grid positions with equal probability. This has expected time complexity of $\Theta(w)$ since the expected time for search, insertion and removal in the table is constant. In addition, we must look at the connectivity at each time step, and the same observations we made in Subsection 2.7.1 apply here. It is of great help checking the existence of simple components first, since it is usually much faster, and it is a sufficient condition for non-connectivity. Hence, the algorithm requires space $\Theta(w)$ and takes time $O(Twh)$, but it usually runs much faster at the steps when $G(\mathcal{V}_t)$ is disconnected.

We used the same machines and system of computation as for the static case, and the results are summarised in Table 2.10.

Conclusions for the Dynamic Case Experiments

The experimental values obtained for $L\mathcal{C}_T$ and $L\mathcal{D}_T$ are in all cases of roughly the same order of magnitude as the values predicted by the theoretical model. However, the level of accuracy is much higher for the smaller values of d and gets poorer for the largest d's, exactly as in the static situation. Again the reason may be that in these last cases, the considered amount of walkers w is quite small, while in Subsection 2.6.1 w is required to grow to infinity.

We observe as well that the average length of the disconnected periods is larger than that of the connected periods and it is much closer to the predicted value. Here is a plausible explanation to this: We were studying situations where ideally, in the limit, there should be one giant component and an average of $\mu = \log 2$ small (indeed simple) components. In this case the probability of connectivity would be $\mathbf{P}(\mathcal{C}) = e^{-\mu} = 1/2$, and moreover we would have

$$L\mathcal{C}_T = \frac{1}{1 - e^{-\mu\left(1-e^{-b\varrho}\right)}} = \frac{e^{\mu} - 1}{1 - e^{-\mu\left(1-e^{-b\varrho}\right)}} = L\mathcal{D}_T.$$

But as shown in Figures 2.12, 2.13 and 2.14 the real observed probability of connectivity is mostly below the theoretical predictions at the limit, and this fact is stressed for the largest values of d. This is the same as saying that the phase transition occurs slightly afterwards for the observed cases than in the theoretical limit, or equivalently that the observed amount of non-giant components is slightly bigger than what we would asymptotically expect. This explains why in our experiments $L\mathcal{C}_T < L\mathcal{D}_T$.

We note that this deviation between the observed number of non-giant components and μ, gets amplified in the expressions of $L\mathcal{C}_T$ and $L\mathcal{D}_T$ since μ appears there exponentially. So let's try the following: let us use the average number of non-giant components we observed (see Table 2.7) as the value of μ in the expressions in Theorem 2.6.14. Then, in Table 2.11 we compare the obtained values with our observations. The new predictions turn out to be much closer to the experimental quantities.

This gives reasonable evidence for the validity of Theorem 2.6.14, but also restricts its applicability to the cases where the number of non-giant components is close to the expected number μ of simple components in the limit.

2.8 Conclusions and Open Problems

In this work we have characterised connectivity issues of a very large set of moving agents, which move through a prescribed real or virtual graph. We believe it is the first time that these kind of characterisations have been obtained, and it could open a interesting line of research. We gave characterisations for the cycle and the grid. The results obtained for the grid could easily be extended to the grid with diagonals. Also, an approach similar to ours should work with the k-dimensional toroidal grid, but a suitable substitute for the Geometric Lemma needs to be found.

In our model we use a fixed number w of walkers. One could alternatively place walkers randomly so that each vertex is occupied independently with probability p. For example, they could be Poissonly distributed at each vertex, with parameter λ such that $1-e^{-\lambda} = p$. In the static case these would bear a similar relation to our model as between the random graph models $G(n,m)$ and $G(n,p)$, and as for that case, one would expect similar properties when p is approximately w/N (or, more precisely, $1-e^{-w/N}$ to capture the case that w is close to or greater than N). This would simplify some of our analysis (e.g. the proof of Lemma 2.4.2). However, it would be difficult to deduce all the results for our model in such a way. For one thing, some of the properties we study are not convex in the required sense (see [46]). There are also other obstacles to using models with independent occupancy probabilities. For instance, if $w(s_0/N)^2 \neq o(1)$ then $(1-S_0/N))^w$ as in Lemma 2.4.2 is not asymptotic to $e^{-S_0\varrho}$.

Further work is the extension of the results presented for the cycle and the toroidal grid to other families of graphs, not necessarily to model realistic networks. One interesting case is the n-dimensional hypercube of $N = 2^n$ vertices, as the number of neighbours of a vertex is not constant. Another further project is to study the connectivity of walkers when the underlying topology has obstacles which can interfere with communication.

$N = 1000 \times 1000$		Experimental average	Theoretical value
$d = 3$ $w = 555377$	Occupied vertices	426140.57	426144.12
	Prob. of connectivity	0.54	0.50
	Number of components	1.68	1.69
	Size of the biggest comp.	426139.80	426143.43
	Av. size of other comp.	1.14	1
$d = 7$ $w = 106128$	Occupied vertices	100674.83	100690.47
	Prob. of connectivity	0.40	0.50
	Number of components	1.89	1.69
	Size of the biggest comp.	100673.72	100689.78
	Av. size of other comp.	1.23	1
$d = 10$ $w = 50804$	Occupied vertices	49533.84	49535.06
	Prob. of connectivity	0.39	0.50
	Number of components	1.95	1.69
	Size of the biggest comp.	49532.60	49534.36
	Av. size of other comp.	1.31	1
$d = 32$ $w = 4113$	Occupied vertices	4104.37	4104.55
	Prob. of connectivity	0.37	0.50
	Number of components	1.97	1.69
	Size of the biggest comp.	4102.96	4103.86
	Av. size of other comp.	1.53	1
$d = 100$ $w = 301$	Occupied vertices	301.00	300.95
	Prob. of connectivity	0.36	0.50
	Number of components	2.16	1.69
	Size of the biggest comp.	298.52	300.27
	Av. size of other comp.	2.02	1
$d = 145$ $w = 122$	Occupied vertices	122.00	121.99
	Prob. of connectivity	0.19	0.50
	Number of components	2.38	1.69
	Size of the biggest comp.	118.05	121.30
	Av. size of other comp.	2.69	1

Table 2.7: Contrasted results at the phase transition for $N = 1000 \times 1000$.

$N = 3000 \times 3000$		Experimental average	Theoretical value
$d = 3$	Occupied vertices	4309968.88	4309990.64
	Prob. of connectivity	0.49	0.50
	Number of components	1.67	1.69
$w = 5866110$	Size of the biggest comp.	4309968.18	4309989.95
	Av. size of other comp.	1.05	1
$d = 8$	Occupied vertices	833825.19	833827.19
	Prob. of connectivity	0.37	0.50
	Number of components	1.87	1.69
$w = 875018$	Size of the biggest comp.	833824.16	833826.49
	Av. size of other comp.	1.22	1
$d = 14$	Occupied vertices	271795.67	271796.38
	Prob. of connectivity	0.39	0.5
	Number of components	1.86	1.69
$w = 275985$	Size of the biggest comp.	271794.6	271795.69
	Av. size of other comp.	1.25	1
$d = 55$	Occupied vertices	14525.60	14526.26
	Prob. of connectivity	0.41	0.5
	Number of components	1.86	1.69
$w = 14538$	Size of the biggest comp.	14524.48	14525.57
	Av. size of other comp.	1.34	1
$d = 208$	Occupied vertices	718.97	718.97
	Prob. of connectivity	0.29	0.50
	Number of components	2.10	1.69
$w = 719$	Size of the biggest comp.	717.16	718.28
	Av. size of other comp.	1.58	1
$d = 375$	Occupied vertices	176.99	177.00
	Prob. of connectivity	0.28	0.50
	Number of components	2.29	1.69
$w = 177$	Size of the biggest comp.	174.20	176.31
	Av. size of other comp.	1.98	1

Table 2.8: Contrasted results at the phase transition for $N = 3000 \times 3000$.

$N = 10000 \times 10000$		Experimental average	Theoretical value
$d = 9$	Occupied vertices	8679487.90	8679449.86
	Prob. of connectivity	0.52	0.5
	Number of components	1.71	1.69
$w = 9079434$	Size of the biggest comp.	8679487.05	8679449.16
	Av. size of other comp.	1.23	1
$d = 22$	Occupied vertices	1426204.88	1426198.05
	Prob. of connectivity	0.40	0.50
	Number of components	1.89	1.69
$w = 1436466$	Size of the biggest comp.	1426203.82	1426197.36
	Av. size of other comp.	1.19	1
$d = 100$	Occupied vertices	55915.73	55915.36
	Prob. of connectivity	0.38	0.50
	Number of components	1.97	1.69
$w = 55931$	Size of the biggest comp.	55914.51	55914.67
	Av. size of other comp.	1.3	1
$d = 464$	Occupied vertices	1824.97	1824.98
	Prob. of connectivity	0.37	0.50
	Number of components	2.04	1.69
$w = 1825$	Size of the biggest comp.	1823.39	1824.29
	Av. size of other comp.	1.48	1
$d = 1086$	Occupied vertices	249.00	249.00
	Prob. of connectivity	0.33	0.50
	Number of components	2.23	1.70
$w = 249$	Size of the biggest comp.	246.25	248.30
	Av. size of other comp.	2.32	1

Table 2.9: Contrasted results at the phase transition for $N = 10000 \times 10000$.

$N = 1000 \times 1000$		Experimental average	Theoretical expectation
$d = 3$	$\mathcal{LC}_T$	1.93	2.05
$w = 555377$	$\mathcal{LD}_T$	2.14	2.05
$d = 7$	$\mathcal{LC}_T$	2.05	2.41
$w = 106128$	$\mathcal{LD}_T$	2.70	2.41
$d = 10$	$\mathcal{LC}_T$	2.28	2.79
$w = 50804$	$\mathcal{LD}_T$	3.17	2.79
$d = 32$	$\mathcal{LC}_T$	4.89	6.75
$w = 4113$	$\mathcal{LD}_T$	7.56	6.75
$d = 100$	$\mathcal{LC}_T$	14.14	25.36
$w = 301$	$\mathcal{LD}_T$	27.86	25.13
$d = 145$	$\mathcal{LC}_T$	18.97	41.80
$w = 122$	$\mathcal{LD}_T$	55.20	42.09

Table 2.10: Contrasted results for the dynamic process ($N = 1000 \times 1000$).

$N = 1000 \times 1000$		Experimental average	Modified prediction
$d = 3$	$\mathcal{LC}_T$	1.93	2.08
$w = 555377$	$\mathcal{LD}_T$	2.14	2.02
$d = 7$	$\mathcal{LC}_T$	2.05	2.01
$w = 106128$	$\mathcal{LD}_T$	2.70	2.88
$d = 10$	$\mathcal{LC}_T$	2.28	2.20
$w = 50804$	$\mathcal{LD}_T$	3.17	2.49
$d = 32$	$\mathcal{LC}_T$	4.89	4.97
$w = 4113$	$\mathcal{LD}_T$	7.56	8.15
$d = 100$	$\mathcal{LC}_T$	14.14	15.26
$w = 301$	$\mathcal{LD}_T$	27.86	33.42
$d = 145$	$\mathcal{LC}_T$	18.97	21.35
$w = 122$	$\mathcal{LD}_T$	55.20	63.51

Table 2.11: New predictions, using the observed average number of non-giant components instead of μ.

3

Dynamic Random Geometric Graphs

3.1 Introduction

In the present chapter, we study the evolution of connectivity of a random geometric graph over the unit torus $[0,1)^2$ as the vertices, also denoted agents, move randomly around. In particular, starting from a given random geometric graph, where the radius r is the known connectivity threshold r_c (see Section 3.2), each agent chooses independently and uniformly at random an angle $\alpha \in [0, 2\pi)$, and moves a distance s in that direction for a period of m steps. Then a new angle is selected independently for each agent, and the process repeats. In the literature, this mobility model is denoted as the *random walk* model [38, 47]. Our main result is that we provide precise asymptotic results for the expected number of steps that the graph remains connected once it becomes connected, and the expected number of steps the graph remains disconnected once it becomes disconnected. These expressions will be given as a function of the number of vertices n, the number of steps in the same direction m and the step size s. As it will be indicated in Section 3.3, the proof techniques will be different for different sizes s of basic step. In addition, for the static model of random geometric graphs where the radius r is the known connectivity threshold, we provide asymptotic bounds on the probability of occurrence of components according to their sizes. All the computations were made using the usual Euclidean distance in the torus, but similar results can be obtained for any ℓ_p-normed distance, $1 \leq p \leq \infty$, with the only prize of changing a constant factor in the expressions. Moreover, our results also can be extended to any d-dimensional unit torus of bounded dimension.

3.2 Static Properties

We first state some of the known results about static random geometric graphs, which will be the starting point to derive our results. We refer to [66] for further detail.

Given a set of n agents and a positive real $r = r(n)$, each agent is placed at some random position in the unit torus $[0,1)^2$ selected independently and u.a.r. Throughout the chapter, we denote by $X_i = (x_i, y_i)$ the random position of agent i for $i \in \{1,\dots,n\}$, and let $\mathcal{X} = \mathcal{X}(n) = \bigcup_{i=1}^{n}\{X_i\}$. We note that with probability 1 no two agents choose the same position and thus we restrict the attention to the case that $|\mathcal{X}| = n$. We define $G(\mathcal{X}; r)$ as the random graph having $\mathcal{X}$ as the vertex set, and with an edge connecting each pair of vertices X_i and X_j in $\mathcal{X}$ at distance $d(X_i, X_j) \le r$, where $d(\cdot,\cdot)$ denotes the Euclidean distance in the torus.

$$d(X,Y) = \min\{||X - Y + Z|| \ : \ Z \in \mathbb{Z}^2\}, \qquad \forall X, Y \in [0,1)^2$$

Unless otherwise stated, all our stated results are asymptotic as $n \to \infty$. We assume hereinafter that $r = o(1)$. Otherwise a Balls and Bins argument shows that $G(\mathcal{X}; r)$ is trivially a.a.s. connected.

Let K_1 be the random variable counting the number of isolated vertices in $G(\mathcal{X}; r)$. Then, by multiplying the probability that one vertex is isolated by the number of vertices we obtain

Lemma 3.2.1. $\mathbf{E}K_1 = n(1-\pi r^2)^{n-1} = ne^{-\pi r^2 n - O(r^4 n)}$.

Define $\mu = ne^{-\pi r^2 n}$. Observe from Lemma 3.2.1 that this parameter μ is closely related to $\mathbf{E}K_1$. More precisely, $\mu = o(1)$ iff $\mathbf{E}K_1 = o(1)$, and if $\mu = \Omega(1)$ then $\mathbf{E}K_1 \sim \mu$. Moreover the asymptotic behaviour of μ characterises the connectivity of $G(\mathcal{X}; r)$. In fact (see [64, 65]),

Theorem 3.2.2.

- *If $\mu \to 0$, then a.a.s. $G(\mathcal{X}; r)$ is connected.*
- *If $\mu = \Theta(1)$, then a.a.s. $G(\mathcal{X}; r)$ consists of one giant component of size $> n/2$ and a Poisson number (with parameter μ) of isolated vertices.*
- *If $\mu \to \infty$, then a.a.s. $G(\mathcal{X}; r)$ is disconnected.*

From the definition of μ we deduce that $\mu = \Theta(1)$ iff $r = \sqrt{\frac{\log n \pm O(1)}{\pi n}}$. Therefore as a weaker consequence we conclude that the property of connectivity of $G(\mathcal{X}; r)$ exhibits a sharp threshold at $r_c = \sqrt{\frac{\log n}{\pi n}}$. Theorem 3.2.2 also implies that, if $\mu = \Theta(1)$, the components of size 1 (i.e. isolated vertices) are predominant and have the main contribution to the connectivity of $G(\mathcal{X}; r)$. In fact if $\mathcal{C}$ (respectively $\mathcal{D}$) denotes the event that $G(\mathcal{X}; r)$ is connected (respectively disconnected), we have the following

Corollary 3.2.3. *Assume that $\mu = \Theta(1)$. Then*

$$\mathbf{P}(\mathcal{C}) \sim \mathbf{P}(K_1 = 0) \sim e^{-\mu}, \qquad \mathbf{P}(\mathcal{D}) \sim \mathbf{P}(K_1 > 0) \sim 1 - e^{-\mu}.$$

Observe that we used the fact that, if $\mu = \Theta(1)$, the probability that $G(\mathcal{X}; r)$ has some component of size greater than 1 other than the giant component is $o(1)$. We give more accurate bounds on this probability. Moreover we characterise the probability of having components of any fixed size. Before stating this more precisely we need some definitions.

Given a component Γ of $G(\mathcal{X}; r)$, we say that Γ is *embeddable* if it is contained in some square with sides parallel to the axes of the torus and length $1 - 2r$. In other words, Γ is embeddable if it can be mapped into the square $[r, 1-r]^2$ by a translation in the torus. Embeddable components do not wrap around the torus. Throughout the chapter, in all geometrical descriptions involving an embeddable component Γ, we assume that Γ is contained in $[r, 1-r]^2$ and regard the torus $[0, 1)^2$ as the unit square and $d(\cdot, \cdot)$ as the usual Euclidean distance. This assumption is often not explicitly mentioned. Hence terms as "left", "right", "above" and "below" are globally defined.

On the other hand, components which are not embeddable must have size at least $\Omega(1/r)$. Note that sometimes several non-embeddable components can coexist together (see Figure 3.1). However, there are some non-embeddable components which are so spread around the torus that do not allow any room for other non-embeddable ones. We call these components *solitary*, and by definition we can have at most one solitary component. We cannot disprove the existence of this solitary component, since with probability $1 - o(1)$ there exists one giant component of this nature. For components which are not solitary

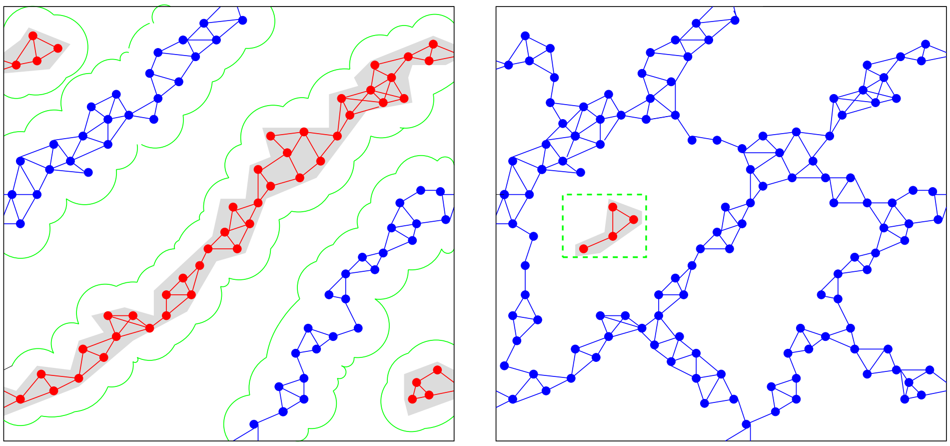


Figure 3.1: Non-embeddable components on the unit torus; to the left two non-embeddable and non-solitary components (one with shaded background), to the right a solitary non-embeddable and an embeddable component (the latter with shaded background)

(i.e., either embeddable or non-embeddable but able to coexist with other non-embeddable ones), we give asymptotic bounds on the probability of their existence according to their size.

Given a fixed integer $\ell \geq 1$, let K_ℓ be the number of components in $G(\mathcal{X}; r)$ of size exactly ℓ. For large enough n, we can assume these components to be embeddable, since $r = o(1)$. Moreover, for any fixed $\epsilon > 0$, let $K'_{\epsilon,\ell}$ be the number of components of size exactly ℓ which have all their vertices at distance at most ϵr from their leftmost one. Finally, let $\widetilde{K}_\ell$ denote the number of components of size at least ℓ which are not solitary. Figure 3.2

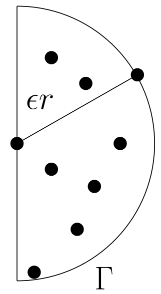


Figure 3.2: A component Γ of size exactly $\ell = 9$ and with its vertices at distance at most ϵr from the leftmost one

shows an example of a component that contributes to $K'_{\epsilon,9}$.

Notice that $K'_{\epsilon,\ell} \leq K_\ell \leq \widetilde{K}_\ell$. However, in the following results we show that asymptotically all the weight in the probability that $\widetilde{K}_\ell > 0$ comes from components which also contribute to $K'_{\epsilon,\ell}$ for ϵ arbitrarily small. This means that the more common components of size at least ℓ are cliques of size exactly ℓ with all their vertices close together.

Lemma 3.2.4. *Let $\ell \geq 2$ be a fixed integer, and $0 < \epsilon < 1/2$ also fixed. Assume that $\mu = \Theta(1)$. Then,*

$$\mathbf{E}K'_{\epsilon,\ell} = \Theta(1/\log^{\ell-1} n)$$

Proof. First observe that with probability 1, for each component Γ which contributes to $K'_{\epsilon,\ell}$, Γ has a unique leftmost vertex X_i and the vertex X_j in Γ at greatest distance from X_i is also unique. Hence, we can restrict our attention to this case.

Fix an arbitrary set of indices $J \subset \{1, \dots, n\}$ of size $|J| = \ell$, with two distinguished elements i and j. Denote by $\mathcal{Y} = \bigcup_{k\in J} X_k$ the set of random points in $\mathcal{X}$ with indices in J. Let $\mathcal{E}$ be the following event: All vertices in $\mathcal{Y}$ are at distance at most ϵr from X_i and to the right of X_i; vertex X_j is the one in $\mathcal{Y}$ with greatest distance from X_i; and the vertices of $\mathcal{Y}$ form a component of $G(\mathcal{X}; r)$. If $\mathbf{P}(\mathcal{E})$ is multiplied by the number of possible choices of i, j and the remaining $\ell - 2$ elements of J, we get

$$\mathbf{E}K'_{\epsilon,\ell} = n(n-1)\binom{n-2}{\ell-2}\mathbf{P}(\mathcal{E}). \tag{3.1}$$

In order to bound the probability of $\mathcal{E}$ we need some definitions. Let $\rho = d(X_i, X_j)$ and let $\mathcal{S}$ be the set of all points in the torus $[0,1)^2$ which are at distance at most r from some vertex in $\mathcal{Y}$ (see Figure 3.3 for an example). Note that ρ and $\mathcal{S}$ depend on the set of random points $\mathcal{Y}$. We first need bounds on $\mathsf{Area}(\mathcal{S})$ in terms of ρ. Observe that $\mathcal{S}$ is contained in the circle of radius $r + \rho$ and centre X_i, and then

$$\mathsf{Area}(\mathcal{S}) \leq \pi(r+\rho)^2. \tag{3.2}$$

Now let $i_\mathsf{L} = i$, i_R, i_T and i_B be respectively the indices of the leftmost, rightmost, topmost and bottommost vertices in $\mathcal{Y}$ (some of these indices possibly equal). Assume w.l.o.g. that the vertical length of $\mathcal{Y}$ (i.e., the vertical distance between X_{i_T} and X_{i_B}) is at least $\rho/\sqrt{2}$. Otherwise, the horizontal length of $\mathcal{Y}$ has this property and we can rotate the descriptions

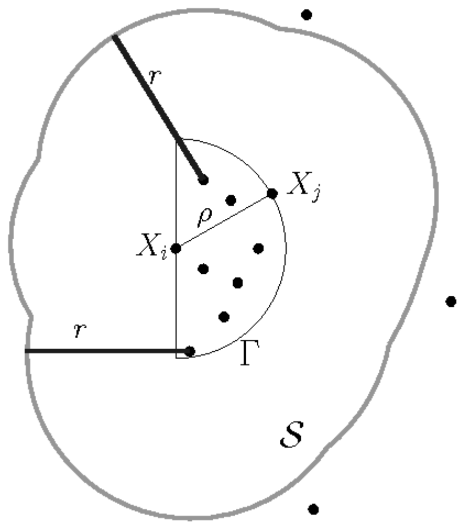


Figure 3.3: The set $\mathcal{S}$ for the component Γ of Figure 3.2

in the argument. The upper half-circle with centre X_{i_T} and the lower half-circle with centre X_{i_B} are disjoint and are contained in $\mathcal{S}$. If X_{i_R} is at greater vertical distance from X_{i_T} than from X_{i_B}, then consider the rectangle of height $\rho/(2\sqrt{2})$ and width $r - \rho/(2\sqrt{2})$ with one corner on X_{i_R} and above and to the right of X_{i_R}. Otherwise, consider the same rectangle below and to the right of X_{i_R}. This rectangle is also contained in $\mathcal{S}$ and its interior does not intersect the previously described half-circles. Analogously, we can find another rectangle with the same properties of height $\rho/(2\sqrt{2})$ and width $r - \rho/(2\sqrt{2})$, to the left of X_{i_L} and either above or below X_{i_L}. Hence,

$$\mathsf{Area}(\mathcal{S}) \geq \pi r^2 + 2\left(\frac{\rho}{2\sqrt{2}}\right)\left(r - \frac{\rho}{2\sqrt{2}}\right). \tag{3.3}$$

Figure 3.4 illustrates the bounds on $\mathsf{Area}(\mathcal{S})$ obtained in (3.2) and (3.3). From (3.2), (3.3) and the fact that $\rho < r/2$, we can write

$$\pi r^2\left(1 + \frac{1}{6}\frac{\rho}{r}\right) < \mathsf{Area}(\mathcal{S}) < \pi r^2\left(1 + \frac{5}{2}\frac{\rho}{r}\right) < \frac{9\pi}{4}r^2. \tag{3.4}$$

Now consider the probability P that the $n - \ell$ vertices not in $\mathcal{Y}$ lie outside $\mathcal{S}$. Then $P = (1 - \mathsf{Area}(\mathcal{S}))^{n-\ell}$. Moreover, by (3.4) and using the fact that $e^{-x-x^2} \leq 1 - x \leq e^{-x}$ for all $x \in [0, 1/2]$, we obtain

$$e^{-(1+5\rho/(2r))\pi r^2 n - (9\pi r^2/4)^2 n} < P < \frac{e^{-(1+\rho/(6r))\pi r^2 n}}{(1 - 9\pi r^2/4)^\ell},$$

and after a few manipulations

$$\left(\frac{\mu}{n}\right)^{1+5\rho/(2r)} e^{-(9\pi r^2/4)^2 n} < P < \left(\frac{\mu}{n}\right)^{1+\rho/(6r)} \frac{1}{(1 - 9\pi r^2/4)^\ell}. \tag{3.5}$$

Event $\mathcal{E}$ can also be described as follows: There is some non-negative real $\rho \leq \epsilon r$ such that X_j is placed at distance ρ from X_i and to the right of X_i; all the remaining vertices

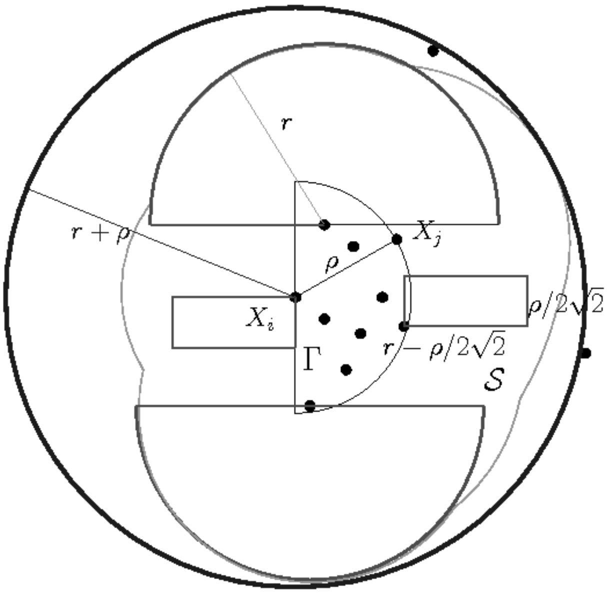


Figure 3.4: Bounds on $\mathsf{Area}(\mathcal{S})$

in $\mathcal{Y}$ are inside the half-circle of centre X_i and radius ρ; and the $n-\ell$ vertices not in $\mathcal{Y}$ lie outside $\mathcal{S}$. Hence, $\mathbf{P}(\mathcal{E})$ can be bounded from above (below) by integrating with respect to ρ the probability density function of $d(X_i, X_j)$ times the probability that the remaining $\ell-2$ selected vertices lie inside the right half-circle of centre X_i and radius ρ times the upper (lower) bound on P we obtained in (3.5):

$$\Theta(1)\, I(5/2) \le \mathbf{P}(\mathcal{E}) \le \Theta(1)\, I(1/6), \tag{3.6}$$

where

$$\begin{aligned} I(\beta) &= \int_0^{\epsilon r} \pi\rho \left(\frac{\pi}{2}\rho^2\right)^{\ell-2} \frac{1}{n^{1+\beta\rho/r}}\, d\rho. \\ &= \frac{2}{n}\left(\frac{\pi}{2}r^2\right)^{\ell-1} \int_0^{\epsilon} x^{2\ell-3} n^{-\beta x}\, dx \end{aligned} \tag{3.7}$$

Since ℓ is fixed, for $\beta = 5/2$ or $\beta = 1/6$,

$$\begin{aligned} I(\beta) &= \Theta\left(\frac{\log^{\ell-1} n}{n^\ell}\right) \int_0^{\epsilon} x^{2\ell-3} n^{-\beta x}\, dx \\ &= \Theta\left(\frac{\log^{\ell-1} n}{n^\ell}\right) \frac{(2\ell-3)!}{(\beta \log n)^{2\ell-2}} \\ &= \Theta\left(\frac{1}{n^\ell \log^{\ell-1} n}\right), \end{aligned} \tag{3.8}$$

and the statement follows from (3.1), (3.6) and (3.8). □

Next, we bound the probability of having non-solitary components of size at least $\ell \ge 2$ which are not ℓ-cliques of small diameter.

Lemma 3.2.5. *Let $\ell \geq 2$ be a fixed integer. Let $\epsilon > 0$ be also fixed. Assume that $\mu = \Theta(1)$. Then*
$$\mathbf{P}(\widetilde{K}_\ell - K'_{\epsilon,\ell} > 0) = O(1/\log^\ell n).$$

Proof. We assume throughout this proof that $\epsilon \leq 10^{-18}$, and prove the claim for this case. The case $\epsilon > 10^{-18}$ follows from the fact that $(\widetilde{K}_\ell - K'_{\epsilon,\ell}) \leq (\widetilde{K}_\ell - K'_{10^{-18},\ell})$.

Consider all the possible components in $G(\mathcal{X}; r)$ which are not solitary. Remove from these components the ones of size at most ℓ and diameter at most ϵr, and denote by M the number of remaining components. By construction $\widetilde{K}_\ell - K'_{\epsilon,\ell} \leq M$, and therefore it is sufficient to prove that $\mathbf{P}(M > 0) = O(1/\log^\ell n)$. The components counted by M are classified into several types, according to their size and diameter. We deal with each type separately.

Part 1. Consider all the possible components in $G(\mathcal{X}; r)$ which have diameter at most ϵr and size between $\ell + 1$ and $\log n/37$. Call them components of *type* 1, and let M_1 denote their number.

For each k, $\ell + 1 \leq k \leq \log n/37$, let E_k be the expected number of components of *type* 1 and size k. We observe that these components have all of their vertices at distance at most ϵr from the leftmost one. Therefore, we can apply the same argument we used for bounding $\mathbf{E}K'_{\epsilon,\ell}$ in the proof of Lemma 3.2.4. Note that (3.1), (3.6) and (3.7) are also valid for sizes not fixed but depending on n. Thus we obtain

$$E_k \leq O(1)n(n-1)\binom{n-2}{k-2} I(1/6),$$

where $I(1/6)$ is defined in (3.7). We use the fact that $\binom{n-2}{k-2} \leq (\frac{ne}{k-2})^{k-2}$ and get

$$E_k = O(1) \log n \left(\frac{e}{2}\frac{\log n}{k-2}\right)^{k-2} \int_0^\epsilon x^{2k-3} n^{-x/6}\, dx \tag{3.9}$$

The expression $x^{2k-3}n^{-x/6}$ can be maximised for $x \in \mathbb{R}^+$ by elementary techniques, and we deduce that
$$x^{2k-3}n^{-x/6} \leq \left(\frac{2k-3}{(e/6)\log n}\right)^{2k-3}.$$

Then we can bound the integral in (3.9) and get

$$\begin{aligned} E_k &= O(1)\log n \left(\frac{e}{2}\frac{\log n}{k-2}\right)^{k-2} \epsilon \left(\frac{2k-3}{(e/6)\log n}\right)^{2k-3} \\ &= O(1)k\left(\frac{36}{2e}\frac{(2k-3)^2}{(k-2)\log n}\right)^{k-2}. \end{aligned}$$

Note that for $k \leq \log n/37$ the expression $k\left(\frac{36}{2e}\frac{(2k-3)^2}{(k-2)\log n}\right)^{k-2}$ is decreasing with respect to k. Hence we can write

$$E_k = O\left(\frac{1}{\log^{\ell+1} n}\right), \qquad \forall k \ : \ \ell + 3 \leq k \leq \frac{1}{37}\log n.$$

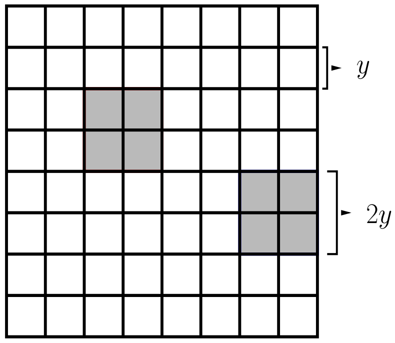


Figure 3.5: The tessellation for counting components of *type* 2 with two particular boxes shaded.

Moreover the bounds $E_{\ell+1} = O(1/\log^{\ell} n)$ and $E_{\ell+2} = O(1/\log^{\ell+1} n)$ are obtained from Lemma 3.2.4, and hence

$$\mathbf{E}M_1 = \sum_{k=\ell+1}^{\frac{1}{37}\log n} E_k = O\left(\frac{1}{\log^{\ell} n}\right) + O\left(\frac{1}{\log^{\ell+1} n}\right) + \frac{\log n}{37} O\left(\frac{1}{\log^{\ell+1} n}\right) = O\left(\frac{1}{\log^{\ell} n}\right),$$

so that $\mathbf{P}(M_1 > 0) \leq \mathbf{E}M_1 = O(1/\log^{\ell} n)$.

Part 2. Consider all the possible components in $G(\mathcal{X}; r)$ which have diameter at most ϵr and size greater than $\log n/37$. Call them components of *type* 2, and let M_2 denote their number.

We tessellate the torus with square cells of side $y = \lfloor(\epsilon r)^{-1}\rfloor^{-1}$ with $y \geq \epsilon r$ but also $y \sim \epsilon r$. We define a box to be a square of side $2y$ consisting of the union of 4 cells of the tessellation. Consider the set of all possible boxes. Note that any component of *type* 2 must be fully contained in some box. (see Figure 3.5).

Let us fix a box b. Let W be the number of vertices which are deployed inside b. Clearly W has a binomial distribution with mean $\mathbf{E}W = (2y)^2 n \sim (2\epsilon)^2 \log n/\pi$. By setting $\delta = \frac{\log n}{37\mathbf{E}W} - 1$ and applying Chernoff inequality to W, we have

$$\mathbf{P}(W > \frac{1}{37}\log n) = \mathbf{P}(W > (1+\delta)\mathbf{E}W) \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mathbf{E}W} = n^{-\frac{(\log(1+\delta) - \frac{\delta}{1+\delta})}{37}}.$$

Note that $\delta \sim \frac{\pi}{148\epsilon^2} - 1 > e^{79}$, and therefore

$$\mathbf{P}(W > \frac{1}{37}\log n) < n^{-2.1}.$$

Then taking a union bound over the set of all $\Theta(r^{-1}) = \Theta(n/\log n)$ boxes, the probability that there is some box with more than $\frac{1}{37}\log n$ vertices is $O(1/(n^{1.1}\log n))$. Since each component of *type* 2 is contained in some box, we have

$$\mathbf{P}(M_2 > 0) = O(1/(n^{1.1}\log n)).$$

Part 3. Consider all the possible components in $G(\mathcal{X}; r)$ which are embeddable and have diameter at least ϵr. Call them components of *type* 3, and let M_3 denote their number.

We tessellate the torus into square cells of side αr, for some $\alpha = \alpha(\epsilon) > 0$ fixed but small enough. The required α will be specified later in the argument. Let Γ be a component of *type* 3. Let $\mathcal{S} = \mathcal{S}_\Gamma$ be the set of all points in the torus $[0,1)^2$ which are at distance at most r from some vertex in Γ. Remove from $\mathcal{S}$ the vertices of Γ and the edges (represented by straight segments) and denote by $\mathcal{S}'$ the outer connected topological component of the remaining set. By construction, $\mathcal{S}'$ must contain no vertex in $\mathcal{X}$ (see Figure 3.6, left picture).

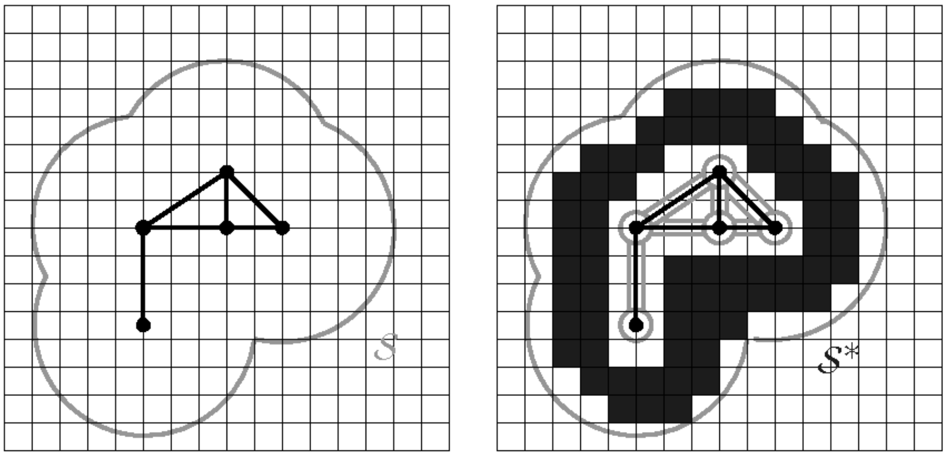


Figure 3.6: The tessellation for counting components of *type* 3

Now let i_{L}, i_{R}, i_{T} and i_{B} be respectively the indices of the leftmost, rightmost, topmost and bottommost vertices in Γ. Some of these indices are possibly equal. Assume w.l.o.g. that the vertical length of Γ (i.e. the vertical distance between $X_{i_{\mathsf{T}}}$ and $X_{i_{\mathsf{B}}}$) is at least $\epsilon r/\sqrt{2}$. Otherwise, the horizontal length of Γ has this property and we can rotate the descriptions in the argument. The upper half-circle with centre $X_{i_{\mathsf{T}}}$ and the lower half-circle with centre $X_{i_{\mathsf{B}}}$ are disjoint and are contained in $\mathcal{S}'$. If $X_{i_{\mathsf{R}}}$ is at greater vertical distanc(i.e., the vertical distance between $X_{i_{\mathsf{T}}}$ and $X_{i_{\mathsf{B}}}$) e from $X_{i_{\mathsf{T}}}$ thanrom $X_{i_{\mathsf{B}}}$, then consider the rectangle of height $\epsilon r/(2\sqrt{2})$ and width $r - \epsilon r/(2\sqrt{2})$ with one corner on $X_{i_{\mathsf{R}}}$ and above and to the right of $X_{i_{\mathsf{R}}}$. Otherwise, consider the same rectangle below and to the right of $X_{i_{\mathsf{R}}}$. This rectangle is also contained in $\mathcal{S}'$ and its interior does not intersect the previously described half-circles. Analogously, we can find another rectangle of height $\epsilon r/(2\sqrt{2})$ and width $r - \epsilon r/(2\sqrt{2})$ to the left of $X_{i_{\mathsf{L}}}$ and either above or below $X_{i_{\mathsf{L}}}$ with the same properties. Hence, taking into account that $\epsilon \le 10^{-18}$, we have

$$\mathsf{Area}(\mathcal{S}') \ge \pi r^2 + 2\left(\frac{\epsilon r}{2\sqrt{2}}\right)\left(r - \frac{\epsilon r}{2\sqrt{2}}\right) > \left(1 + \frac{\epsilon}{5}\right)\pi r^2. \tag{3.10}$$

Let $\mathcal{S}^*$ be the union of all the cells in the tessellation which are fully contained in $\mathcal{S}'$. We loose a bit of area compared to $\mathcal{S}'$. However, if α was chosen small enough, we can guarantee that $\mathcal{S}^*$ is topologically connected and has area $\mathsf{Area}(\mathcal{S}^*) \ge (1 + \epsilon/6)\pi r^2$. This α can be chosen to be the same for all components of *type* 3.

Hence, we showed that the event $(M_3 > 0)$ implies that some connected union of cells $\mathcal{S}^*$ of area $\mathsf{Area}(\mathcal{S}^*) \ge (1 + \epsilon/6)\pi r^2$ contains no vertices. By removing some cells from $\mathcal{S}^*$, we can assume that $(1 + \epsilon/6)\pi r^2 \le \mathsf{Area}(\mathcal{S}^*) < (1 + \epsilon/6)\pi r^2 + \alpha^2 r^2$. Let $\mathcal{S}^*$ be any union of cells with these properties. Note that there are $\Theta(1/r^2) = \Theta(n/\log n)$ many possible choices for $\mathcal{S}^*$. The probability that $\mathcal{S}^*$ contains no vertices is

$$(1 - \mathsf{Area}(\mathcal{S}^*))^n \leq e^{-(1+\epsilon/6)\pi r^2 n} = \left(\frac{\mu}{n}\right)^{1+\epsilon/6}.$$

Therefore, we can take the union bound over all the $\Theta(n/\log n)$ possible sets $\mathcal{S}^*$, and obtain an upper bound of the probability that there is some component of the *type* 3:

$$\mathbf{P}(M_3 > 0) \leq \Theta\left(\frac{n}{\log n}\right)\left(\frac{\mu}{n}\right)^{1+\epsilon/6} = \Theta\left(\frac{1}{n^{\epsilon/6}\log n}\right).$$

Part 4. Consider all the possible components in $G(\mathcal{X}; r)$ which are not embeddable and not solitary either. Call them components of *type* 4, and let M_4 denote their number.

We tessellate the torus $[0,1)^2$ into $\Theta(n/\log n)$ small square cells of side length αr, where $\alpha > 0$ is a sufficiently small positive constant. The required α will be specified later in the argument.

Let Γ be a component of *type* 4. Let $\mathcal{S} = \mathcal{S}_\Gamma$ be the set of all points in the torus $[0,1)^2$ which are at distance at most r from some vertex in Γ. Remove from $\mathcal{S}$ the vertices of Γ and the edges (represented by straight segments) and denote by $\mathcal{S}'$ the remaining set. By construction, $\mathcal{S}'$ must contain no vertex in $\mathcal{X}$.

Suppose there is a horizontal or a vertical band of width $2r$ in $[0,1)^2$ which does not intersect the component Γ. Assume w.l.o.g. that it is the topmost horizontal band consisting of all points with the y-coordinate in $[1-2r, 1)$. Let us divide the torus into vertical bands of width $2r$. All of them must contain at least one vertex of Γ, since otherwise Γ would be embeddable. Select any 9 consecutive vertical bands and pick one vertex of Γ with maximal y-coordinate in each one. For each one of these 9 vertices, we select the left upper quarter-circle centred at the vertex if the vertex is closer to the right side of the band or the right upper quarter-circle otherwise. These nine quarter-circles are disjoint and must contain no vertices by construction. Moreover, they belong to the same connected component of the set $\mathcal{S}'$, which we denote by $\mathcal{S}''$, and which has an area of $\mathsf{Area}(\mathcal{S}'') \geq (9/4)\pi r^2$. Let $\mathcal{S}^*$ be the union of all the cells in the tessellation of the torus which are completely contained in $\mathcal{S}''$. We lose a bit of area compared to $\mathcal{S}''$. However, as usual, by choosing α small enough we can guarantee that $\mathcal{S}^*$ is connected and it has an area of $\mathsf{Area}(\mathcal{S}^*) \geq (11/5)\pi r^2$. Note that this α can be chosen to be the same for all components Γ of this kind.

Suppose otherwise that all horizontal and vertical bands of width $2r$ in $[0,1)^2$ contain at least one vertex of Γ. Since Γ is not solitary it must be possible that it coexists with some other non-embeddable component Γ'. Then all vertical bands or all horizontal bands of width $2r$ must also contain some vertex of Γ'. Assume w.l.o.g. the vertical bands do. Let us divide the torus into vertical bands of width $2r$. We can find a simple path Π with vertices in Γ' which passes through 11 consecutive bands. For each one of the 9 internal bands, pick the uppermost vertex of Γ in the band below Π (in the torus sense). As before each one of these vertices contributes with a disjoint quarter-circle which must be empty of vertices, and by the same argument we obtain a connected union of cells of the tessellation, which we denote by $\mathcal{S}^*$, with $\mathsf{Area}(\mathcal{S}^*) \geq (11/5)\pi r^2$ and containing no vertices.

Hence, we showed that the event $(M_4 > 0)$ implies that some connected union of cells $\mathcal{S}^*$ with $\mathsf{Area}(\mathcal{S}^*) \geq (11/5)\pi r^2$ contains no vertices. By repeating the same argument we used for components of *type* 3, but replacing $(1+\epsilon/6)\pi r^2$ by $(11/5)\pi r^2$, we get

$$\mathbf{P}(M_4 > 0) = \Theta\left(\frac{1}{n^{6/5}\log n}\right). \qquad \square$$

Lemma 3.2.6. *Let $\ell \geq 2$ be a fixed integer. Let $0 < \epsilon < 1/2$ be fixed. Assume that $\mu = \Theta(1)$. Then*

$$\mathbf{E}[K'_{\epsilon,\ell}]_2 = O(1/\log^{2\ell-2} n).$$

Proof. As in the proof of Lemma 3.2.4, we assume that each component Γ which contributes to $K'_{\epsilon,\ell}$ has a unique leftmost vertex X_i, and the vertex X_j in Γ at greatest distance from X_i is also unique. In fact, this happens with probability 1.

Choose any two disjoint subsets of $\{1,\dots,n\}$ of size ℓ each, namely J_1 and J_2, with four distinguished elements $i_1, j_1 \in J_1$ and $i_2, j_2 \in J_2$. For $k \in \{1,2\}$, denote by $\mathcal{Y}_k = \bigcup_{l\in J_k} X_l$ the set of random points in $\mathcal{X}$ with indices in J_k. Let $\mathcal{E}$ be the event that the following conditions hold for both $k=1$ and $k=2$: All vertices in $\mathcal{Y}_k$ are at distance at most ϵr from X_{i_k} and to the right of X_{i_k}; vertex X_{j_k} is the one in $\mathcal{Y}_k$ with greatest distance from X_{i_k}; and the vertices of $\mathcal{Y}_k$ form a component of $G(\mathcal{X};r)$. If $\mathbf{P}(\mathcal{E})$ is multiplied by the number of possible choices of i_k, j_k and the remaining vertices of J_k, we get

$$\mathbf{E}[K'_{\epsilon,\ell}]_2 = O(n^{2\ell})\mathbf{P}(\mathcal{E}). \tag{3.11}$$

In order to bound the probability of $\mathcal{E}$ we need some definitions. For each $k \in \{1,2\}$, let $\rho_k = d(X_{i_k}, X_{j_k})$ and let $\mathcal{S}_k$ be the set of all the points in the torus $[0,1)^2$ which are at distance at most r from some vertex in $\mathcal{Y}_k$. Notice that ρ_k and $\mathcal{S}_k$ depend on the set of random points $\mathcal{Y}_k$. Also define $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$.

Let $\mathcal{F}$ be the event that $d(X_{i_1}, X_{i_2}) > 3r$. This holds with probability $1 - O(r^2)$. In order to bound $\mathbf{P}(\mathcal{E} \mid \mathcal{F})$, we apply a similar approach to the one in the proof of Lemma 3.2.4. In fact, observe that if $\mathcal{F}$ holds then $\mathcal{S}_1 \cap \mathcal{S}_2 = \emptyset$. Therefore in view of (3.4) we can write

$$\pi r^2(2 + (\rho_1+\rho_2)/(6r)) < \mathsf{Area}(\mathcal{S}) < \frac{18\pi}{4} r^2, \tag{3.12}$$

and using the same elementary techniques that gave us (3.5) we get

$$(1 - \mathsf{Area}(\mathcal{S}))^{n-2\ell} < \left(\frac{\mu}{n}\right)^{2+(\rho_1+\rho_2)/(6r)} \frac{1}{(1 - 18\pi r^2/4)^{2\ell}}. \tag{3.13}$$

Observe that $\mathcal{E}$ can also be described as follows: For each $k \in \{1,2\}$ there is some non-negative real $\rho_k \leq \epsilon r$ such that X_{j_k} is placed at distance ρ_k from X_{i_k} and to the right of X_{i_k}; all the remaining vertices in $\mathcal{Y}_k$ are inside the half-circle of centre X_{i_k} and radius ρ_k; and the $n-\ell$ vertices not in $\mathcal{Y}_k$ lie outside $\mathcal{S}_k$. In fact, rather than this last condition we only require for our bound that all vertices in $\mathcal{X} \setminus (\mathcal{Y}_1 \cup \mathcal{Y}_2)$ are placed outside $\mathcal{S}$. Clearly, this has probability $(1 - \mathsf{Area}(\mathcal{S}))^{n-2\ell}$. Then, from (3.13) and following an analogous argument to the one that leads to (3.6), we obtain the bound

$$\begin{aligned}\mathbf{P}(\mathcal{E} \mid \mathcal{F}) &\leq \Theta(1) \int_0^{\epsilon r}\int_0^{\epsilon r} \pi\rho_1 \left(\frac{\pi}{2}\rho_1^2\right)^{\ell-2} \pi\rho_2 \left(\frac{\pi}{2}\rho_2^2\right)^{\ell-2} \frac{1}{n^{2+(\rho_1+\rho_2)/(6r)}}\, d\rho_1 d\rho_2 \\ &= \Theta(1)\, I(1/6)^2,\end{aligned}$$

where $I(1/6)$ is defined in (3.7). Thus from (3.8) we conclude

$$\mathbf{P}(\mathcal{E} \wedge \mathcal{F}) \leq \Theta(1)\, P(\mathcal{F})\, I(1/6)^2 = O\left(\frac{1}{n^{2\ell}\log^{2\ell-2} n}\right). \tag{3.14}$$

Otherwise, suppose that $\mathcal{F}$ does not hold (i.e., $d(X_{i_1}, X_{i_2}) \leq 3r$). Observe that $\mathcal{E}$ implies that $d(X_{i_1}, X_{i_2}) > r$, since X_{i_1} and X_{i_2} must belong to different components. Hence the circles with centres on X_{i_1} and X_{i_2} and radius r have an intersection of area less than $(\pi/2)r^2$. These two circles are contained in $\mathcal{S}$ and then we can write $\mathsf{Area}(\mathcal{S}) \geq (3/2)\pi r^2$. Note that $\mathcal{E}$ implies that all vertices in $\mathcal{X} \setminus (\mathcal{Y}_1 \cup \mathcal{Y}_2)$ are placed outside $\mathcal{S}$ and that for each $k \in \{1, 2\}$ all the vertices in $\mathcal{Y}_k \setminus \{X_{i_k}\}$ are at distance at most ϵr and to the right of X_{i_k}. This gives us the following rough bound

$$\mathbf{P}(\mathcal{E} \mid \overline{\mathcal{F}}) \leq \left(\frac{\pi}{2}(\epsilon r)^2\right)^{2\ell-2}\left(1 - \frac{3\pi}{2}r^2\right)^{n-2\ell} = O(1)\left(\frac{\log n}{n}\right)^{2\ell-2}\left(\frac{\mu}{n}\right)^{3/2}.$$

Multiplying this by $\mathbf{P}(\overline{\mathcal{F}}) = O(r^2) = O(\log n/n)$ we obtain

$$\mathbf{P}(\mathcal{E} \wedge \overline{\mathcal{F}}) = O\left(\frac{\log^{2\ell-1} n}{n^{2\ell+1/2}}\right), \tag{3.15}$$

which is negligible compared to (3.14). The statement of the lemma follows from (3.11), (3.14) and (3.15). □

Theorem 3.2.7. *Let $\ell \geq 2$ be a fixed integer. Let $0 < \epsilon < 1/2$ be fixed. Assume that $\mu = \Theta(1)$. Then*

$$\mathbf{P}(\widetilde{K}_\ell > 0) \sim \mathbf{P}(K_\ell > 0) \sim \mathbf{P}(K'_{\epsilon,\ell} > 0) = \Theta\left(\frac{1}{\log^{\ell-1} n}\right).$$

Proof. From Corollary 1.12 in [15], we have

$$\mathbf{E}K'_{\epsilon,\ell} - \frac{1}{2}\mathbf{E}[K'_{\epsilon,\ell}]_2 \leq \mathbf{P}(K'_{\epsilon,\ell} > 0) \leq \mathbf{E}K'_{\epsilon,\ell},$$

and therefore by Lemmata 3.2.4 and 3.2.6 we obtain

$$\mathbf{P}(K'_{\epsilon,\ell} > 0) = \Theta(1/\log^{\ell-1} n).$$

Combining this with Lemma 3.2.5, yields the statement. □

3.3 Dynamic Properties

We define the dynamic model as follows. Given a positive real $s = s(n)$ and a positive integer $m = m(n)$, we consider the following random process $(\mathcal{X}_t)_{t\in\mathbb{Z}} = (\mathcal{X}_t(n, s, m))_{t\in\mathbb{Z}}$: At time step $t = 0$, n agents are scattered independently and u.a.r. over the torus $[0, 1)^2$, as in the static model. Moreover each agent chooses u.a.r. an angle $\alpha \in [0, 2\pi)$, and moves in the direction of α, travelling distance s at each time step. These directions are changed every m steps for all agents. More formally, for each agent i and for each interval $[t, t + m]$

with $t \in \mathbb{Z}$ divisible by m, an angle in $[0, 2\pi)$ is chosen independently and u.a.r., and this angle determines the direction of i between time steps t and $t+m$. Note that we are also considering negative steps, which is interpreted as if the agents were already moving around the torus ever before step $t = 0$. We extend the notation from the static model, and denote by $X_{i,t} = (x_{i,t}, y_{i,t})$ the position of each agent i at time t. Also let $\mathcal{X}_t = \bigcup_{i=1}^n X_{i,t}$ be the set of positions of the agents at time t. Furthermore, given a positive $r = r(n) \in \mathbb{R}$ such that $r = o(1)$, a random graph process can be derived from $(\mathcal{X}_t)_{t\in\mathbb{Z}}$. For any $t \in \mathbb{Z}$, the vertex set of $G(\mathcal{X}_t; r)$ is $\mathcal{X}_t$, and we join by an edge all pairs of vertices in $\mathcal{X}_t$ which are at Euclidean distance at most r. We derive asymptotic results on $\big(G(\mathcal{X}_t; r)\big)_{t\in\mathbb{Z}}$ as $n \to \infty$.

We use the following lemma proven in [62].

Lemma 3.3.1 (Nain *et al.*)**.** *At any fixed step $t \in \mathbb{Z}$, the vertices are distributed over the torus $[0, 1)^2$ independently and u.a.r. Consequently for any $t \in \mathbb{Z}$, $G(\mathcal{X}_t; r)$ has the same distribution as $G(\mathcal{X}; r)$.*

In the remaining of the section, we focus our attention around the threshold of connectivity obtained in Section 3.2 and we assume that $\mu = \Theta(1)$, or equivalently

$$r = \sqrt{\frac{\log n \pm O(1)}{\pi n}}.$$

In order to prove the main statement of this section, we first require some technical results which involve only two arbitrary consecutive steps t and $t+1$ of $(\mathcal{X}_t)_{t\in\mathbb{Z}}$. In this context t is considered to be an arbitrary fixed integer, and it is often omitted from notation whenever it is understood. Thus for each $i \in \{1, \dots, n\}$, the random positions $X_{i,t}$ and $X_{i,t+1}$ of agent i at times t and $t+1$ are simply denoted by $X_i = (x_i, y_i)$ and $X'_i = (x'_i, y'_i)$. Let also $\mathcal{X} = \mathcal{X}_t$ and $\mathcal{X}' = \mathcal{X}_{t+1}$. Note that the random points X_i and X'_i are not independent. In fact if $2\pi z_i \in [0, 2\pi)$ is the angle in which the agent i moves between times t and $t+1$, then $x'_i = x_i + s\cos(2\pi z_i)$ and $y'_i = y_i + s\sin(2\pi z_i)$, where all the sums involving coordinates are taken $\mod 1$. This motivates an alternative description of the model at times t and $t+1$ in terms of a three dimensional placement of the agents, in which the third dimension is interpreted as a normalised angle. For each $i \in \{1, \dots, n\}$, define the random point $\widehat{X}_i = (x_i, y_i, z_i) \in [0, 1)^3$, and also let $\widehat{\mathcal{X}} = \bigcup_{i=1}^n \widehat{X}_i$. Observe that by Lemma 3.3.1 all the random points $\widehat{X}_i$ are chosen independently and u.a.r. from the 3-torus $[0, 1)^3$, and also that $\widehat{\mathcal{X}}$ encodes all the information of the model at times t and $t+1$. In fact, if we map $[0, 1)^3$ onto $[0, 1)^2$ by the following surjections

$$\begin{aligned} \pi_1 : [0,1)^3 &\to [0,1)^2 & \qquad \pi_2 : [0,1)^3 &\to [0,1)^2 \\ (x,y,z) &\mapsto (x,y) & (x,y,z) &\mapsto (x + s\cos(2\pi z), y + s\sin(2\pi z)), \end{aligned}$$

we can recover the positions of agent i at times t and $t+1$ from $\widehat{X}_i$ and write $X_i = \pi_1(\widehat{X}_i)$ and $X'_i = \pi_2(\widehat{X}_i)$. Moreover, for any measurable set $\mathcal{A} \subseteq [0, 1)^2$, the events $X_i \in \mathcal{A}$ and $X'_i \in \mathcal{A}$ are respectively equivalent to the events $\widehat{X}_i \in \pi_1^{-1}(\mathcal{A})$ and $\widehat{X}_i \in \pi_2^{-1}(\mathcal{A})$ in this new setting. Furthermore, by setting $\mathcal{A}_z = \mathcal{A} - (s\cos(2\pi z), s\sin(2\pi z))$ we get

$$\mathsf{Vol}(\pi_2^{-1}(\mathcal{A})) = \int_{[0,1)} \Big(\int_{\mathcal{A}_z} dxdy \Big) dz = \mathsf{Area}(\mathcal{A}).$$

In addition, observe that $\mathsf{Vol}(\pi_1^{-1}(\mathcal{A})) = \mathsf{Vol}(\mathcal{A} \times [0,1)) = \mathsf{Area}(\mathcal{A})$, and hence we have

$$\mathsf{Area}(\mathcal{A}) = \mathsf{Vol}(\pi_1^{-1}(\mathcal{A})) = \mathsf{Vol}(\pi_2^{-1}(\mathcal{A})). \tag{3.16}$$

In view of Lemma 3.3.1, for any measurable sets $\mathcal{A} \subseteq [0,1)^2$ and $\mathcal{B} \subseteq [0,1)^3$,

$$\mathbf{P}(X_i \in \mathcal{A}) = \mathsf{Area}(\mathcal{A}), \qquad \mathbf{P}(X_i' \in \mathcal{A}) = \mathsf{Area}(\mathcal{A}), \qquad \mathbf{P}(\widehat{X}_i \in \mathcal{B}) = \mathsf{Vol}(\mathcal{B}),$$

which is naturally compatible with (3.16).

Now we define some sets which will repeatedly appear in this section. For each $i \in \{1, \dots, n\}$, consider the sets

$$\mathcal{R}_i = \{X \in [0,1)^2 \ : \ d(X, X_i) \leq r\} \quad \text{and} \quad \mathcal{R}_i' = \{X \in [0,1)^2 \ : \ d(X, X_i') \leq r\},$$

and also let $\widehat{\mathcal{R}}_i = \pi_1^{-1}(\mathcal{R}_i)$ and $\widehat{\mathcal{R}}_i' = \pi_2^{-1}(\mathcal{R}_i')$ be their counterparts in $[0,1)^3$. Note that for $i, j \in \{1, \dots, n\}$, we have that $\widehat{X}_i \in \widehat{\mathcal{R}}_j$ iff $d(X_i, X_j) \leq r$, and similarly that $\widehat{X}_i \in \widehat{\mathcal{R}}_j'$ iff $d(X_i', X_j') \leq r$, where each of these events occurs with probability exactly $\mathsf{Vol}(\widehat{\mathcal{R}}_i) = \mathsf{Vol}(\widehat{\mathcal{R}}_i') = \pi r^2$. Also observe that X_i is isolated in $G(\mathcal{X}; r)$ iff $(\widehat{\mathcal{X}} \setminus \{\widehat{X}_i\}) \cap \widehat{\mathcal{R}}_i = \emptyset$, and that analogously X_i' is isolated in $G(\mathcal{X}'; r)$ iff $(\widehat{\mathcal{X}} \setminus \{\widehat{X}_i\}) \cap \widehat{\mathcal{R}}_i' = \emptyset$. We need the following

Lemma 3.3.2. *Assume $\mu = \Theta(1)$. There exists a constant $\epsilon > 0$ such that for large enough n the following statements are true : For any $i, j \in \{1, \dots, n\}$ (possibly $i = j$),*

(i). if $d(X_i, X_j) > r$ then $\mathsf{Vol}(\widehat{\mathcal{R}}_i \cap \widehat{\mathcal{R}}_j) \leq \frac{\pi}{2} r^2$,

(ii). if $s < r/7$ and $d(X_i, X_j) > r - 2s$ then $\mathsf{Vol}((\widehat{\mathcal{R}}_i \cup \widehat{\mathcal{R}}_i') \cap (\widehat{\mathcal{R}}_j \cup \widehat{\mathcal{R}}_j')) \leq (1-\epsilon)\pi r^2$,

(iii). if $s \geq r/7$ and $s = O(r)$ then $\mathsf{Vol}(\widehat{\mathcal{R}}_i \cap \widehat{\mathcal{R}}_j') \leq (1-\epsilon)\pi r^2$,

(iv). if $s = \omega(r)$ then $\mathsf{Vol}(\widehat{\mathcal{R}}_i \cap \widehat{\mathcal{R}}_j') = O(r^3 \frac{s+1}{s}) = o(r^2)$.

Proof.

Statement (i). Assume w.l.o.g. that the segment $\overline{X_i X_j}$ is vertical and that X_i is above X_j. Let $\mathcal{S} \subset [0,1)^2$ be the upper half-circle with centre X_i and radius r, and $\widehat{\mathcal{S}} = \pi_1^{-1}(\mathcal{S}) = \mathcal{S} \times [0,1) \subset [0,1)^3$. Notice that $\mathsf{Vol}(\widehat{\mathcal{S}}) = \pi r^2/2$, $\widehat{\mathcal{S}} \subset \widehat{\mathcal{R}}_i$ and $\widehat{\mathcal{S}} \cap \widehat{\mathcal{R}}_j = \emptyset$, and the statement follows.

Statement (ii). The distance between X_i' and X_j' is greater than $3r/7$, since $d(X_i', X_j') \geq d(X_i, X_j) - 2s > r - 4s$. Let $\mathcal{S}_i$ (respectively $\mathcal{S}_j$) be the set of points in $[0,1)^2$ at distance at most $8r/7$ from X_i' (respectively X_j') Note that $\mathcal{S}_i$ and $\mathcal{S}_j$ are two circles of radius $8r/7$ with centres at distance greater than $3r/7$. Then straightforward computations show that $\mathsf{Area}(\mathcal{S}_i \cap \mathcal{S}_j)$ is at most $(1-\epsilon)\pi r^2$ for some $\epsilon > 0$. We define $\widehat{\mathcal{S}}_i = \pi_1^{-1}(\mathcal{S}_i)$ and $\widehat{\mathcal{S}}_j = \pi_1^{-1}(\mathcal{S}_j)$. We have, $\widehat{\mathcal{S}}_i \supset \widehat{\mathcal{R}}_i \cup \widehat{\mathcal{R}}_i'$ and $\widehat{\mathcal{S}}_j \supset \widehat{\mathcal{R}}_j \cup \widehat{\mathcal{R}}_j'$. Hence,

$$\mathsf{Vol}((\widehat{\mathcal{R}}_i \cup \widehat{\mathcal{R}}_i') \cap (\widehat{\mathcal{R}}_j \cup \widehat{\mathcal{R}}_j')) \leq \mathsf{Vol}(\widehat{\mathcal{S}}_i \cap \widehat{\mathcal{S}}_j) = \mathsf{Area}(\mathcal{S}_i \cap \mathcal{S}_j) \leq (1-\epsilon)\pi r^2.$$

Statement (iii). Let $k \in \{1, \dots, n\}$ be different from i and j. Observe that $\mathsf{Vol}(\widehat{\mathcal{R}}_i \setminus \widehat{\mathcal{R}}'_j)$ is the probability that $d(X_i, X_k) \le r$ but $d(X'_j, X'_k) > r$. Suppose that $d(X_i, X_k) \le r$ but also $d(X'_j, X_k) > 13r/14$, which happens with probability at least $(1 - 13^2/14^2)\pi r^2$. Let α be the angle of $\overrightarrow{X'_j X_k}$ with respect to the horizontal axis. Recall that agent k moves between time steps t and $t+1$ towards a direction $2\pi z_k$, where z_k is the third coordinate of $\widehat{X}_k$. If $2\pi z_k \in [\alpha - \pi/3, \alpha + \pi/3]$, then the agent increases its distance with respect to X'_j by at least $s/2 \ge r/14$, and thus $d(X'_j, X'_k) > r/14 + 13r/14 = r$. This range of directions has probability 1/3. Summarising, we proved that $\mathsf{Vol}(\widehat{\mathcal{R}}_i \setminus \widehat{\mathcal{R}}'_j) \ge (1 - 13^2/14^2)\pi r^2/3$, and the statement follows.

Statement (iv). Given $k \in \{1, \dots, n\}$ different from i and j, observe that $\mathsf{Vol}(\widehat{\mathcal{R}}_i \cap \widehat{\mathcal{R}}'_j)$ is the probability that $d(X_k, X_i) \le r$ and $d(X'_k, X'_j) \le r$. Suppose first that $s < 1/2$. We claim that the probability that $d(X'_k, X'_j) \le r$ conditional upon any fixed outcome of X_k is at most $(2+\epsilon)r/s$ for some $\epsilon > 0$, no matter which particular point X_k is chosen. In fact, assume $X_k \ne X'_j$ and let α be the angle of $\overrightarrow{X_k X'_j}$ with respect to the horizontal axis. If agent k moves between steps t and $t+1$ towards a direction $2\pi z_k$ not in $[\alpha - \arcsin(r/s), \alpha + \arcsin(r/s)]$ then $d(X'_k, X'_j) > r$. Hence, $\mathsf{Vol}(\widehat{\mathcal{R}}_i \cap \widehat{\mathcal{R}}'_j)$ is at most $\mathbf{P}(d(X_k, X_i) \le r) = \pi r^2$ times $(2+\epsilon)r/s$, which satisfies the claim. The case $X_k = X'_j$ is trivial.

The case $s \ge 1/2$ is a bit more delicate, since agent k may loop many times around the torus while moving between steps t and $t+1$. In fact, as we move along the circumference of radius s centred on X_k, we cross the axes of the torus $\Theta(1+s)$ times. This gives the extra factor $(1+s)$ in the statement, which is negligible when $s = o(1)$ but grows large when $s = \omega(1)$. ☐

For each $i \in \{1, \dots, n\}$, we define $\widehat{\mathcal{Q}}_i := \widehat{\mathcal{R}}'_i \setminus \widehat{\mathcal{R}}_i$ and $\widehat{\mathcal{Q}}'_i := \widehat{\mathcal{R}}_i \setminus \widehat{\mathcal{R}}'_i$. Given any two agents i and j, observe that $\widehat{X}_i \in \widehat{\mathcal{Q}}'_j$ iff $\widehat{X}_j \in \widehat{\mathcal{Q}}'_i$ iff $d(X_i, X_j) \le r$ and $d(X'_i, X'_j) > r$, i.e. the agents are joined by an edge at time t but not at time $t+1$. This holds with probability $\mathsf{Vol}(\widehat{\mathcal{Q}}_i) = \mathsf{Vol}(\widehat{\mathcal{Q}}'_i)$, which does not depend on the particular agents and of t. This probability will be denoted by q hereinafter. The value of this parameter depends on the asymptotic relation between r and s and is given in the following

Lemma 3.3.3. *The probability that two different agents $i, j \in \{1, \dots, n\}$ are at distance at most r at time t but greater than r at time $t+1$ is $q \le \pi r^2$, which also satisfies*

$$q \sim \begin{cases} \frac{4}{\pi} sr & \text{if } s = o(r), \\ \Theta(r^2) & \text{if } s = \Theta(r), \\ \pi r^2 & \text{if } s = \omega(r). \end{cases}$$

Proof. The first bound on q is immediate from the definition of q and the fact that $\mathsf{Vol}(\widehat{\mathcal{R}}_i) = \pi r^2$. In order to obtain the second statement, we consider separate cases.

Case 1 ($s \le \epsilon r$, for some fixed but small enough $\epsilon > 0$). In order to compute the probability that $\widehat{X}_j \in \widehat{\mathcal{Q}}'_i$, we express $\widehat{X}_j = (x_j, y_j, z_j)$ in new coordinates (ρ, θ, z), where $\rho = d(X_j, X'_i)$, θ is the angle between the horizontal axis and $\overrightarrow{X_i X_j}$, and $z = z_j$. Then we integrate an element of volume over the region $\widehat{\mathcal{Q}}'_i$ in terms of these coordinates. Let $\xi = d(X_j, X_i)$, so that (ξ, θ, z) are the usual cylindrical coordinates (see Figure 3.7). From the law of cosines, we can establish the relation between ρ and ξ and write

$$\rho = \sqrt{\xi^2 + s^2 - 2\xi s\cos\theta} \quad \text{and} \quad \xi = \sqrt{\rho^2 - s^2\sin^2\theta} + s\cos\theta. \tag{3.17}$$

Now observe that the minimum value that ρ can take is $r - s$, since X_j must lie outside the

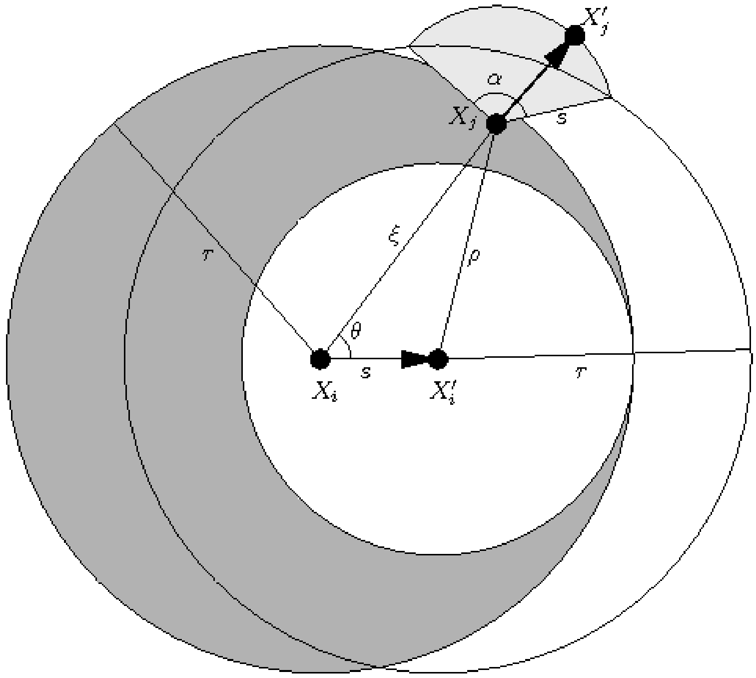


Figure 3.7: Two agents that are at distance at most r at time t but greater than r at time $t+1$

circle of radius $r - s$ and centre X_i'. Otherwise by the triangular inequality $d(X_i', X_j') \le r$, the agents i and j would share an edge at step $t+1$. On the other hand, X_j must lie inside the circle of radius r centred on X_i, and therefore (by setting $\xi = r$ in (3.17)) the maximum value that ρ can achieve is $\sqrt{r^2 + s^2 - 2rs\cos\theta}$.

Moreover, let α be the angle determined from the range of all possible values of $2\pi z$ (i.e., possible directions for agent j to move). Again by the law of cosines,

$$\alpha = 2\arccos\left(\frac{r^2 - s^2 - \rho^2}{2s\rho}\right).$$

Finally from (3.17) and the change of variables formula, we can determine the element of volume in coordinates (ρ, θ, z):

$$dxdydz = \xi\, d\xi d\theta dz = \frac{\xi\rho}{\xi - s\cos\theta}\, d\rho d\theta dz.$$

Using the fact that $r - 2s \le \xi \le r$, we can write

$$\frac{\xi\rho}{\xi - s\cos\theta} = \rho\left(1 \pm O\left(\frac{s}{r}\right)\right).$$

In view of all the above, we deduce

$$
\begin{aligned}
q &= \int_{\widehat{\mathcal{Q}}'_i} dxdydz \\
&= \int_0^{2\pi} \int_{r-s}^{\sqrt{r^2+s^2-2rs\cos\theta}} \frac{\alpha}{2\pi} \frac{\xi\rho}{\xi - s\cos\theta}\, d\rho d\theta \\
&= \left(1 \pm O\left(\frac{s}{r}\right)\right) \int_0^{2\pi} \int_{r-s}^{\sqrt{r^2+s^2-2rs\cos\theta}} \frac{1}{\pi} \arccos\left(\frac{r^2 - s^2 - \rho^2}{2s\rho}\right) \rho\, d\rho d\theta \\
&= \left(1 \pm O\left(\frac{s}{r}\right)\right) 2 \int_0^{\pi} \frac{1}{2\pi} \Bigg(- rs\sin\theta - \theta r^2 \\
&\qquad\qquad + (r^2 + s^2 - 2rs\cos\theta) \arccos \frac{r\cos\theta - s}{\sqrt{r^2 + s^2 - 2rs\cos\theta}} \Bigg) d\theta.
\end{aligned}
$$

Looking at the Taylor series with respect to s/r of the expression inside the integral divided by r^2, we get

$$
\begin{aligned}
q &= \left(1 \pm O\left(\frac{s}{r}\right)\right) \int_0^{\pi} r^2 \left(-\frac{2\theta\cos\theta}{\pi}\frac{s}{r} + O\left(\left(\frac{s}{r}\right)^2\right)\right) d\theta \\
&= \left(1 \pm O\left(\frac{s}{r}\right)\right) \frac{4}{\pi} sr.
\end{aligned} \tag{3.18}
$$

Case 2 ($\epsilon r < s < r/7$). Recall that $\mathcal{R}_i$ is the circle of radius r and centre X_i. Take the chord in $\mathcal{R}_i$ which is perpendicular to the segment $\overline{X_i X'_i}$ and at distance r from X'_i. This chord divides $\mathcal{R}_i$ into two regions. One of them, call it $\mathcal{S}$, has the property that all the points inside are at distance at least r from X'_i and moreover $\mathsf{Area}(\mathcal{S}) \geq \epsilon\sqrt{2\epsilon - \epsilon^2} r^2$. Suppose that $X_j \in \mathcal{S}$ (i.e., the agent j is in $\mathcal{S}$ at time t), which happens with probability at least $\epsilon\sqrt{2\epsilon - \epsilon^2} r^2$. Let us now consider the circle centred on X'_i and passing through X_j. We observe that $d(X'_j, X'_i) > d(X_j, X'_i)$ with probability at least $1/2$, since it is sufficient that the direction $2\pi z_j$ in which agent j moves lies in the outer side of the tangent of that circle at X_j. Therefore, the probability that $d(X_j, X_i) \leq r$ and $d(X'_j, X'_i) > r$, or equivalently $\widehat{X}_j \in \widehat{\mathcal{Q}}'_i$, is at least $\frac{1}{2}\epsilon\sqrt{2\epsilon - \epsilon^2} r^2$.
Case 3 ($s \geq r/7$). We can write

$$
q = \mathsf{Vol}(\widehat{\mathcal{Q}}'_i) = \mathsf{Vol}(\widehat{\mathcal{R}}_i \setminus \widehat{\mathcal{R}}'_i) = \mathsf{Vol}(\widehat{\mathcal{R}}_i) - \mathsf{Vol}(\widehat{\mathcal{R}}_i \cap \widehat{\mathcal{R}}'_i),
$$

and the result follows from the statements (i) and (iv) in Lemma 3.3.2. □

We also need the following technical result, which allows us to compute the probability that a given subset of $[0,1)^3$ contains no points in $\widehat{\mathcal{X}}$, but some other subsets contain at least one. The proof is immediate from Lemma 2.4.1.

Lemma 3.3.4. *For any fixed integer $k \geq 0$, let $\widehat{\mathcal{S}}_0, \ldots, \widehat{\mathcal{S}}_k$ be pairwise disjoint subsets of $[0,1)^3$, with volumes $s_0, \ldots, s_k$ respectively. If $\sum_{i=0}^k s_i = o(N)$, then*

$$
\mathbf{P}\left((\widehat{\mathcal{S}}_0 \cap \widehat{\mathcal{X}} = \emptyset) \wedge \bigwedge_{i=1}^{k} (\widehat{\mathcal{S}}_i \cap \widehat{\mathcal{X}} \neq \emptyset) \right) \sim (1 - s_0)^n \prod_{i=1}^{k} \left(1 - e^{-s_i n}\right).
$$

We are now in position to study the changes experienced by the isolated vertices between two consecutive steps t and $t+1$. Extending the notation in Section 3.2, we denote by $K_{1,t}$ the number of isolated vertices of $G(\mathcal{X}_t; r)$. Also, for any two consecutive steps t and $t+1$, we define the following random variables: B_t is the number of agents i such that X_i is not isolated in $G(\mathcal{X}_t; r)$ but X'_i is isolated in $G(\mathcal{X}_{t+1}; r)$; D_t is the number of agents i such that X_i is isolated in $G(\mathcal{X}_t; r)$ but X'_i is not isolated in $G(\mathcal{X}_{t+1}; r)$; S_t is the number of agents i such that X_i and X'_i are both isolated in $G(\mathcal{X}_t; r)$ and $G(\mathcal{X}_{t+1}; r)$. For simplicity, we often denote them by B, D and S whenever t and $t+1$ are understood. Note that B and D have the same distribution, since any creation of an isolated vertex corresponds to a destruction of an isolated vertex in the time-reversed process and vice versa.

To state the following result, we recall the definition of asymptotic mutual independence from (2.34)

Proposition 3.3.5. *Assume $\mu = \Theta(1)$. Then for any two consecutive steps,*

$$\mathbf{E}B = \mathbf{E}D \sim \mu(1 - e^{-qn}) \quad \text{and} \quad \mathbf{E}S \sim \mu e^{-qn}.$$

Moreover we have that

(i). If $s = o(1/rn)$, then $\mathbf{P}(B > 0) \sim \mathbf{E}B$; $\mathbf{P}(D > 0) \sim \mathbf{E}D$; S is asymptotically Poisson; and $(B > 0)$, $(D > 0)$ and S are asymptotically mutually independent.

(ii). If $s = \Theta(1/rn)$, then B, D and S are asymptotically mutually independent Poisson.

(iii). If $s = \omega(1/rn)$, then B and D are asymptotically Poisson; $\mathbf{P}(S > 0) \sim \mathbf{E}S$; and B, D and $(S > 0)$ are asymptotically mutually independent.

Proof. The central ingredient in the proof is the computation of the joint factorial moments $\mathbf{E}([B]_{\ell_1}[D]_{\ell_2}[S]_{\ell_3})$ of these variables. In particular we find the asymptotic values of $\mathbf{E}B$, $\mathbf{E}D$ and $\mathbf{E}S$. Moreover, In the case $s = \Theta\big(1/(rn)\big)$, we show that for any fixed naturals ℓ_1, ℓ_2 and ℓ_3 we have

$$\mathbf{E}([B]_{\ell_1}[D]_{\ell_2}[S]_{\ell_3}) \sim (\mathbf{E}B)^{\ell_1}(\mathbf{E}D)^{\ell_2}(\mathbf{E}S)^{\ell_3}. \tag{3.19}$$

Then, the result follows from Theorem 1.23 in Bollobás' [15]. The other cases are more delicate since (3.19) does not always hold for extreme values of s, and we obtain a weaker result. In the case $s = o\big(1/(rn)\big)$, we compute the moments for any natural ℓ_3 but only for $\ell_1, \ell_2 \in \{0, 1, 2\}$ and obtain

$$\begin{aligned}
&\mathbf{E}([B]_{\ell_1}[D]_{\ell_2}[S]_{\ell_3}) \sim (\mathbf{E}B)^{\ell_1}(\mathbf{E}D)^{\ell_2}(\mathbf{E}S)^{\ell_3}, \quad \text{if } \ell_1, \ell_2 < 2,\\
&\mathbf{E}([B]_2[D]_{\ell_2}[S]_{\ell_3}) = o(\mathbf{E}(B\,[D]_{\ell_2}[S]_{\ell_3})),\\
&\mathbf{E}([B]_{\ell_1}[D]_2[S]_{\ell_3}) = o(\mathbf{E}([B]_{\ell_1}D\,[S]_{\ell_3})).
\end{aligned} \tag{3.20}$$

From the previous equations and by extending to several variables the upper and lower bounds given in [15], Section 1.4, we deduce that $(B > 0)$, $(D > 0)$ and S satisfy (2.34) and also

$$\mathbf{P}(B > 0) \sim \mathbf{E}B, \quad \mathbf{P}(D > 0) \sim \mathbf{E}D \quad \text{and} \quad \mathbf{P}(S = k) \sim e^{-\mathbf{E}S}\frac{(\mathbf{E}S)^k}{k!} \quad \forall k \in \mathbb{N}.$$

Similarly, for the case $s = \omega\big(1/(rn)\big)$, we compute the moments for any naturals ℓ_1 and ℓ_2 but only for $\ell_3 \in \{0,1,2\}$ and obtain

$$\begin{aligned} &\mathbf{E}([B]_{\ell_1}[D]_{\ell_2}[S]_{\ell_3}) \sim (\mathbf{E}B)^{\ell_1}(\mathbf{E}D)^{\ell_2}(\mathbf{E}S)^{\ell_3}, \quad \text{if } \ell_3 < 2, \\ &\mathbf{E}([B]_{\ell_1}[D]_{\ell_2}[S]_2) = o(\mathbf{E}([B]_{\ell_1}[D]_{\ell_2}S)) \end{aligned} \tag{3.21}$$

From this and by using once more upper and lower bounds given in Section 1.4 of [15], we conclude that B, D and $(S > 0)$ satisfy (2.34) and also

$$\mathbf{P}(B = k) \sim e^{-\mathbf{E}B}\frac{(\mathbf{E}B)^k}{k!} \quad \forall k \in \mathbb{N},$$
$$\mathbf{P}(D = k) \sim e^{-\mathbf{E}D}\frac{(\mathbf{E}D)^k}{k!} \quad \forall k \in \mathbb{N} \quad \text{and} \quad \mathbf{P}(S > 0) \sim \mathbf{E}S.$$

We proceed to compute the moments. First, define for each $i \in \{1,\dots,n\}$ B_i, D_i and S_i as the indicator functions of the following events respectively: X_i is not isolated in $G(\mathcal{X}_t; r)$ but X_i' is isolated in $G(\mathcal{X}_{t+1}; r)$; X_i is isolated in $G(\mathcal{X}_t; r)$ but X_i' is not isolated in $G(\mathcal{X}_{t+1}; r)$; X_i and X_i' are both isolated in $G(\mathcal{X}_t; r)$ and $G(\mathcal{X}_{t+1}; r)$. This allows us to write

$$B = \sum_{i=1}^n B_i, \quad D = \sum_{i=1}^n D_i, \quad S = \sum_{i=1}^n S_i.$$

Note that $B_i = 1$ iff all points in $\widehat{\mathcal{X}} \setminus \{\widehat{X}_i\}$ are outside $\widehat{\mathcal{R}}_i'$ but at least one is inside $\widehat{\mathcal{Q}}_i'$; $D_i = 1$ iff all points in $\widehat{\mathcal{X}} \setminus \{\widehat{X}_i\}$ are outside $\widehat{\mathcal{R}}_i$ but at least one is inside $\widehat{\mathcal{Q}}_i$; and finally $S_i = 1$ iff all points in $\widehat{\mathcal{X}} \setminus \{\widehat{X}_i\}$ are outside $\widehat{\mathcal{R}}_i \cup \widehat{\mathcal{R}}_i' = \widehat{\mathcal{R}}_i \cup \widehat{\mathcal{Q}}_i = \widehat{\mathcal{R}}_i' \cup \widehat{\mathcal{Q}}_i'$.

Now given any fixed naturals ℓ_1, ℓ_2, ℓ_3 with $\ell = \ell_1 + \ell_2 + \ell_3$, we choose an ordered tuple J of ℓ different agents $i_1, \dots, i_\ell \in \{1,\dots,n\}$, and define

$$\mathcal{E} = \bigwedge_{a=1}^{\ell_1}(B_{i_a} = 1) \wedge \bigwedge_{b=\ell_1+1}^{\ell_1+\ell_2}(D_{i_b} = 1) \wedge \bigwedge_{c=\ell_1+\ell_2+1}^{\ell}(S_{i_c} = 1). \tag{3.22}$$

Observe that $\mathbf{P}(\mathcal{E})$ does not depend of the particular tuple J, and multiplying it by the number $[n]_\ell$ of ordered choices of J, we get

$$\mathbf{E}([B]_{\ell_1}[D]_{\ell_2}[S]_{\ell_3}) = [n]_\ell \mathbf{P}(\mathcal{E}) \tag{3.23}$$

By relabelling the agents in J we assume hereinafter that $J = (1,\dots,\ell)$, and we call $\widehat{\mathcal{Y}} = \bigcup_{i=1}^{\ell}\{\widehat{X}_i\}$. Moreover, we define the set

$$\widehat{\mathcal{R}} = \bigcup_{i=1}^{\ell_1}\widehat{\mathcal{R}}_i' \cup \bigcup_{i=\ell_1+1}^{\ell_1+\ell_2}\widehat{\mathcal{R}}_i \cup \bigcup_{i=\ell_1+\ell_2+1}^{\ell}(\widehat{\mathcal{R}}_i \cup \widehat{\mathcal{R}}_i'),$$

and the collection of sets

$$\widehat{\boldsymbol{\mathcal{Q}}} = \{\widehat{\mathcal{Q}}_1', \dots, \widehat{\mathcal{Q}}_{\ell_1}', \widehat{\mathcal{Q}}_{\ell_1+1}, \dots, \widehat{\mathcal{Q}}_{\ell_1+\ell_2}\},$$

which play an important role in the computation of $\mathbf{P}(\mathcal{E})$. It is useful to call $\widehat{\mathcal{Q}}_i^* = \widehat{\mathcal{Q}}_i'$ for $1 \le i \le \ell_1$, $\widehat{\mathcal{Q}}_i^* = \widehat{\mathcal{Q}}_i$ for $\ell_1 + 1 \le i \le \ell_1 + \ell_2$, so that we can write $\widehat{\boldsymbol{\mathcal{Q}}} = \{\widehat{\mathcal{Q}}_1^*, \dots, \widehat{\mathcal{Q}}_{\ell_1+\ell_2}^*\}$.

Case 1 ($s = \Theta\big(1/(rn)\big)$). We say that an agent $i \in J$ is *restricted* if there is some other $j \in J$ with $j > i$ such that $d(X_i, X_j) \le 2r + 4s$. Let $\mathcal{F}$ be the event that there are no restricted agents in J, i.e. $d(X_i, X_j) > 2r + 4s$ for all $i, j \in J$ $(i \ne j)$. This has probability $1 - O(r^2)$. Suppose first that $\mathcal{F}$ holds and compute the probability of $\mathcal{E}$ conditional upon that. We observe that $\mathcal{F}$ implies that for any $i, j \in J$ $(i \ne j)$ we must have $\widehat{\mathcal{R}}_i \cap \widehat{\mathcal{R}}_j = \emptyset$, $\widehat{\mathcal{R}}'_i \cap \widehat{\mathcal{R}}'_j = \emptyset$ and $\widehat{\mathcal{R}}_i \cap \widehat{\mathcal{R}}'_j = \emptyset$. Then $\mathsf{Vol}(\widehat{\mathcal{R}}) = \ell\pi r^2 + \ell_3 q$, and the sets in $\widehat{\boldsymbol{\mathcal{Q}}}$ are pairwise disjoint and also disjoint from $\widehat{\mathcal{R}}$. Moreover observe that, conditional upon $\mathcal{F}$, $\mathcal{E}$ is equivalent to the event that all points in $\widehat{\mathcal{X}} \setminus \widehat{\mathcal{Y}}$ lie outside $\widehat{\mathcal{R}}$, but at least one belongs to each $\widehat{\mathcal{Q}}^*_i \in \widehat{\boldsymbol{\mathcal{Q}}}$. From all the above, the probability of $\mathcal{E}$ can be easily obtained from Lemmata 3.3.3 and 3.3.4:

$$\begin{aligned}\mathbf{P}(\mathcal{E} \wedge \mathcal{F}) &= (1 - O(r^2))\,\mathbf{P}(\mathcal{E} \mid \mathcal{F}) \\ &\sim (1 - \ell\pi r^2 - \ell_3 q)^n (1 - e^{-qn})^{\ell_1+\ell_2} \\ &\sim \left(\frac{\mu}{n}\right)^{\ell} (1 - e^{-qn})^{\ell_1+\ell_2} e^{-\ell_3 qn}. \end{aligned} \tag{3.24}$$

We claim that $\mathbf{P}(\mathcal{E} \wedge \mathcal{F})$ is the main contribution to $\mathbf{P}(\mathcal{E})$. In fact if $\mathcal{F}$ does not hold (i.e. some of the points in $\widehat{\mathcal{Y}}$ are at distance at most $2r + 4s$), then $\mathbf{P}(\mathcal{E} \mid \overline{\mathcal{F}})$ is larger than the expression in (3.24), but this is balanced out by the fact that $\mathbf{P}(\overline{\mathcal{F}})$ is small. Before proving this claim, define $\mathcal{H}$ to be the event that $d(X_i, X_j) > r - 2s$ for all $i, j \in J$ $(i \ne j)$. Notice that $\mathcal{E}$ implies $\mathcal{H}$, since otherwise, for some $i, j \in J$, X_i and X_j would be joined by an edge in $G(\mathcal{X}_t; r)$ and also X'_i and X'_j in $G(\mathcal{X}_{t+1}; r)$, which is not compatible with $\mathcal{E}$. Therefore we only need to see that $\mathbf{P}(\mathcal{E} \wedge \overline{\mathcal{F}}) = \mathbf{P}(\overline{\mathcal{F}} \wedge \mathcal{H})\mathbf{P}(\mathcal{E} \mid \overline{\mathcal{F}} \wedge \mathcal{H})$ is negligible compared to (3.24).

Suppose then that $\mathcal{H}$ holds and also that $p > 0$ of the agents in J are restricted (i.e. $\mathcal{F}$ does not hold). This happens with probability $O(r^{2p})$. In this case, we deduce that $\mathsf{Vol}(\widehat{\mathcal{R}}) \ge (\ell - p)\pi r^2 + \epsilon\pi r^2$, since each unrestricted agent in J contributes at least πr^2 to $\mathsf{Vol}(\widehat{\mathcal{R}})$ and the first restricted one gives by Lemma 3.3.2 (ii) the term $\epsilon\pi r^2$. Moreover, $\mathcal{E}$ implies that all points in $\widehat{\mathcal{X}} \setminus \widehat{\mathcal{Y}}$ lie outside of $\widehat{\mathcal{R}}$, which has probability $\big(1 - \mathsf{Vol}(\widehat{\mathcal{R}})\big)^{n-\ell} = O(1/n^{\ell-p+\epsilon})$. Summarising, the weight in $\mathbf{P}(\mathcal{E} \wedge \overline{\mathcal{F}})$ coming from situations with p restricted agents is $O(r^{2p}/n^{\ell-p+\epsilon}) = O(\log^p n/n^{\ell+\epsilon})$, and is thus negligible compared to (3.24). Hence $\mathbf{P}(\mathcal{E}) \sim \mathbf{P}(\mathcal{E} \wedge \mathcal{F})$, and the required condition on the moments announced in (3.19) follows from (3.23) and (3.24).

Case 2 ($s = o\big(1/(rn)\big)$). Defining $\mathcal{F}$ and $\mathcal{H}$ as in the case $s = \Theta\big(1/(rn)\big)$ and by an analogous argument, we obtain

$$\mathbf{P}(\mathcal{E} \wedge \mathcal{F}) \sim \left(\frac{\mu}{n}\right)^{\ell} (1 - e^{-qn})^{\ell_1+\ell_2} e^{-\ell_3 qn} \sim \left(\frac{\mu}{n}\right)^{\ell} (qn)^{\ell_1+\ell_2} \tag{3.25}$$

However, the analysis of the case that $\mathcal{F}$ does not hold is slightly more delicate here. Indeed, there is an additional $o(1)$ factor in (3.25), namely $(qn)^{\ell_1+\ell_2}$, which forces us to get tighter bounds on $\mathbf{P}(\mathcal{E} \wedge \overline{\mathcal{F}} \wedge \mathcal{H})$ than the ones obtained before. Unlike in the case $s = \Theta\big(1/(rn)\big)$, we also need to consider the role of $\widehat{\boldsymbol{\mathcal{Q}}}$ when $\mathcal{F}$ does not hold, and special care must be taken with several new situations which do not occur otherwise. For instance, since the elements of $\widehat{\boldsymbol{\mathcal{Q}}}$ are not necessarily disjoint, then for $\widehat{\mathcal{Q}}^*_i, \widehat{\mathcal{Q}}^*_j \in \widehat{\boldsymbol{\mathcal{Q}}}$ the condition that both contain some element of $\widehat{\mathcal{X}}$ can be satisfied by having just a single point in $\widehat{\mathcal{Q}}^*_i \cap \widehat{\mathcal{Q}}^*_j \cap \widehat{\mathcal{X}}$. Moreover, if $\ell_1 \ge 2$ and $1 \le i, j \le \ell_1$ (or $\ell_2 \ge 2$ and $\ell_1 + 1 \le i, j \le \ell_1 + \ell_2$), the previous condition is also satisfied if $\widehat{X}_j \in \widehat{\mathcal{Q}}^*_i$, which is equivalent to $\widehat{X}_i \in \widehat{\mathcal{Q}}^*_j$. If the latter situation occurs, we say that i and j *collaborate*.

We first bound the weight in $\mathbf{P}(\mathcal{E}\wedge\overline{\mathcal{F}})$ due to situations in which there are no pairs of elements in J which collaborate. We need some definitions. Let $J_1=\{1,\ldots,\ell_1+\ell_2\}$ and $\widehat{\mathcal{Y}}_1=\bigcup_{i=1}^{\ell_1+\ell_2}\{\widehat{X}_i\}$, and consider the class $\mathscr{P}$ of partitions of J_1. Namely, a partition of J_1 is a collection of subsets of J_1, here denoted *blocks*, which are disjoint and have union J_1. The size of a partition is the number of blocks, and for each block we call *leader* to the maximal element in the block. Given a partition $P=\{A_1,\ldots,A_k\}\in\mathscr{P}$ and also $i_1,\ldots,i_k\in\{1,\ldots,n\}\setminus J$, let $\mathcal{E}_{P,i_1,\ldots,i_k}$ be the following event: For each block A_j of P, we have $\widehat{X}_{i_j}\in\bigcap_{i\in A_j}\widehat{\mathcal{Q}}_i^*$ and moreover all the points in $\widehat{\mathcal{X}}\setminus(\widehat{\mathcal{Y}}\cup\{i_1,\ldots,i_k\})$ lie outside of $\widehat{\mathcal{R}}$. We wish to bound the probability of $\mathcal{E}_{P,i_1,\ldots,i_k}\wedge\overline{\mathcal{F}}\wedge\mathcal{H}$. Notice that if $\mathcal{E}_{P,i_1,\ldots,i_k}$ holds, then all the $\ell_1+\ell_2-k$ non-leader elements in J_1 must be restricted, and possibly some other p' agents in J are restricted too. Moreover, $\mathcal{F}$ does not hold iff this p' satisfies $0<\ell_1+\ell_2-k+p'<\ell$. Given any p' with that property, suppose that p' is exactly the number of restricted agents in J which are either in $J\setminus J_1$ or are leaders of some block. We condition upon this and also upon $\mathcal{H}$, which has probability $r^{2p'}$. Then for each block A_j with leader l_j, event $\mathcal{E}_{P,i_1,\ldots,i_k}$ requires that $\widehat{X}_{i_j}\in\widehat{\mathcal{Q}}_{l_j}^*$ and for all $i\in A_j$ $(i\neq l_j)$ $\widehat{X}_i\in(\widehat{\mathcal{Q}}_{i_j}\cup\widehat{\mathcal{Q}}'_{i_j})$. In addition, since the number of restricted agents in J is $\ell_1+\ell_2-k+p'>0$, arguing as in the case $s=\Theta\big(1/(rn)\big)$, we have $\mathsf{Vol}(\widehat{\mathcal{R}})\geq(\ell_3+k-p')\pi r^2+\epsilon\pi r^2$. Then the contribution to $\mathbf{P}(\mathcal{E}_{P,i_1,\ldots,i_k}\wedge\overline{\mathcal{F}}\wedge\mathcal{H})$ for this particular p' is

$$O(r^{2p'})q^k(2q)^{\ell_1+\ell_2-k}(1-\mathsf{Vol}(\widehat{\mathcal{R}}))^{n-\ell-k}=O\left(\frac{\log^{p'}n}{n^{\ell+k+\epsilon}}\right)(qn)^{\ell_1+\ell_2},$$

so for some $0<\epsilon'<\epsilon$, we can write

$$\mathbf{P}(\mathcal{E}_{P,i_1,\ldots,i_k}\wedge\overline{\mathcal{F}}\wedge\mathcal{H})=O\left(\frac{1}{n^{\ell+k+\epsilon'}}\right)(qn)^{\ell_1+\ell_2}.$$

Finally observe that if there are no pairs of elements in J which collaborate, then $\mathcal{E}\wedge\overline{\mathcal{F}}$ implies that $\mathcal{E}_{P,i_1,\ldots,i_k}\wedge\overline{\mathcal{F}}\wedge\mathcal{H}$ holds for some $P\in\mathscr{P}$ of size k and some $i_1,\ldots,i_k\in\{1,\ldots,n\}\setminus J$, and therefore has probability

$$O\left(n^k\right)O\left(\frac{1}{n^{\ell+k+\epsilon'}}\right)(qn)^{\ell_1+\ell_2}=O\left(\frac{1}{n^{\ell+\epsilon'}}\right)(qn)^{\ell_1+\ell_2},\tag{3.26}$$

which is negligible compared to (3.25). In particular, if $\ell_1,\ell_2<2$, then no pair of elements in J collaborates and then $\mathbf{P}(\mathcal{E})\sim\mathbf{P}(\mathcal{E}\wedge\mathcal{F})$. Hence, the first line of (3.20) follows from (3.23) and (3.25).

We extend the approach above to deal with situations in which some pair of elements in J collaborate. Unfortunately, their contribution to $\mathbf{P}(\mathcal{E}\wedge\overline{\mathcal{F}}\wedge\mathcal{H})$ may be larger than (3.25) if s tends to 0 fast. Hence we restrict ℓ_1 and ℓ_2 to be at most 2 and prove only (3.20). If $\ell_1=2$ let $\mathcal{E}_1$ be the following event: $\widehat{X}_1\in\widehat{\mathcal{Q}}'_2$; $\widehat{\mathcal{R}}$ contains no points in $\widehat{\mathcal{X}}\setminus\widehat{\mathcal{Y}}$; and for each natural $(3\leq i\leq 2+\ell_2)$, $\widehat{\mathcal{Q}}_i$ contains some point in $\widehat{\mathcal{X}}\setminus\widehat{\mathcal{Y}}$. Similarly if $\ell_2=2$ let $\mathcal{E}_2$ be the following event: $\widehat{X}_{\ell_1+1}\in\widehat{\mathcal{Q}}_{\ell_1+2}$; $\widehat{\mathcal{R}}$ contains no points in $\widehat{\mathcal{X}}\setminus\widehat{\mathcal{Y}}$; and for each natural i $(1\leq i\leq\ell_1)$, $\widehat{\mathcal{Q}}'_i$ contains some point in $\widehat{\mathcal{X}}\setminus\widehat{\mathcal{Y}}$. Finally if $\ell_1=\ell_2=2$ let $\mathcal{E}_{1,2}$ be the following event: $\widehat{X}_1\in\widehat{\mathcal{Q}}'_2$ and $\widehat{X}_3\in\widehat{\mathcal{Q}}_4$. In order to compute $\mathbf{P}(\mathcal{E}_1\wedge\mathcal{H})$, we can repeat the same argument as above, but imposing that $\widehat{X}_1\in\widehat{\mathcal{Q}}'_2$ and ignoring other conditions on $\widehat{\mathcal{Q}}'_1$ and

$\widehat{\mathcal{Q}}'_2$. We obtain that for some $\epsilon' > 0$

$$\mathbf{P}(\mathcal{E}_1 \wedge \mathcal{H}) = O\left(\frac{1}{n^{\ell-1+\epsilon'}}\right) q(qn)^{\ell_2} = O\left(\frac{1}{n^{\ell+\epsilon'}}\right)(qn)^{1+\ell_2}, \tag{3.27}$$

and similarly

$$\mathbf{P}(\mathcal{E}_2 \wedge \mathcal{H}) = O\left(\frac{1}{n^{\ell+\epsilon'}}\right)(qn)^{\ell_1+1} \quad \text{and} \quad \mathbf{P}(\mathcal{E}_{1,2} \wedge \mathcal{H}) = O\left(\frac{1}{n^{\ell+\epsilon'}}\right)(qn)^2. \tag{3.28}$$

Observe that if some agents in J collaborate, then $\mathcal{E}\wedge\overline{\mathcal{F}}$ implies that $\mathcal{E}_1\wedge\mathcal{H}$, $\mathcal{E}_2\wedge\mathcal{H}$ or $\mathcal{E}_{1,2}\wedge\mathcal{H}$ hold. Unfortunately, from (3.25), (3.27) and (3.28) we cannot guarantee that $\mathbf{P}(\mathcal{E} \wedge \overline{\mathcal{F}})$ is smaller than $\mathbf{P}(\mathcal{E} \wedge \mathcal{F})$, but in any case, by multiplying these probabilities by $[n]_\ell$ in view of (3.23), we complete the proof of (3.20).

Case 3 ($s = \omega\big(1/(rn)\big)$ but also $s = O(r)$). Following the same notation as in the case $s = \Theta\big(1/(rn)\big)$ and by an analogous argument we obtain

$$\mathbf{P}(\mathcal{E} \wedge \mathcal{F}) \sim \left(\frac{\mu}{n}\right)^\ell (1 - e^{-qn})^{\ell_1+\ell_2} e^{-\ell_3 qn} \sim \left(\frac{\mu}{n}\right)^\ell e^{-\ell_3 qn} \tag{3.29}$$

If $\ell_3 \leq 1$, we claim that this is the main contribution to $\mathbf{P}(\mathcal{E})$. In fact, suppose that $\mathcal{H}$ holds and also that $p > 0$ of the agents in J are restricted (i.e. $\mathcal{F}$ does not hold). This happens with probability $O(r^{2p})$. Since $\ell_3 \leq 1$, then the only possible event which contributes to S required in the definition of $\mathcal{E}$ is $(S_\ell = 1)$ (cf. (3.22)). This involves agent ℓ which cannot be restricted by definition. Then we deduce that $\mathsf{Vol}(\widehat{\mathcal{R}}) \geq (\ell - p)\pi r^2 + \ell_3 q + \epsilon\pi r^2$, since the unrestricted agents in J contribute $(\ell - p)\pi r^2 + \ell_3 q$ to $\mathsf{Vol}(\widehat{\mathcal{R}})$ and the first restricted one gives the term $\epsilon\pi r^2$, by Lemma 3.3.2 (ii) and (iii). Therefore, the probability of $\mathcal{E}$ in this situation is $O(e^{-\ell_3 qn}/n^{\ell-p+\epsilon})$, which combined with the probability $O(r^{2p})$ that p agents are restricted has negligible weight compared to (3.29). Hence, $\mathbf{P}(\mathcal{E}) \sim \mathbf{P}(\mathcal{E} \wedge \mathcal{F})$, and the first line of (3.21) follows from (3.23) and (3.29).

Unfortunately, if $\ell_3 = 2$ and we have p restricted agents in J, we can only assure that $\mathsf{Vol}(\widehat{\mathcal{R}}) \geq (\ell - p)\pi r^2 + q + \epsilon\pi r^2$, and then for some $0 < \epsilon' < \epsilon$

$$\mathbf{P}(\mathcal{E} \wedge \overline{\mathcal{F}}) = O\left(\frac{r^{2p}}{n^{\ell-p+\epsilon}}\right) e^{-qn} = O\left(\frac{1}{n^{\ell+\epsilon'}}\right) e^{-qn}, \tag{3.30}$$

which may have significant contribution to $\mathbf{P}(\mathcal{E})$ if s is large enough. But in any case, in view of (3.23), (3.29) and (3.30), we verify that the second line of (3.21) is satisfied.

Case 4 ($s = \omega(r)$). Let $\mathcal{F}'$ be the event that for any $i, j \in J$ $(i \neq j)$ we have that $d(X_i, X_j) > 2r$ and $d(X'_i, X'_j) > 2r$. This event has probability $1 - O(r^2)$. Observe that if $\mathcal{F}'$ holds, then for any $i, j \in J$ $(i \neq j)$ we must have $\widehat{\mathcal{R}}_i \cap \widehat{\mathcal{R}}_j = \emptyset$, $\widehat{\mathcal{R}}'_i \cap \widehat{\mathcal{R}}'_j = \emptyset$ and $\widehat{\mathcal{R}}_i \cap \widehat{\mathcal{R}}'_j = \emptyset$. Therefore, $\mathsf{Vol}(\widehat{\mathcal{R}}) = \ell\pi r^2 + \ell_3 q$ and the sets in $\widehat{\boldsymbol{\mathcal{Q}}}$ are pairwise disjoint and also disjoint from $\widehat{\mathcal{R}}$. Then in view of Lemmata 3.3.3 and 3.3.4, and by the same argument that leads to (3.24)

$$\mathbf{P}(\mathcal{E} \wedge \mathcal{F}') \sim \left(\frac{\mu}{n}\right)^\ell (1 - e^{-qn})^{\ell_1+\ell_2} e^{-\ell_3 qn} \sim \left(\frac{\mu}{n}\right)^\ell e^{-\ell_3 qn}. \tag{3.31}$$

The remaining of the argument is analogous to Case 3 but replacing $\mathcal{F}$ with $\mathcal{F}'$ and using Lemma 3.3.2 (iv). ☐

Taking into account that $K_{1,t} = D_t + S_t$ and $K_{1,t+1} = S_t + B_t$, Proposition 3.3.5 completely characterises the number of isolated vertices at two consecutive steps in the case $s = \Theta\big(1/(rn)\big)$. For the other ranges of s, the result is weaker but still sufficient for our further purposes. We remark that if $s = o(1/(rn)$, then creations and destructions of isolated vertices are rare, but a Poisson number of isolated vertices is present at both consecutive steps. Otherwise if $s = \omega(1/(rn)$, then the isolated vertices which are present at both consecutive steps are rare since, but a Poisson number of them are created and also a Poisson number destroyed.

Now in order to characterise the connectivity of $\big(G(\mathcal{X}_t;r)\big)_{t\in\mathbb{Z}}$, we need to bound the probability of the event that components other than isolated vertices and the giant one appear at some step. We know by Theorem 3.2.2 that a.a.s. this does not occur at one single step t. However during long periods of time this event could affect the connectivity and must be considered.

Extending the notation in Section 3.2, given a step t let $\widetilde{K}_{2,t}$ be the number of non-solitary components other than isolated vertices occurring at step t. We show that they have a negligible effect compared to isolated vertices in the dynamic evolution of connectivity.

Lemma 3.3.6. *Assume that* $\mu = \Theta(1)$ *and* $s = o\big(1/(rn)\big)$. *Then,*

- $\mathbf{P}(\widetilde{K}_{2,t} > 0 \wedge \widetilde{K}_{2,t+1} = 0) = \mathbf{P}(\widetilde{K}_{2,t} = 0 \wedge \widetilde{K}_{2,t+1} > 0) = o(srn)$,
- $\mathbf{P}(\widetilde{K}_{2,t} > 0 \wedge B_t > 0) = o(srn)$.

Proof. Recall from Lemma 3.3.3 that if $s = o\big(1/(rn)\big)$ then $q = \Theta(rs)$. It suffices to prove that $\mathbf{P}(\widetilde{K}_{2,t} > 0 \wedge \widetilde{K}_{2,t+1} = 0) = o(qn)$ and $\mathbf{P}(\widetilde{K}_{2,t} > 0 \wedge B_t > 0) = o(qn)$, since $(\widetilde{K}_{2,t} = 0 \wedge \widetilde{K}_{2,t+1} > 0)$ corresponds in the time-reversed process to $(\widetilde{K}_{2,t} > 0 \wedge \widetilde{K}_{2,t+1} = 0)$, and thus they have the same probability.

Consider all the possible components in $G(\mathcal{X};r)$ which are not solitary and have size at least 2. They are classified into several types according to their size and diameter, and we deal with each type separately. Then if we denote by M_i the number of components of *type* i in $G(\mathcal{X}_t;r)$, we must show for each i that

$$\mathbf{P}(M_i > 0 \wedge \widetilde{K}_{2,t+1} = 0) = o(qn) \quad \text{and} \quad \mathbf{P}(M_i > 0 \wedge B_t > 0) = o(qn). \tag{3.32}$$

Also we need one definition which helps to describe the changes of edges between $G(\mathcal{X}_t;r)$ and $G(\mathcal{X}_{t+1};r)$. For each $i \in \{1,\dots,n\}$ we define $\widehat{\mathcal{P}}_i = \widehat{\mathcal{Q}}_i \cup \widehat{\mathcal{Q}}'_i = \widehat{\mathcal{R}}_i \Delta \widehat{\mathcal{R}}'_i$ (where Δ denotes the symmetric difference of sets). Given also $j \in \{1,\dots,n\}$, see that $\widehat{X}_j \in \widehat{\mathcal{P}}_i$ iff $\widehat{X}_i \in \widehat{\mathcal{P}}_j$ iff agents i and j share an edge either at time t or at time $t+1$ but not at both times, which happens with probability $\mathsf{Vol}(\widehat{\mathcal{P}}_i) = 2q$.

Each part in this proof is labelled by a number followed by a prime ($'$) in order to avoid confusion with the parts in the proof of Lemma 3.2.5, which are often referred to. Moreover, we write for simplicity Part i (p.L. 3.2.5) to denote Part i in the proof of Lemma 3.2.5.

We set throughout this proof $\epsilon = 10^{-18}$.

Part $1'$. Consider all the possible components in $G(\mathcal{X};r)$ which have diameter at most ϵr and size between 2 and $\log n/37$. Call them components of *type* 1, and let M_1 denote their number at time t. This definition is similar to the one in Part 1 (p.L. 3.2.5), but also includes components of size 2, covered by Lemma 3.2.4.

Given any $i \in \{1,\dots,n\}$, let $\mathcal{E}_i$ be the following event: There exists a component Γ of *type* 1 in $G(\mathcal{X} \setminus \{X_i\}; r)$ and moreover for some $j \in \{1,\dots,n\}$ such that X_j is a vertex of Γ we have that $\widehat{X}_i \in \widehat{\mathcal{P}}_j$. In order to compute the probability of $\mathcal{E}_i$, note that the arguments in the proofs of Lemmata 3.2.4 and 3.2.5 are still valid if we replace $\mathcal{X}$ by $\mathcal{X} \setminus \{X_i\}$ (i.e. we ignore agent i in the model). Hence, the probability of having some component in $G(\mathcal{X} \setminus \{X_i\}; r)$ of *type* 1 and size at least $\ell \geq 2$ is $O(1/\log^{\ell-1} n)$. Suppose first that $G(\mathcal{X} \setminus \{X_i\}; r)$ has some component Γ of *type* 1 and size between 3 and $\log n/37$. This happens with probability $O(1/\log^2 n)$. Conditional upon this, the probability that $\widehat{X}_i \in \widehat{\mathcal{P}}_j$ for some $j \in \{1,\dots,n\}$ with X_j being a vertex of Γ is at most $\log n/37$ times $2q$. This contributes to the probability of $\mathcal{E}_i$ by $O(1/\log^2 n)(\log n/37)(2q) = O(q/\log n)$. Otherwise suppose that $G(\mathcal{X} \setminus \{X_i\}; r)$ has some component Γ of *type* 1 and size exactly 2. This happens with probability $O(1/\log n)$. Conditional upon this, the probability that $\widehat{X}_i \in \widehat{\mathcal{P}}_j$ for some $j \in \{1,\dots,n\}$ with X_j being a vertex of Γ is at most two times $2q$. This also contributes to the probability of $\mathcal{E}_i$ by $O(1/\log n)(4q) = O(q/\log n)$, and therefore $\mathbf{P}(\mathcal{E}_i) = O(q/\log n)$.

Given any $i_1, i_2 \in \{1,\dots,n\}$ $(i_1 \neq i_2)$, let $\mathcal{F}_{i_1,i_2}$ be the following event: There exists a component Γ of *type* 1 in $G(\mathcal{X}\setminus\{X_{i_2}\}; r)$ and moreover $\widehat{\mathcal{R}}'_{i_1} \cap (\widehat{\mathcal{X}} \setminus \{\widehat{X}_{i_1}, \widehat{X}_{i_2}\}) = \emptyset$. To derive the probability of $\mathcal{F}_{i_1,i_2}$, we distinguish two cases according to the distance between X_{i_1} and Γ. Suppose first that for some $h \in \{1,\dots,n\} \setminus \{i_1, i_2\}$ we have that $r < d(X_{i_1}, X_h) \leq 3r$, which happens with probability $O(r^2) = O(\log n/n)$. Let $\mathcal{S}_h$ be the set of points in $[0,1)^2$ at distance greater than ϵr but at most r from X_h, and let $\mathcal{S}_{i_1}$ be the circle with centre X_{i_1} and radius $r-2s$. At least one half-circle of $\mathcal{S}_{i_1}$ has all points at distance greater than r from X_h, so $\mathsf{Area}(\mathcal{S}_h \cup \mathcal{S}_{i_1}) \geq (1-\epsilon^2)\pi r^2 + \pi(r-2s)^2/2 \geq (5/4)\pi r^2$. Notice that, if $\mathcal{F}_{i_1,i_2}$ holds for some component Γ which contains a vertex X_h such that $d(X_{i_1}, X_h) \leq 3r$, then we must have $d(X_{i_1}, X_h) > r$ and moreover $\mathcal{S}_h \cup \mathcal{S}_{i_1}$ must contain no point in $\mathcal{X} \setminus \{X_{i_1}, X_{i_2}\}$, which occurs with probability $(1 - \mathsf{Area}(\mathcal{S}_h \cup \mathcal{S}_{i_1}))^{n-2} = O(1/n^{5/4})$. Therefore, multiplying this by the probability that $d(X_{i_1}, X_h) \leq 3r$ and taking the union bound over the $n-2$ possible choices of h, the contribution to $\mathbf{P}(\mathcal{F}_{i_1,i_2})$ due to situations of this type is $O(n(\log n/n)/n^{5/4}) = O(\log n/n^{5/4})$, and in particular is $O(1/(n \log n))$. On the other hand, we claim that the probability that $\mathcal{F}_{i_1,i_2}$ holds for some component Γ with all vertices at distance greater than $3r$ from X_{i_1} is also $O(1/(n\log n))$. In order to prove this last claim, we follow all the notation in the proof of Lemma 3.2.4 for the remaining of the paragraph, and also define $\widehat{\mathcal{S}} = \pi_1^{-1}(\mathcal{S})$ and $\widehat{\mathcal{Y}} = \pi_1^{-1}(\mathcal{Y})$. We can repeat the same computations the proof of Lemma 3.2.4 but, instead of asking that all the $n-\ell$ points in $\mathcal{X} \setminus \mathcal{Y}$ lie outside of $\mathcal{S}$, we require that all the $n-\ell-2$ points in $\widehat{\mathcal{X}} \setminus (\widehat{\mathcal{Y}} \cup \{\widehat{X}_{i_1}, \widehat{X}_{i_2}\})$ lie outside of $\widehat{\mathcal{S}} \cup \widehat{\mathcal{R}}'_{i_1}$. This last fact occurs with probability $\widehat{P} = (1 - \mathsf{Vol}(\widehat{\mathcal{S}} \cup \widehat{\mathcal{R}}'_{i_1}))^{n-\ell-2}$, which plays a role analogous to that of P in the proof of Lemma 3.2.4. If X_{i_1} is at distance greater than $3r$ from any point in $\mathcal{Y}$, then $\widehat{\mathcal{S}}$ and $\widehat{\mathcal{R}}'_{i_1}$ are disjoint. Therefore from (3.4) we get

$$\pi r^2 \left(2 + \frac{1}{6}\frac{\rho}{r}\right) < \mathsf{Vol}(\widehat{\mathcal{S}} \cup \widehat{\mathcal{R}}'_{i_1}) < \frac{13\pi}{4} r^2, \tag{3.33}$$

and an argument analogous to that leading to (3.5) shows that

$$\widehat{P} < \left(\frac{\mu}{n}\right)^{2+\rho/(6r)} \frac{1}{(1-13\pi r^2/4)^{\ell+1}}. \tag{3.34}$$

Repeating the same computations in the proof of Lemma 3.2.4, but replacing P with $\widehat{P}$, proves the claim for components of *type* 1 of fixed size $\ell \geq 2$. This is extended to all components of *type* 1 by arguing as in Part 1 (p.L. 3.2.5). As a result, we conclude that $\mathbf{P}(\mathcal{F}_{i_1,i_2}) = O(1/(n\log n))$.

Now we proceed to prove (3.32) for components of *type* 1. First observe that the event $(M_1 > 0 \wedge \widetilde{K}_{2,t+1} = 0)$ implies that $\mathcal{E}_i$ holds for some $i \in \{1,\dots,n\}$, since the only way for a component of *type* 1 to disappear within one time step is getting joined to something else. Therefore,

$$\mathbf{P}(M_1 > 0 \wedge \widetilde{K}_{2,t+1} = 0) \leq \sum_{i=1}^{n} \mathbf{P}(\mathcal{E}_i) = O\left(\frac{qn}{\log n}\right).$$

Notice that $(M_1 > 0 \wedge B_t > 0)$ implies that $\mathcal{F}_{i_1,i_2}$ holds and moreover $\widehat{X}_{i_2} \in \widehat{\mathcal{Q}}'_{i_1}$, for some $i_1, i_2 \in \{1,\dots,n\}$ $(i_1 \neq i_2)$. Then,

$$\mathbf{P}(M_1 > 0 \wedge B_t > 0) \leq \sum_{i_1,i_2} \mathbf{P}\big(\mathcal{F}_{i_1,i_2} \wedge (\widehat{X}_{i_2} \in \widehat{\mathcal{Q}}'_{i_1})\big) = O\left(\frac{n^2 q}{n\log n}\right) = O\left(\frac{qn}{\log n}\right).$$

Part $2'$. Consider all the possible components in $G(\mathcal{X};r)$ which have diameter at most ϵr and size greater than $\log n/37$. Call them components of *type* 2, and let M_2 denote their number at time t.

Repeat the same tessellation of $[0,1)^2$ into cells as in Part 2 (p.L. 3.2.5), and also consider the set of square boxes defined there. Given any box b and $i,j \in \{1,\dots,n\}$ $(i \neq j)$, we define $\mathcal{E}_{b,i,j}$ to be the event that box b contains more than $\log n/37 - 1$ points of $\mathcal{X} \setminus \{X_i\}$ and moreover $\widehat{X}_i \in \widehat{\mathcal{P}}_j$. Observe that each of the events $(M_2 > 0 \wedge \widetilde{K}_{2,t+1} = 0)$ and $(M_2 > 0 \wedge B_t > 0)$ implies that $\mathcal{E}_{b,i,j}$ holds for some box b and $i,j \in \{1,\dots,n\}$. Then, by repeating the argument in Part 2 (p.L. 3.2.5), but ignoring X_i and also replacing $\log n/37$ with $\log n/37 - 1$, we deduce that

$$\mathbf{P}(M_2 > 0 \wedge \widetilde{K}_{2,t+1} = 0) \leq O\left(\frac{1}{n^{1.1}\log n}\right) \sum_{i,j} \mathbf{P}(\widehat{X}_j \in \widehat{\mathcal{P}}_i) = O\left(\frac{qn}{n^{0.1}\log n}\right),$$

and the same bound applies to $\mathbf{P}(M_2 > 0 \wedge B_t > 0)$.

Part $3'$. Consider all the possible components in $G(\mathcal{X};r)$ which are not embeddable and not solitary. Call them components of *type* 4, and let M_4 denote their number at time t.

Repeat the same tessellation of $[0,1)^2$ into cells as in Part 4 (p.L. 3.2.5), and observe that each of the events $(M_4 > 0 \wedge \widetilde{K}_{2,t+1} = 0)$ and $(M_4 > 0 \wedge B_t > 0)$ implies that for some $i,j \in \{1,\dots,n\}$ there exists some connected union $\mathcal{S}^*$ of cells in the tessellation with $\mathsf{Area}(\mathcal{S}^*) \geq (11/5)\pi r^2$ such that $\mathcal{S}^* \cap (\mathcal{X} \setminus \{X_i\}) = \emptyset$ and moreover $\widehat{X}_i \in \widehat{\mathcal{P}}_j$. Hence, from Part 4 (p.L. 3.2.5) but replacing $\mathcal{X}$ with $\mathcal{X} \setminus \{X_i\}$, we obtain

$$\mathbf{P}(M_4 > 0 \wedge \widetilde{K}_{2,t+1} = 0) \leq O\left(\frac{1}{n^{6/5}\log n}\right) \sum_{i,j} \mathbf{P}(\widehat{X}_j \in \widehat{\mathcal{P}}_i) = O\left(\frac{qn}{n^{1/5}\log n}\right),$$

and the same bound applies to $\mathbf{P}(M_4 > 0 \wedge B_t > 0)$.

Part $4'$. The embeddable components with diameter at least ϵr treated in Part 3 (p.L. 3.2.5) are here divided into two types. First consider all the possible components in $G(\mathcal{X};r)$ of diameter between ϵr and $6\sqrt{2}r$. Call them components of *type* 3a, and let M_{3a} denote their number at time t.

We tessellate the torus $[0,1)^2$ into square cells of side αr, for some fixed but small enough $\alpha > 0$. From Part 3 (p.L. 3.2.5), if $G(\mathcal{X}_t; r)$ has some component of this type, then there exists a topologically connected union $\mathcal{S}^*$ of cells with $\mathsf{Area}(\mathcal{S}^*) \geq (1+\epsilon/6)\pi r^2$ which contains no point in $\mathcal{X}$. By removing some extra cells from $\mathcal{S}^*$, we can assume that the number of cells in $\mathcal{S}^*$ is exactly $\lceil \frac{(1+\epsilon/6)\pi}{\alpha^2} \rceil$. Now for each $i, j \in \{1, \ldots, n\}$ and each union $\mathcal{S}^*$ of $\lceil \frac{(1+\epsilon/6)\pi}{\alpha^2} \rceil$ cells that is topologically connected, let $\mathcal{E}_{i,j,\mathcal{S}^*}$ be the following event: $\mathcal{S}^*$ contains no points in $\mathcal{X} \setminus \{X_i, X_j\}$, X_j is at distance at least $2r$ from all the points in $\mathcal{S}^*$; $\widehat{\mathcal{R}}'_j$ contains no points in $\widehat{\mathcal{X}} \setminus \{\widehat{X}_i, \widehat{X}_j\}$; and moreover $\widehat{X}_i \in \widehat{\mathcal{P}}_j$. Notice that if X_j is at distance at least $2r$ from all the points in $\mathcal{S}^*$, then $\pi_1^{-1}(\mathcal{S}^*)$ and $\widehat{\mathcal{R}}'_j$ are disjoint. Hence, $\mathsf{Vol}(\pi_1^{-1}(\mathcal{S}^*) \cup \widehat{\mathcal{R}}'_j) \geq (2+\epsilon/6)\pi r^2$ and

$$\mathbf{P}(\mathcal{E}_{i,j,\mathcal{S}^*}) \leq \left(1 - \mathsf{Vol}(\pi_1^{-1}(\mathcal{S}^*) \cup \widehat{\mathcal{R}}'_j)\right)^{n-2} (2q) = O\left(\frac{q}{n^{2+\epsilon/6}}\right).$$

Similarly, let $\mathcal{F}_{i,j,\mathcal{S}^*}$ be the following event: $\mathcal{S}^*$ contains no points in $\mathcal{X} \setminus \{X_i, X_j\}$; X_j is at distance at most $2r$ from some point in $\mathcal{S}^*$; and moreover $\widehat{X}_i \in \widehat{\mathcal{P}}_j$. Notice that the probability that X_j is at distance at most $2r$ from some point in $\mathcal{S}^*$ is $O(r^2) = O(\log n/n)$. Hence,

$$\mathbf{P}(\mathcal{F}_{i,j,\mathcal{S}^*}) \leq (1 - \mathsf{Area}(\mathcal{S}^*))^{n-2}\, O\left(\frac{\log n}{n}\right)(2q) = O\left(\frac{q \log n}{n^{2+\epsilon/6}}\right).$$

Finally, observe that each of the events $(M_{3a} > 0 \wedge \widetilde{K}_{2,t+1} = 0)$ and $(M_{3a} > 0 \wedge B_t > 0)$ implies that either $\mathcal{E}_{i,j,\mathcal{S}^*}$ or $\mathcal{F}_{i,j,\mathcal{S}^*}$ hold, for some $i, j \in \{1, \ldots, n\}$ and some topologically connected union $\mathcal{S}^*$ of cells. Therefore, the probabilities of $(M_{3a} > 0 \wedge \widetilde{K}_{2,t+1} = 0)$ and $(M_{3a} > 0 \wedge B_t > 0)$ are at most

$$\sum_{i,j,\mathcal{S}^*} \mathcal{E}_{i,j,\mathcal{S}^*} + \sum_{i,j,\mathcal{S}^*} \mathcal{F}_{i,j,\mathcal{S}^*} = O\left(\frac{qn}{n^{\epsilon/6}}\right).$$

Part $5'$. Finally consider all the possible components in $G(\mathcal{X}; r)$ which are embeddable and have diameter at least $6\sqrt{2}r$. Call them components of *type* 3b, and let M_{3b} denote their number at time t.

We tessellate the torus into square cells of side αr, for some fixed but small enough $\alpha > 0$. Our goal is to show that if $G(\mathcal{X}_t; r)$ has some component of *type* 3b, then there exists some topologically connected union $\mathcal{S}^*$ of cells with $\mathsf{Area}(\mathcal{S}^*) \geq (11/5)\pi r^2$ and which does not contain any vertex in $\mathcal{X}$. Then, arguing as in Part $3'$, we conclude that both $\mathbf{P}(M_{3b} > 0 \wedge \widetilde{K}_{2,t+1} = 0)$ and $\mathbf{P}(M_{3b} > 0 \wedge B_t > 0)$ are $O\left(qn/(n^{1/5} \log n)\right)$. We now proceed to prove the claim on the union of cells $\mathcal{S}^*$. Given a component Γ of *type* 3b in $G(\mathcal{X}_t; r)$, let $\mathcal{S}'$, i_{T} and i_{B} be defined as in Part 3 (p.L. 3.2.5). Then, by repeating the same argument in there but replacing ϵr with $6\sqrt{2}r$, we can assume w.l.o.g. that the vertical distance between $X_{i_{\mathsf{T}}}$ and $X_{i_{\mathsf{B}}}$ is at least $6r$, and claim that the upper half-circle with centre $X_{i_{\mathsf{T}}}$ and radius r and the lower half-circle with centre $X_{i_{\mathsf{B}}}$ and radius r must be disjoint and contained in $\mathcal{S}'$. Now, consider the region of points in the torus $[0,1)^2$ with the y-coordinate between that of $X_{i_{\mathsf{T}}}$ and $X_{i_{\mathsf{B}}}$, and split this region into three horizontal bands of the same width. Observe that each band has width at least $2r$ and hence must contain some vertex of

Γ. For each of these bands, pick the rightmost vertex of Γ in the band. We select the right lower quarter-circle of radius r centred at the vertex if the vertex is closer to the top of the band, or otherwise the right upper quarter-circle. We also perform the symmetric operation and choose three more quarter-circles to the left of the leftmost vertices in the three bands. All these six quarter-circles together with the two half-circles previously described are by construction mutually disjoint and contained in $\mathcal{S}'$. Therefore $\mathsf{Area}(\mathcal{S}') \geq (5/2)\pi r^2$. Let $\mathcal{S}^*$ be the union of all the cells in the tessellation which are fully contained in $\mathcal{S}'$. We loose a bit of area compared to $\mathcal{S}'$. However, if α was chosen small enough, we can guarantee that $\mathcal{S}^*$ is topologically connected and also $\mathsf{Area}(\mathcal{S}^*) \geq (11/5)\pi r^2$. This α can be chosen to be the same for all components of *type* 3b. □

Now we can characterise the connectivity of $\big(G(\mathcal{X}_t; r)\big)_{t\in\mathbb{Z}}$ at two consecutive steps. We denote by $\mathcal{C}_t$ the event that $G(\mathcal{X}_t; r)$ is connected, and by $\mathcal{D}_t = \overline{\mathcal{C}_t}$ the event that $G(\mathcal{X}_t; r)$ is disconnected.

Corollary 3.3.7. *Assume that* $\mu = \Theta(1)$. *Then,*

$$\mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1}) \sim e^{-\mu}(1 - e^{-\mathbf{E}B}), \qquad \mathbf{P}(\mathcal{D}_t \wedge \mathcal{C}_{t+1}) \sim e^{-\mu}(1 - e^{-\mathbf{E}B})$$
$$\mathbf{P}(\mathcal{C}_t \wedge \mathcal{C}_{t+1}) \sim e^{-\mu}e^{-\mathbf{E}B}, \qquad \mathbf{P}(\mathcal{D}_t \wedge \mathcal{D}_{t+1}) \sim 1 - 2e^{-\mu} + e^{-\mu}e^{-\mathbf{E}B}$$

Proof. First observe that $K_{1,t} = S_t + D_t$ and $K_{1,t+1} = S_t + B_t$. Therefore we have

$$\mathbf{P}(K_{1,t} = 0 \wedge K_{1,t+1} > 0) = \mathbf{P}(S_t = 0 \wedge D_t = 0 \wedge B_t > 0),$$

and by Proposition 3.3.5 we get

$$\mathbf{P}(K_{1,t} = 0 \wedge K_{1,t+1} > 0) \sim e^{-\mathbf{E}S - \mathbf{E}D}(1 - e^{-\mathbf{E}B}) \sim e^{-\mu}(1 - e^{-\mathbf{E}B}). \tag{3.35}$$

We want to relate this probability with $\mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1})$. In fact, by partitioning $(K_{1,t} = 0 \wedge K_{1,t+1} > 0)$ and $(\mathcal{C}_t \wedge \mathcal{D}_{t+1})$ into disjoint events, we obtain

$$\mathbf{P}(K_{1,t} = 0 \wedge K_{1,t+1} > 0) = \mathbf{P}(\mathcal{C}_t \wedge K_{1,t+1} > 0) + \mathbf{P}(\mathcal{D}_t \wedge K_{1,t} = 0 \wedge K_{1,t+1} > 0),$$
$$\mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1}) = \mathbf{P}(\mathcal{C}_t \wedge K_{1,t+1} > 0) + \mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1} \wedge K_{1,t+1} = 0),$$

and thus we can write

$$\mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1}) = \mathbf{P}(K_{1,t} = 0 \wedge K_{1,t+1} > 0) + P_1 - P_2, \tag{3.36}$$

where $P_1 = \mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1} \wedge K_{1,t+1} = 0)$ and $P_2 = \mathbf{P}(\mathcal{D}_t \wedge K_{1,t} = 0 \wedge K_{1,t+1} > 0)$.

Suppose that $s = o\big(1/(rn)\big)$. In this case, $\mathbf{P}(K_{1,t} = 0 \wedge K_{1,t+1} > 0) = \Theta(srn)$ (see (3.35) and Proposition 3.3.5). Also observe that $\mathcal{D} \wedge (X = 0)$ implies that $\widetilde{X} > 0$. In fact, we must have at least two components of size greater than 1, so at least one of these must be non-solitary. Then, we have that $P_1 \leq \mathbf{P}(\widetilde{K}_{2,t} = 0 \wedge \widetilde{K}_{2,t+1} > 0)$ and $P_2 \leq \mathbf{P}(\widetilde{K}_{2,t} > 0 \wedge B_t > 0)$, and from Lemma 3.3.6 we get

$$P_1, P_2 = o\big(\mathbf{P}(K_{1,t} = 0 \wedge K_{1,t+1} > 0)\big). \tag{3.37}$$

Otherwise if $s = \Omega\big(1/(rn)\big)$, then $\mathbf{P}(K_{1,t} = 0 \wedge K_{1,t+1} > 0) = \Theta(1)$. In this case, we simply use the fact that $P_1 \leq \mathbf{P}(\widetilde{K}_{2,t+1} > 0) = o(1)$ and $P_2 \leq \mathbf{P}(\widetilde{K}_{2,t} > 0) = o(1)$ (see Theorem 3.2.7 and Lemma 3.3.1), and deduce that (3.37) also holds.

Finally, the asymptotic expression of $\mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1})$ is obtained from (3.35), (3.36) and (3.37). Moreover, by considering the time-reversed process, we deduce that $\mathbf{P}(\mathcal{D}_t \wedge \mathcal{C}_{t+1}) = \mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1})$. The remaining probabilities in the statement are computed from Corollary 3.2.3, Lemma 3.3.1, and using the fact that

$$\mathbf{P}(\mathcal{C}_t \wedge \mathcal{C}_{t+1}) = \mathbf{P}(\mathcal{C}_t) - \mathbf{P}(\mathcal{C}_t \wedge \mathcal{D}_{t+1}),$$
$$\mathbf{P}(\mathcal{D}_t \wedge \mathcal{D}_{t+1}) = \mathbf{P}(\mathcal{D}_t) - \mathbf{P}(\mathcal{D}_t \wedge \mathcal{C}_{t+1}). \qquad \square$$

Given any event $\mathcal{E}$ in the static model $G(\mathcal{X};r)$, we denote by $\mathcal{E}_t$ the event that $\mathcal{E}$ holds at time t. In the $\big(G(\mathcal{X}_t;r)\big)_{t\in\mathbb{Z}}$ model, we define $L_t(\mathcal{E})$ analogously to its definition in Chapter 2, which in fact is not specific of any particular random process. Namely, $L_t(\mathcal{E})$ is the number of consecutive steps that $\mathcal{E}$ holds starting at step t. Note that, as in Chapter 2, the distribution of $L_t(\mathcal{E})$ does not depend on t. Therefore, Lemma 2.5.10 also applies to the present setting.

Our main goal in this section is to compute the expected time that the graph of walkers remains (dis)connected, after a point in time at which it becomes (dis)connected. More precisely, define

$$L\mathcal{C}_{\text{av}} = \mathbf{E}(L_t(\mathcal{C}) \mid \mathcal{D}_{t-1} \wedge \mathcal{C}_t) \qquad \text{and} \qquad L\mathcal{D}_{\text{av}} = \mathbf{E}(L_t(\mathcal{D}) \mid \mathcal{C}_{t-1} \wedge \mathcal{D}_t).$$

Then, from Lemma 2.5.10, in order to find the asymptotic value of $L\mathcal{C}_{\text{av}}$ and $L\mathcal{D}_{\text{av}}$ it suffices to see that $\mathbf{E}(L(\mathcal{C})) < +\infty$ and $\mathbf{E}(L(\mathcal{D})) < +\infty$.

Lemma 3.3.8. *Let $b = b(n)$ be the smallest natural number such that $(b-3)ms \geq 3\sqrt{2}/2$. Then, there exists $p = p(n) > 0$ such that: for any fixed circle $\mathcal{R} \subset [0,1)^2$ of radius $r/2$, any $i \in \{1,\dots,n\}$, any $t \in \mathbb{Z}$, and conditional upon any particular position of $X_{i,t}$ in the torus, the probability that $X_{i,t+bm} \in \mathcal{R}$ is at least p.*

Proof. Fix arbitrary positions for $X_{i,t}$ and for the centre X of circle $\mathcal{R}$. Let t' be the smallest integer such that $t' \mid m$ and $t' \geq t$. Observe that t' is the first time after t when agent i selects a new angle. Let $h = t' - t$, which satisfies $0 \leq h < m$. The particular point $X_{i,t}$ is irrelevant in our argument, and we restrict our attention to the position of agent i at the times when it chooses a new angle, and also to the final position. For simplicity, we denote $Y_k = X_{i,t'+km}$ $(0 \leq k \leq b-1)$ and $Y_b = X_{i,t+bm}$. Observe that

$$d(Y_{k+1}, Y_k) = ms, \quad \forall k \,:\, 0 \leq k \leq b-2, \quad \text{and} \quad d(Y_b, Y_{b-1}) = (m-h)s. \tag{3.38}$$

Recall that, if α_k denotes the angle in which agent i moves between Y_k and Y_{k+1}, then each α_k is selected uniformly and independently at random from the interval $[0, 2\pi)$.

In order to prove the statement, we compute a lower bound on the probability of a strategy that is sufficient for agent i to reach $\mathcal{R}$ at time $t + bm$. We start from an arbitrary point $Y_0 \in [0,1)^2$ and build a sequence of points $Y_0, \dots, Y_b$ satisfying (3.38) such that $d(Y_b, X) \leq r/2$. The strategy consists in imposing some restrictions on the angles $\alpha_0, \dots, \alpha_b$. For the sake of simplicity in the geometrical descriptions, it is convenient to allow $Y_0, \dots, Y_b$ and X to lie in $\mathbb{R}^2$ rather than into the torus $[0,1)^2$. Once the construction of the sequence of points is completed, we map them back to the torus by the canonical projection. Hence, we assume hereinafter that Y_0 and X are two arbitrary points in $\mathbb{R}^2$ such that $d(Y_0, X) \leq \sqrt{2}/2$, which is the maximal distance in the torus $[0,1)^2$. For each k,

$0 \le k \le b-4$, we restrict α_k to be in $[\theta_k - \pi/6, \theta_k + \pi/6] \pmod{2\pi}$, where θ_k is the angle of $\overrightarrow{Y_k X}$ with respect to the horizontal axis. We claim that, with this choice of angle, the distance between Y_k and X is decreased at each step by at least $ms/3$ until it is at most ms. In fact by the law of cosines,

$$d(Y_{k+1}, X) \le \sqrt{\big(d(Y_k, X)\big)^2 + (ms)^2 - \sqrt{3}d(Y_k, X)ms}, \tag{3.39}$$

and therefore, if $d(Y_k, X) > ms$, we can write

$$\begin{aligned} d(Y_{k+1}, X) &\le \sqrt{\big(d(Y_k, X)\big)^2 + \Big(1 + \frac{2}{3} - \sqrt{3}\Big)(ms)^2 - \frac{2}{3}d(Y_k, X)ms} \\ &\le \sqrt{\big(d(Y_k, X)\big)^2 + \frac{1}{9}(ms)^2 - \frac{2}{3}d(Y_k, X)ms} \\ &= d(Y_k, X) - \frac{1}{3}ms. \end{aligned} \tag{3.40}$$

Otherwise, if $d(Y_k, X) \le ms$, then from (3.39) we deduce that also

$$d(Y_{k+1}, X) \le \sqrt{(1 - \sqrt{3})\big(d(Y_k, X)\big)^2 + (ms)^2} \le ms. \tag{3.41}$$

We claim that $d(Y_{b-3}, X) \le ms$. Suppose otherwise that $d(Y_{b-3}, X) > ms$. Then in view of (3.39), (3.40) and (3.41), for all k such that $0 \le k \le b-4$ we also have $d(Y_k, X) > ms$, and moreover

$$d(Y_{b-3}, X) \le d(Y_0, X) - (b-3)\frac{ms}{3} \le \frac{\sqrt{2}}{2} - (b-3)\frac{ms}{3} \le 0,$$

which contradicts the assumption, and proves the claim.

Now, let $Z \in \mathbb{R}^2$ be the only point on the line containing Y_{b-3} and X satisfying $d(Z, X) = (m-h)s$ and such that X lies on the segment $\overline{Y_{b-3}Z}$. Moreover, denote by W one of the two points on the perpendicular bisector of segment $\overline{Y_{b-3}Z}$ which satisfy $d(W, Y_{b-3}) = ms$. We want to set the angles α_{b-3}, α_{b-2} and α_{b-1} so that Y_{b-2}, Y_{b-1} and Y_b are close to W, Z and X respectively. Indeed, if ϕ_{b-3}, ϕ_{b-2} and ϕ_{b-1} are respectively the angles between the horizontal axis and $\overrightarrow{Y_{b-3}W}$, $\overrightarrow{WZ}$ and $\overrightarrow{ZX}$, then by imposing that $\alpha_k \in [\phi_k - \epsilon r/(ms), \phi_k + \epsilon r/(ms)] \pmod{2\pi}$ for some small enough $\epsilon > 0$, we achieve that $d(Y_b, X) \le r/2$ and thus $Y_b \in \mathcal{R}$.

As a conclusion, the probability of choosing all the angles according to the strategy described is

$$p := (1/6)^{b-3}\Theta\big((r/(ms))^3\big),$$

and this completes the proof. □

The next lemma allows us to apply Lemma 2.5.10.

Lemma 3.3.9. $\mathbf{E}(L(\mathcal{C})) < +\infty$ *and* $\mathbf{E}(L(\mathcal{D})) < +\infty$.

Proof. Fix one circle $\mathcal{R} \subset [0,1)^2$ of radius $r/2$, and take b as in the statement of Lemma 3.3.8. Since the agents choose their angles independently from each other, we have that, conditional upon any arbitrary $\mathcal{X}_t$, the probability that after bm steps all agents end up inside $\mathcal{R}$ is

$$\mathbf{P}(\mathcal{X}_{t+bm} \subset \mathcal{R} \mid \mathcal{X}_t) \ge p^n, \tag{3.42}$$

for some $p = p(n) > 0$. Observe that for any $t \in \mathbb{Z}$ the event $(\mathcal{X}_t \subset \mathcal{R})$ implies that $G(\mathcal{X}_t; r)$ is a clique, since all pairs of vertices in $\mathcal{X}_t$ are at distance at most r, and thus $G(\mathcal{X}_t; r)$ is connected. Consequently, for any $d \in \mathbb{N}$, we can write

$$\mathbf{P}\Big(\bigwedge_{k=0}^{d} \mathcal{D}_{t+kbm}\Big) \le (1-p^n)\mathbf{P}\Big(\bigwedge_{k=0}^{d-1} \mathcal{D}_{t+kbm}\Big) \le \mathbf{P}(\mathcal{D}_t)(1-p^n)^d. \tag{3.43}$$

Notice that the equation $L_t(\mathcal{D}) = \sum_{k=0}^{\infty} 1[\mathcal{D}_t] \cdots 1[\mathcal{D}_{t+k}]$, is satisfied pointwise, for every element in the probability space $(\mathcal{X}_t)_{t\in\mathbb{Z}}$. Therefore, by the Monotone Convergence Theorem, (3.43) and the fact that $p > 0$, we conclude

$$\begin{aligned}\mathbf{E}(L_t(\mathcal{D})) &= \sum_{k=0}^{\infty} \mathbf{P}(\mathcal{D}_t \wedge \cdots \wedge \mathcal{D}_{t+k}) \\ &\le \sum_{d=0}^{\infty} bm\, \mathbf{P}\Big(\bigwedge_{k=0}^{d} \mathcal{D}_{t+kbm}\Big) \\ &\le bm\, \mathbf{P}(\mathcal{D}_t) \sum_{d=0}^{\infty} (1-p^n)^d < +\infty.\end{aligned}$$

The same kind of argument shows that $\mathbf{E}(L(\mathcal{C})) < +\infty$. In this case we fix two circles $\mathcal{R}$ and $\mathcal{R}'$ in $[0,1)^2$ of radius $r/2$ with centres at distance greater than $2r$. Observe that for any $t \in \mathbb{Z}$ the event $\big((\mathcal{X}_t \setminus \{X_{1,t}\}) \subset \mathcal{R}\big) \wedge (X_{1,t} \in \mathcal{R}')$ implies that $G(\mathcal{X}_t; r)$ is disconnected. Moreover, from Lemma 3.3.8, we obtain an analogue to (3.42)

$$\mathbf{P}\Big((\mathcal{X}_{t+bm} \setminus \{X_{1,t+bm}\}) \subset \mathcal{R}\big) \wedge (X_{1,t+bm} \in \mathcal{R}') \mid \mathcal{X}_t\Big) \ge p^n, \tag{3.44}$$

and the argument follows as in the previous case but replacing $\mathcal{D}$ with $\mathcal{C}$. □

We are now ready to prove our main theorem which characterises the expected number of steps the graph remains (dis)connected once it becomes (dis)connected.

Theorem 3.3.10.

$$L\mathcal{C}_{\mathrm{av}} \sim \frac{1}{(1-e^{-\mathbf{E}B})} = \begin{cases} \frac{\pi}{4srn} & \textit{if } srn = o(1), \\ \frac{1}{(1-e^{-4srn/\pi})} & \textit{if } srn = \Theta(1), \\ 1 & \textit{if } srn = \omega(1), \end{cases} \qquad \textit{and}$$

$$L\mathcal{D}_{\mathrm{av}} \sim \frac{e^{\mu}-1}{(1-e^{-\mathbf{E}B})} = \begin{cases} \frac{\pi(e^{\mu}-1)}{4srn} & \textit{if } srn = o(1), \\ \frac{e^{\mu}-1}{(1-e^{-4srn/\pi})} & \textit{if } srn = \Theta(1), \\ e^{\mu}-1 & \textit{if } srn = \omega(1). \end{cases}$$

Proof. By Lemma 3.3.9, we have that $\mathbf{E}(L_k(\mathcal{C})) < +\infty$, $\mathbf{E}(L_k(\mathcal{D})) < +\infty$. Then we can apply Lemma 2.5.10, and the result follows by Corollaries 3.2.3 and 3.3.7. □

3.4 Conclusions.

In this chapter, we have introduced the dynamic random geometric graph and studied the expected length of the connectivity and disconnectivity periods, considering different step sizes s and different lengths m during which the angle remains invariant, always considering the static connectivity threshold $r = r_c$. We believe that a similar analysis can be performed for other values of r as well.

The *random walk mobility* model simulates the behaviour of a swarm of mobile agents as sensors or robots, which move randomly to monitor an unknown territory or to search in it. There exist other models such as the *way-point model*, where each agent chooses randomly a fixed way-point (from a set of pre-determined way-points) and moves there, and when it arrives it chooses another and moves there, and so on [18]. A possible line of future research is to use the techniques and results developed in this chapter to study other models of mobility. Another line of work is to explore the behaviour of other graph properties, such as the chromatic number or the clique number, as the dynamic random geometric graphs evolve.

4

Sharp Threshold for Hamiltonicity of Random Geometric Graphs

4.1 Introduction

Given a graph G on n vertices, a *Hamiltonian cycle* is a simple cycle that visits each vertex of G exactly once. A graph is said to be *Hamiltonian* if it contains a Hamiltonian cycle. The problem of given a graph, deciding if it is Hamiltonian or not is known to be NP-complete [33]. Two known facts for the Hamiltonicity of random graphs are that almost all d-regular graphs ($d \geq 3$) are Hamiltonian [71], and that in the $\mathcal{G}_{n,p}$ model if $p(n) = (\log n + \log\log n + \omega(n))/n$, then a.a.s. $\mathcal{G}_{n,p}$ is Hamiltonian [52] (see also chapter 8 of [15]).

Random geometric graphs are the randomised version of unit disk graphs. An undirected graph is a *unit disk graph* if its vertices can be put into one-to-one correspondence with circles of equal radius in the plane in such a way that two vertices are joined by an edge iff their corresponding circles intersect [20]. The problem of deciding if a given unit disk graph is Hamiltonian is known to be NP-complete [44].

A natural issue to study is the existence of Hamiltonian cycles in a random geometric graph $G(\mathcal{X}; r) = G\big(\mathcal{X}(n); r(n)\big)$. Penrose in his book [66] posed it as an open problem whether exactly at the point where $G(\mathcal{X}; r)$ gets 2-connected, the graph also becomes Hamiltonian a.a.s. Petit in [68] proved that for $r = \omega(\sqrt{\log n/n})$, $G(\mathcal{X}; r)$ is Hamiltonian a.a.s. and he also gave a distributed algorithm to find a Hamiltonian cycle in $G(\mathcal{X}; r)$ with his choice of radius. In the present chapter, we find the sharp threshold of this property for a random geometric graph over the unit square under any ℓ_p-normed distance. In fact, let p ($1 \leq p \leq \infty$) be arbitrary but fixed throughout the chapter, and let $\mathcal{G}$ denote a random geometric graph $G(\mathcal{X}; r)$ over $[0, 1]^2$ with respect to ℓ_p. Let α_p be the area of the unit disk in the ℓ_p norm, and recall from Section 2.2 that the connectedness of $\mathcal{G}$ has a sharp threshold at $r = r(n) = \sqrt{\log n/(\alpha_p n)}$. We first show the following

Theorem 4.1.1. *The property that a random geometric graph $\mathcal{G} = G(\mathcal{X}; r)$ contains a Hamiltonian cycle exhibits a sharp threshold at $r = \sqrt{\frac{\log n}{\alpha_p n}}$, where α_p is the area of the unit disk in the ℓ_p norm. More precisely, for any $\epsilon > 0$,*

- *if $r = \sqrt{\frac{\log n}{(\alpha_p+\epsilon)n}}$, then a.a.s. $\mathcal{G}$ contains no Hamiltonian cycle,*
- *if $r = \sqrt{\frac{\log n}{(\alpha_p-\epsilon)n}}$, then a.a.s. $\mathcal{G}$ contains a Hamiltonian cycle.*

And as a corollary of the proof, we describe a linear time algorithm that finds a Hamiltonian cycle in $G(\mathcal{X}; r)$ a.a.s., provided that $r \geq \sqrt{\frac{\log n}{(\alpha_p-\epsilon)n}}$ for some fixed $\epsilon > 0$.

4.2 Proof of Theorem 4.1.1

To prove Theorem 4.1.1, note that the lower bound of the threshold is trivial. In fact, if $r = \sqrt{\frac{\log n}{(\alpha_p+\epsilon)n}}$, then a.a.s. $\mathcal{G}$ is disconnected [65] and hence it cannot contain any Hamiltonian cycle. To simplify the proof of the upper bound, we need some auxiliary definitions and lemmas. In the remainder of the section, we assume that $r = \sqrt{\frac{\log n}{(\alpha_p-\epsilon)n}}$ for some fixed $\epsilon > 0$, and we show that a.a.s. $\mathcal{G}$ contains a Hamiltonian cycle.

Let us take $y = \lceil \frac{2}{r} \rceil^{-1}$. Intuitively, y is close to $r/2$ but slightly smaller. We divide $[0,1]^2$ into squares of side length y. Call this the *initial tessellation* of $[0,1]^2$. Two different squares R and S are defined to be *friends* if they are either adjacent (i.e., they share at least one corner) or there exists at least one other square T adjacent to both R and S. Thus, each square has at most 24 friends. Then, we create a second and finer tessellation of $[0,1]^2$ by dividing each square into k^2 new squares of side length $y/k \sim r/(2k)$, for some large enough but fixed $k = k(\epsilon) \in \mathbb{N}$. We call this the *fine tessellation* of $[0,1]^2$, and we refer to these smaller squares as *cells*. We note that the total numbers of squares and cells are both $\Theta(1/r^2)$. Note that with probability 1, for every fixed n, any vertex will be contained in exactly one cell (and in exactly one square). In the following we always assume this.

We say that a cell is *dense*, if it contains at least 48 vertices of $\mathcal{G}$. If the cell contains at least one vertex but less than 48 vertices, we say the cell is *sparse*. If the cell contains no vertex, the cell is *empty*. Furthermore we define an *animal* to be a union of cells which is topologically connected. The *size* of an animal is the number of different cells it contains. In particular, the squares of the initial tessellation of $[0,1]^2$ are animals of size k^2. An animal is called *dense* if it contains at least one dense cell. If an animal contains no dense cell, but it contains at least one vertex of $\mathcal{G}$, it is called *sparse*.

From hereinafter, all distances in $[0,1]^2$ will be taken in the ℓ_p metric. As usual, the distance between two sets of points P_1 and P_2 in $[0,1]^2$ is the infimum of the distances between any pair of points in P_1 and P_2. Two cells c_1 and c_2 are said to be *close* to each other if

$$\sup_{p_1 \in c_1, p_2 \in c_2} \{\text{distance}(p_1, p_2)\} \leq r.$$

For an arbitrary cell c at distance at least r from the boundary of $[0,1]^2$, let $K = K(n)$ be the number of cells which are close to c and also above and to the right of c. Obviously, K does not depend on the particular cell we chose.

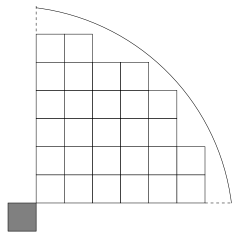


Figure 4.1: Set of cells close, above and to the right of the shaded cell

Lemma 4.2.1. *For any $\eta > 0$, we can choose k sufficiently large such that $K > (\alpha_p - \eta)k^2$ for n large enough.*

Proof. Let c be a cell at distance at least r from the boundary of $[0,1]^2$. Call A the union of the cells which are close to c and also above and to the right of c. Let p be the top right corner point of c. Define the set

$$B = \{q \in [0,1]^2 \cap R : \text{distance}(p,q) \leq r - 4y/k\},$$

where R is the set of points which are above and to the right of p. Observe that $B \subseteq A$. Moreover, if k is chosen large enough, the area of B is at least $\frac{1}{4}(\alpha_p - \eta)r^2$. Thus, A contains at least $\frac{1}{4}(\alpha_p - \eta)r^2/(y/k)^2 > (\alpha_p - \eta)k^2$ cells. □

Lemma 4.2.2. *The following statements are true a.a.s.*

(i). All animals of size $4K$ are dense.

(ii). All animals of size $2K$ which touch any of the four sides of $[0,1]^2$ are dense.

(iii). All cells at distance less than $4y$ from two sides of $[0,1]^2$ are dense.

Proof. Let $0 < \delta < \epsilon$. Taking into account that the side length of each cell is $y/k \geq \frac{1}{2k}\sqrt{\frac{\log n}{(\alpha_p - \delta)n}}$ (but also $y/k \leq c\sqrt{\log n/n}$ for some $c > 0$), the probability that any given cell is not dense (i.e., it contains at most 47 vertices) is

$$\sum_{i=0}^{47} \binom{n}{i} \left(\frac{y^2}{k^2}\right)^i \left(1 - \frac{y^2}{k^2}\right)^{n-i} = \Theta(1) n^{47} \left(\frac{y^2}{k^2}\right)^{47} \left(1 - \frac{y^2}{k^2}\right)^n,$$

since the weight of this sum is concentrated in the last term. Then, plugging in the bounds for y/k, we get that the probability above is

$$O(1)(ny^2/k^2)^{47} e^{-y^2 n/k^2} = O(1)(\log n)^{47} n^{-\frac{1}{4k^2(\alpha_p - \delta)}}.$$

For each one of the cells of a given animal, we can consider the event that this particular cell is not dense. Notice that these events are negatively correlated, as the probability that any particular cell is not dense conditional upon having some other cells with at most 47 vertices is not greater than the unconditional probability. Thus, the probability that a given animal of size $4K$ contains no dense cell is at most

$$\left(O(1)(\log n)^{47} n^{-\frac{1}{4k^2(\alpha_p-\delta)}}\right)^{4K} = O(1)(\log n)^{C} n^{-\frac{K}{k^2(\alpha_p-\delta)}},$$

for some constant C. Let $\rho = \frac{K}{k^2(\alpha_p-\delta)}$. From Lemma 4.2.1 applied with any $0 < \eta < \delta$, by choosing k sufficiently large, we can guarantee that $\rho > 1$. Now note that the number of animals of size $4K$ is $O(1/r^2)$ since for each fixed shape of an animal there are $O(1/r^2)$ many choices and there is only a constant number of shapes. Thus, by taking a union bound over all animals and plugging in the value of r, we get that the probability of having an animal without any dense cell is

$$O(1)(\log n)^{C-1}/n^{\rho-1} = o(1),$$

and (i) holds.

An analogous argument shows that any given animal of size $2K$ is not dense with probability

$$O(1)(\log n)^{C/2} n^{-\rho/2}.$$

Observe that there exist only $O(1/r)$ animals touching any of the four sides of $[0,1]^2$. Hence, the probability that one of these is not dense is

$$O(1)(\log n)^{(C-1)/2}/n^{(\rho-1)/2} = o(1),$$

and (ii) is proved.

To prove (iii), we simply recall that the probability that a given cell is not dense is $o(1)$. By taking a union bound, the same argument holds for a constant number of cells. □

Lemma 4.2.3. *A.a.s., for any cell c_1, there exists a cell c_2 which is dense and close to c_1.*

Proof. Let S be the square of the initial tessellation of $[0,1]^2$ where c_1 is contained, and let A be the animal containing all the cells which are close to c_1 but different from c_1. Suppose that S is at distance at least $2y$ from all sides of $[0,1]^2$. Then, A has size greater than $4K$, and it must contain some dense cell by Lemma 4.2.2 (i) a.a.s.

Otherwise, suppose that S is at distance less than $2y$ from just one side of $[0,1]^2$. Then, A has size greater than $2K$ and it touches one side of $[0,1]^2$, and thus it must contain some dense cell by Lemma 4.2.2 (ii) a.a.s.

Finally, if S is at distance less than $2y$ from two sides of $[0,1]^2$, then all cells in that square must be dense by Lemma 4.2.2 (iii) a.a.s. □

We now consider the following auxiliary graph $\mathcal{G}'$: the vertices of $\mathcal{G}'$ are all those squares belonging to the initial tessellation of $[0,1]^2$ which are dense, and there is an edge between two dense squares R and S if they are friends and there exist cells $c_1 \subset R$ and $c_2 \subset S$ which are dense and close to each other. We observe that the maximal degree of $\mathcal{G}'$ is 24.

Lemma 4.2.4. *A.a.s., $\mathcal{G}'$ is connected.*

Proof. Suppose for contradiction that $\mathcal{G}'$ contains at least two connected components Γ_1 and Γ_2. We denote by D the union of all dense cells which are contained in some vertex (i.e., dense square) of Γ_1, and let $H \supseteq D$ be the union of all cells which are close to some cell contained in D. Note that H is topologically connected, and let the closed curve γ be the outer boundary of H with respect to $\mathbb{R}^2$. Each connected part obtained by removing from γ the intersection with the sides of $[0,1]^2$ is called a *piece* of γ. Define by E the union of all cells in H but not in D. In general, E might have several connected components (animals). Moreover, all cells in E must be not dense, by construction. Note that any cell in D cannot touch any piece of γ. Hence, each piece of γ is touched by exactly one connected component $A \subseteq E$. Observe that, if γ touches some side of $[0,1]^2$, then all connected components of E touching some piece of γ must also touch some side of $[0,1]^2$.

Given any of the four sides s of $[0,1]^2$, the distance between s and Γ_1 is understood to be the distance between s and the dense square of Γ_1 which has the smallest distance to s. We now distinguish between a few cases depending on the fact whether Γ_1 is at distance less than $2y$ from one (or more) side(s) of $[0,1]^2$ or not.

Case 1. Γ_1 is at distance at least $2y$ from any side of $[0,1]^2$.
In this case, let A be the only connected component of E which touches γ. Consider the uppermost dense cell $c \subset D$ (if there are several ones, choose an arbitrary one) and the lowermost dense cell $d \subset D$ (possibly equal to c). Then all cells which are close to c and above c and all cells which are close to d and below d belong to A. Since there are at least as many as $4K$ of these, we have an animal A of size at least $4K$ without any dense cell, which by Lemma 4.2.2 (i) does not happen a.a.s.

Case 2. Γ_1 is at distance less than $2y$ from exactly one side of $[0,1]^2$.
W.l.o.g. we can assume that Γ_1 is at distance less than $2y$ from the bottom side of $[0,1]^2$. Consider the uppermost dense cell $c \subset D$ (if there are several ones, choose an arbitrary one). Let A be the connected component of E which contains all cells which are close to c and above c. Note that there are at least as many as $2K$ of these cells. Moreover, A touches one of the pieces of γ. Hence, we have an animal A of size at least $2K$ without any dense cell and that touches some side of $[0,1]^2$. By Lemma 4.2.2 (ii) this does not happen a.a.s.

Case 3. Γ_1 is at distance less than $2y$ from two opposite sides of $[0,1]^2$.
W.l.o.g. we can assume that Γ_1 is at distance less than $2y$ from the top and the bottom sides of $[0,1]^2$. Among all cells contained in squares of Γ_1 that are at distance less than $4y$ from the top side of $[0,1]^2$, consider the rightmost dense cell c. If c is at distance less than $2y$ from that side, consider all K cells which are close to c and below and to the right of c. Otherwise, if c is at distance at least $2y$ from that side, consider all K cells which are close to c and above and to the right of c. Let A be the connected component of E containing these cells. Similarly, among all cells contained in squares of Γ_1 that are at distance less than $4y$ from the bottom side of $[0,1]^2$, consider the rightmost dense cell d. Again, if d is at distance less than $2y$ from that side, consider all K cells which are close to d and above and to the right of d. Otherwise, if d is at distance at least $2y$ from that side, consider all K cells which are close to d and below and to the right of d. Thus, in either case, we obtain K cells pairwise different from the K previously described ones, and let A' be the connected component containing them. A and A' must be the same, since they touch the same piece of γ. Hence, we have an animal A of size at least $2K$ touching at least one side of $[0,1]^2$ and without any dense cell. By Lemma 4.2.2 (ii) this does not happen a.a.s.

Case 4. Γ_1 is at distance less than $2y$ from one vertical and one horizontal sides of $[0,1]^2$. W.l.o.g. we can assume that Γ_1 is at distance less than $2y$ from the left and the top sides of $[0,1]^2$. Among all cells contained in squares of Γ_1 that are at distance less than $4y$ from the top side of $[0,1]^2$, consider the rightmost dense cell c. If c is at distance less than $2y$ from that side, consider all K cells which are close to c and below and to the right of c. Otherwise, if c is at distance at least $2y$ from that side, consider all K cells which are close to c and above and to the right of c. Let A be the connected component of E containing all these K cells. By construction, all these K cells are at distance less than $4y$ from the top side of $[0,1]^2$. Then, by Lemma 4.2.2 (iii), they must be a.a.s. at distance at least $4y$ from the left side of $[0,1]^2$, since otherwise they would be all dense. Similarly, among all cells contained in squares of Γ_1 that are at distance less than $4y$ from the left side of $[0,1]^2$, consider the lowermost dense cell d. Again, if d is at distance less than $2y$ from that side, consider all K cells which are close to d and below and to the right of d. Otherwise, if d is at distance at least $2y$ from that side, consider all K cells which are close to d and below and to the left of d. Let A' be the connected component of E containing these K cells. By construction, all these K cells are at distance less than $4y$ from the left side of $[0,1]^2$, and hence they must be pairwise different from the K ones previously described a.a.s. (note that we used Lemma 4.2.2 (iii) to prove that the K cells contained in A described above must be at distance at least $4y$ from the top side of $[0,1]^2$). Moreover, A and A' must be the same, since they touch the same piece of γ. Then we have an animal A of size at least $2K$ touching at least one side of $[0,1]^2$ without any dense cell. By Lemma 4.2.2 (ii) this does not happen a.a.s.

Case 5. Γ_1 is at distance less than $2y$ from three sides of $[0,1]^2$.
W.l.o.g. we can assume that Γ_1 is at distance less than $2y$ from the left, top and bottom sides of $[0,1]^2$. The argument is exactly the same as in Case 3, and hence this case does not occur a.a.s.

In case Γ_2 is at distance at least $2y$ from some side of $[0,1]^2$, we can apply one of the above cases with Γ_2 instead of Γ_1. Thus, it suffices to consider the following:

Case 6. Both Γ_1 and Γ_2 are at distance less than $2y$ from all four sides of $[0,1]^2$.
Let Q be the union of all those cells at distance less than $4y$ from both the bottom and left sides of $[0,1]^2$. By Lemma 4.2.2, all the cells in Q must be dense, and thus must belong to squares of the same connected component of $\mathcal{G}'$. W.l.o.g., we can assume that they are not in D (i.e. are not contained in squares of Γ_1). Among all cells contained in squares of Γ_1 that are at distance less than $4y$ from the bottom side of $[0,1]^2$, consider the leftmost dense cell c. If c is at distance less than $2y$ from that side, consider all K cells which are close to c and above and to the left of c. Otherwise, if c is at distance at least $2y$ from that side, consider all K cells which are close to c and below and to the left of c. Let A be the connected component of E containing all these K cells. By construction, all these K cells are at distance less than $4y$ from the bottom side of $[0,1]^2$. Then, by Lemma 4.2.2 (iii), they must be a.a.s. at distance at least $4y$ from the left side of $[0,1]^2$, since otherwise they would be all dense. Similarly, among all cells contained in squares of Γ_1 that are at distance less than $4y$ from the left side of $[0,1]^2$, consider the lowermost dense cell d. Again, if d is at distance less than $2y$ from that side, consider all K cells which are close to d and below and to the right of d. Otherwise, if d is at distance at least $2y$ from that side, consider all K cells which are close to d and below and to the left of d. Let A' be the connected component

of E containing all these K cells. By construction, all these K cells are at distance less than $4y$ from the left side of $[0,1]^2$, and hence they must be pairwise different from the K ones previously described a.a.s. Moreover, A and A' must be the same, since they touch the same piece of γ. Then we have an animal A of size at least $2K$ touching at least one side of $[0,1]^2$ without any dense cell. By Lemma 4.2.2 (ii) this does not happen a.a.s. □

Proof of the upper bound of Theorem 4.1.1. Starting from $\mathcal{G}'$ we construct a new graph $\mathcal{G}''$, by adding some new vertices and edges as follows. Let us consider one fixed sparse square S of the initial tessellation of $[0,1]^2$. For each sparse cell c contained in S, we can a.a.s. find at least one dense cell close to it (by Lemma 4.2.3) which we call the *hook cell* of c (if this cell is not unique, or even the square containing these cell(s) is not unique, take an arbitrary one). This hook cell must lie inside some dense square R, which is a friend of S. Then, that sparse cell c gets the label R. By grouping those ones sharing the same label, we partition the sparse cells of S into at most 24 groups. Each of these groups of sparse cells will be thought as a new vertex, added to graph $\mathcal{G}'$ and connected by an edge to the vertex of $\mathcal{G}'$ described by the common label. By doing this same procedure for all the remaining sparse squares, we obtain the desired graph $\mathcal{G}''$. Those vertices in $\mathcal{G}''$ which already existed in $\mathcal{G}'$ (i.e., dense squares) are called *old*, and those newly added ones are called *new*. Notice that by construction of $\mathcal{G}''$ and by Lemma 4.2.4, $\mathcal{G}''$ must be connected a.a.s.

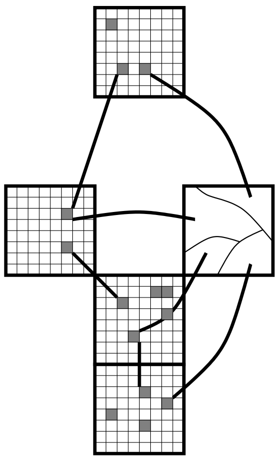


Figure 4.2: Illustration of $\mathcal{G}''$

Now, consider an arbitrary spanning tree $\mathcal{T}$ of $\mathcal{G}''$. Observe that the maximal degree of $\mathcal{T}$ is 24, and that all new vertices of $\mathcal{T}$ have degree one and are connected to old vertices. We use capital letters U, V to denote vertices of $\mathcal{T}$ and reserve the lowercase u, v, w for vertices of $\mathcal{G}$. Fix an arbitrary traversal of $\mathcal{T}$ which, starting at an arbitrary vertex, traverses each edge of $\mathcal{T}$ exactly twice and returns to the starting vertex. Note that such a traversal always exists: fix an arbitrary vertex of $\mathcal{T}$ to be the root vertex and always follow the edge going to the leftmost neighbour of that vertex (the vertex with smallest x-coordinate, and if there are more, the one among them with the smallest y-coordinate) which was not yet visited. Do this recursively for each vertex. When all neighbours of a vertex are visited, we go back to the vertex we came from. We iterate this procedure until all vertices are visited and we are back at the root vertex. This traversal gives an ordering in which we construct our Hamiltonian cycle in $\mathcal{G}$ (i.e., as the Hamiltonian cycle travels along the vertices of $\mathcal{G}$, it will visit the vertices of $\mathcal{T}$ according to this traversal).

Let us give a constructive description of our Hamiltonian cycle. Suppose that at some time we visit an old vertex U of $\mathcal{T}$ and that the next vertex V (w.r.t. the traversal) is also old. Then, there must exist a pair of dense cells $c_1 \subset U$, $c_2 \subset V$ close to each other, and let $u \in c_1$ and $v \in c_2$ be vertices not used so far. In case this is not the last time we visit U (w.r.t. the traversal), immediately after entering vertex w inside U we connect w to u and then u is connected to v. If U is visited for the last time (w.r.t. the traversal), we connect from the entering vertex w all vertices inside U not yet used by an arbitrary Hamiltonian path (note that they form a clique in $\mathcal{G}$) before leaving U via u, and subsequently we connect u to v.

Otherwise, suppose that at some time we visit an old vertex U of $\mathcal{T}$ and that the next vertex V (w.r.t. the traversal) is new. We connect all the vertices inside V (possibly just one) by an arbitrary Hamiltonian path, whose endpoints lie inside the sparse cells $d_1 \subset V$ and $d_2 \subset V$ (possibly equal). Again this is possible since these vertices form a clique in $\mathcal{G}$. Let $c_1 \subset U$ and $c_2 \subset U$ (possibly equal) be the hook cells of d_1 and d_2 (i.e., c_i is a dense cell in U close to the sparse cell d_i in V). Let $u \in c_1$ and $v \in c_2$ be vertices not used so far. Then, immediately after entering vertex w inside U we connect w to u and then u is joined to the corresponding endpoint of the Hamiltonian path connecting the vertices inside V. The other endpoint is connected to v, and so we visit again U.

We observe that at some steps of the above construction we request for unused vertices of $\mathcal{G}$. This is always possible: in fact, each vertex of $\mathcal{T}$ is visited as many times as its degree (at most 24); for each visit of an old vertex U our construction requires exactly two unused vertices $v \in c$, $w \in c$ inside some dense cell $c \subset U$; and c contains at least 48 vertices. By construction, the described cycle is Hamiltonian and the result holds. □

In the following corollary, we give an informal definition of a linear time algorithm that constructs a Hamiltonian cycle for a specific instance of $G(\mathcal{X}; r)$. The procedure is based on the previous constructive proof. We assume that real arithmetic can be done in constant time.

Corollary 4.2.5. *Let $r \geq \sqrt{\frac{\log n}{(\alpha_p - \epsilon)n}}$, for some fixed $\epsilon > 0$. The proof of Theorem 4.1.1 yields an algorithm that a.a.s. produces a Hamiltonian cycle in $G\big(\mathcal{X}(n); r(n)\big)$ in linear time with respect to n.*

Proof. Assume that the input graph satisfies all the conditions required in the proof of Theorem 4.1.1, which happens a.a.s. Assume also that each vertex of the input graph is represented by a pair of coordinates. Observe that the total number of squares is $O(n/\log n)$, and since the number of cells per square is constant, the same holds for the total number of cells. First we compute in linear time the label of the cell and the square where each vertex is contained. At the same time, we can find for each cell (and square) the set of vertices it contains, and mark those cells (squares) which are dense. Now, for the construction of $\mathcal{G}'$, note that each dense square has at most a constant number of friends to which it can be connected. Thus, the edges of $\mathcal{G}'$ can be obtained in time $O(n/\log n)$. In order to construct $\mathcal{G}''$, for each of the $O(n/\log n)$ cells in sparse squares, we compute in constant time its hook cell and the dense square containing it. Since both the number of vertices and the number of edges of $\mathcal{G}''$ are $O(n/\log n)$, we can compute in time $O(n)$ (e.g., by Kruskal's algorithm) an arbitrary spanning tree $\mathcal{T}$ of $\mathcal{G}''$. The traversal and construction of the Hamiltonian cycle is proportional to the number of edges in $\mathcal{T}$ plus the number of vertices in $\mathcal{G}$ and thus can be done in linear time. □

4.3 Conclusion and Outlook

We believe that the above construction can be generalised to obtain sharp thresholds for Hamiltonicity for random geometric graphs in $[0,1]^d$ (d being fixed). However, it seems much more difficult to generalise the results to arbitrary distributions of the vertices. The problem posed by Penrose [66], whether exactly at the point where $G(\mathcal{X};r)$ gets 2-connected the graph also becomes Hamiltonian a.a.s. or not, still remains open.

Bibliography

[1] Milton Abramowitz and Irene A. Stegun, editors. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* U.S. Government Printing Office, tenth edition, 1972.

[2] Dimitris Achlioptas and Ehud Friedgut. A sharp threshold for k-colorability. *Random Structures and Algorithms*, 14(1):63–70, 1999.

[3] Dimitris Achlioptas and Cristopher Moore. Almost all graphs with average degree 4 are 3-colorable. *Journal of Computer and System Sciences*, 67(2):441–471, 2003.

[4] Dimitris Achlioptas and Cristopher Moore. The chromatic number of random regular graphs. In *7th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems* (APPROX'04) *and 8th International Workshop on Randomization and Computation* (RANDOM'04), volume 3122 of *Lecture Notes in Computer Science*, pages 219–228. Springer, 2004.

[5] Dimitris Achlioptas and Assaf Naor. The two possible values of the chromatic number of a random graph. *Annals of Mathematics*, 162(3):1333–1349, 2005.

[6] Pankaj K. Agarwal, Leonidas J. Guibas, Herbert Edelsbrunner, Jeff Erickson, Michael Isard, Sariel Har-Peled, John Hershberger, Christian Jensen, Lydia Kavraki, Patrice Koehl, Ming Lin, Dinesh Manocha, Dimitris Metaxas, Brian Mirtich, David Mount, S. Muthukrishnan, Dinesh Pai, Elisha Sacks, Jack Snoeyink, Subhash Suri, and Ouri Wolefson. Algorithmic issues in modeling motion. *ACM Computing Surveys*, 34(4):550–572, 2002.

[7] Ian F. Akyildiz, Weilian Su, Yogesh Sankarasubramaniam, and Erdal Cayirci. Wireless sensor networks: a survey. *Computer Networks*, 38(4):393–422, 2002.

[8] Martin J.B. Appel and Ralph P. Russo. The connectivity of a graph on uniform points on $[0,1]^d$. *Statistics & Probability Letters*, 60(4):351–357, 2002.

[9] Hilda Assiyatun and Nicholas C. Wormald. Independent sets and 2-independent sets in random d-regular graphs. (in preparation).

[10] Chen Avin and Gunes Ercal. On the cover time of random geometric graphs. In *32nd International Colloquium on Automata, Languages and Programming* (ICALP'05), volume 3580 of *Lecture Notes in Computer Science*, pages 677–689. Springer, 2005.

[11] Andrew D. Barbour, Lars Holst, and Svante Janson. *Poisson Approximation*, volume 2 of *Oxford Studies in Probability.* Oxford University Press, 1992.

[12] Jillian Beardwood, John H. Halton, and John M. Hammersley. The shortest path through many points. *Proceedings of the Cambridge Philosophical Society*, 55:299–327, 1959.

[13] Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1:311–316, 1980.

[14] Béla Bollobás. The chromatic number of random graphs. *Combinatorica*, 8(1):49–55, 1988.

[15] Béla Bollobás. *Random Graphs*, volume 73 of *Cambridge studies in advanced mathematics.* Cambridge University Press, second edition, 2001.

[16] Alfredo Braunstein, Roberto Mulet, Andrea Pagnani, Martin Weigt, and Riccardo Zecchina. Polynomial iterative algorithms for coloring and analyzing random graphs. *Phys. Rev. E*, 68(3), 2003.

[17] Daniel Brélaz. New methods to color the vertices of a graph. *Communications of the ACM*, 22(4):251–256, 1979.

[18] Tracy Camp, Jeff Boleng, and Vanessa Davies. A survey of mobility models for ad hoc network research. *Wireless Communications and Mobile Computing: Special issue on Mobile Ad Hoc Networking: Research, Trends and Applications*, 2(5):483–502, 2002.

[19] Peter Cheeseman, Bob Kanefsky, and William M. Taylor. Where the really hard problems are. In *12th International Joint Conference on Artificial Intelligence* (IJCAI'91), pages 331–340, 1991.

[20] Brent N. Clark, Charles J. Colbourn, and David S. Johnson. Unit disk graphs. *Discrete Mathematics*, 86(1–3):165–177, 1990.

[21] Joseph C. Culberson and Ian P. Gent. Frozen development in graph coloring. *Theor. Comput. Sci.*, 265(1–2):227–264, 2001.

[22] Josep Díaz, Giannis Grammatikopoulos, Alexios Kaporis, Lefteris Kirousis, Xavier Pérez, and Dimitris G. Sotiropoulos. 5-regular graphs are 3-colorable with positive probability. In Gerth Stølting Brodal and Stefano Leonardi, editors, *13th Annual European Symposium on Algorithms* (ESA'05), volume 3669 of *Lecture Notes in Computer Science*, pages 215–225. Springer, 2005.

[23] Josep Díaz, Alexios Kaporis, Graeme Kemkes, Lefteris Kirousis, Xavier Pérez, and Nicholas C. Wormald. On the chromatic number of a random 5-regular graph. (sent for publication, 2006).

[24] Josep Díaz, Alexios Kaporis, Lefteris Kirousis, and Xavier Pérez. Partitioning networks into classes of mutually isolated nodes. In *European Conference on Complex Systems*, 2005.

[25] Josep Díaz, Dieter Mitsche, and Xavier Pérez. Sharp threshold for hamiltonicity of random geometric graphs. *SIAM Journal on Discrete Mathematics*, 21(1):57–65, 2007.

[26] Josep Díaz, Xavier Pérez, Maria J. Serna, and Nicholas C. Wormald. Connectivity for wireless agents moving on a cycle or grid. In Volker Diekert and Bruno Durand, editors, *22nd Annual Symposium on Theoretical Aspects of Computer Science* (STACS'05), volume 3404 of *Lecture Notes in Computer Science*, pages 353–364. Springer, 2005.

[27] Josep Díaz, Xavier Pérez, Maria J. Serna, and Nicholas C. Wormald. Walkers on the cycle and the grid. Research Report CORR 2006-09, Faculty of Mathematics, University of Waterloo, 2006.

[28] Josep Díaz, Jordi Petit, and Maria J. Serna. Random geometric problems on $[0,1]^2$. In Michael Luby, José D. P. Rolim, and Maria J. Serna, editors, *2nd International Workshop Randomization and Approximation Techniques in Computer Science* (RANDOM'98), volume 1518 of *Lecture Notes in Computer Science*, pages 294–306. Springer, 1998.

[29] Olivier Dubois, Yacine Boufkhad, and Jacques Mandler. Typical random 3-sat formulae and the satisfiability threshold. *Electronic Colloquium on Computational Complexity*, 10(7), 2003.

[30] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17–61, 1960.

[31] Alan M. Frieze and Tomasz Łuczak. On the independence and chromatic numbers of random regular graphs. *J. Comb. Theory, Ser. B*, 54(1):123–132, 1992.

[32] Zhicheng Gao and Nicholas C. Wormald. Asymptotic normality determined by high moments, and submap counts of random maps. *Probability Theory and Related Fields*, 130(3):368–376, 2004.

[33] Michael R. Garey and David S. Johnson. *Computers and Intractability; A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., 1990.

[34] Edgar N. Gilbert. Random graphs. *Annals of Mathematical Statistics*, 30:1141–1144, 1959.

[35] Edgar N. Gilbert. Random plane networks. *Journal of the Society for Industrial and Applied Mathematics*, 9(4):533–543, 1961.

[36] Ashish Goel, Sanatan Rai, and Bhaskar Krishnamachari. Monotone properties of random geometric graphs have sharp thresholds. *Annals of Applied Probability*, 15(4):2535–2552, 2005.

[37] G. R. Grimmett and C. J. H. McDiarmid. On colouring random graphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, 77:313–324, 1975.

[38] Roch A. Guérin. Channel occupancy time distribution in a cellular radio system. *IEEE Transactions on Vehicular Technology*, 36(3):89–99, 1987.

[39] Leonidas J. Guibas, John Hershberger, Subhash Suri, and Li Zhang. Kinetic connectivity for unit disks. *Discrete & Computational Geometry*, 25(4):591–610, 2001.

[40] Piyush Gupta and Panganamala R. Kumar. Critical power for asymptotic connectivity in wireless networks. In W. M. McEneaney, G. G. Yin, and Q. Zhang, editors, *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming*, pages 547–566. Birkhäuser, 1999.

[41] Ramin Hekmat. *Ad-hoc Networks: Fundamental Properties and Network Topologies.* Springer, 2006.

[42] Lars Holst. On multiple covering of a circle with random arcs. *Journal of Applied Probability*, 17(1):284–290, 1980.

[43] Xiaoyan Hong, Mario Gerla, Guangyu Pei, and Ching-Chuan Chiang. A group mobility model for ad hoc wireless networks. In *2nd ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems* (MSWiM'99), pages 53–60, 1999.

[44] Alon Itai, Christos H. Papadimitriou, and Jayme Luiz Szwarcfiter. Hamilton paths in grid graphs. *SIAM Journal on Computing*, 11(4):676–686, 1982.

[45] Svante Janson. The second moment method, conditioning and approximation. *Random Discrete Structures*, 76:175–183, 1996.

[46] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Random Graphs.* John Wiley & Sons, 2000.

[47] Amit Jardosh, Elizabeth M. Belding-Royer, Kevin C. Almeroth, and Subhash Suri. Towards realistic mobility models for mobile ad hoc networks. In *9th Annual International Conference on Mobile Computing and Networking* (MOBICOM'03), pages 217–229. ACM Press, 2003.

[48] Esther Jennings and Clayton Okino. On the diameter of sensor networks. In *IEEE Aerospace Conference*, volume 3, pages 1211–1217, 2002.

[49] Alexis C. Kaporis, Lefteris M. Kirousis, and Yannis C. Stamatiou. A note on the non-colorability threshold of a random graph. *Electr. J. Comb.*, 7, 2000.

[50] Richard M. Karp. Probabilistic analysis of partitioning algorithm for the traveling salesman problem in the plane. *Mathematics of Operations Research*, 2(3):209–224, 1977.

[51] Graeme Kemkes and Nicholas C. Wormald. On the chromatic number of a 5-regular random graph. Research Report CORR 2005-22, Faculty of Mathematics, University of Waterloo, 2005.

[52] János Komlós and Endre Szemerédi. Limit distribution for the existence of hamilton cycles in random graphs. *Discrete Mathematics*, 43:55–63, 1983.

[53] Kwindla Hultman Kramer, Nelson Minar, and Pattie Maes. Tutorial: Mobile software agents for dynamic routing. *Mobile Computing and Communications Review*, 3(2):12–16, 1999.

[54] Florent Krząkała, Andrea Pagnani, and Martin Weigt. Threshold values, stability analysis and high-q asymptotics for the coloring problem on random graphs. *Physical Review E*, 70(046705), 2004.

[55] Tomasz Łuczak. The chromatic number of random graphs. *Combinatorica*, 11(1):45–54, 1991.

[56] Tomasz Łuczak and John C. Wierman. The chromatic number of random graphs at the double jump threshold. *Combinatorica*, 9(1):39–49, 1989.

[57] David W. Matula and Luděk Kučera. An expose-and-merge algorithm and the chromatic number of a random graph. In M. Karoński, J. Jaworski, and A. Ruciński, editors, *Random Graphs '87*, pages 175–187. John Wiley & Sons, 1990.

[58] Colin McDiarmid. Random channel assignment in the plane. *Random Structures and Algorithms*, 22(2):187–212, 2003.

[59] Seapahn Meguerdichian, Farinaz Koushanfar, Miodrag Potkonjak, and Mani B. Srivastava. Coverage problems in wireless ad-hoc sensor networks. In *Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies* (INFOCOM'01), volume 3, pages 1380–1387, 2001.

[60] Marc Mézard and Riccardo Zecchina. Random k-satisfiability: from an analytic solution to a new efficient algorithm. *Physical Review E*, 66(056126), 2002.

[61] Michael Molloy. The chromatic number of sparse random graphs. Master's thesis, University of Waterloo, 1992.

[62] Philippe Nain, Don Towsley, Benyuan Liu, and Zhen Liu. Properties of random direction models. *24th Annual Joint Conference of the IEEE Computer and Communications Societies* (INFOCOM'05), 3:1897–1907, 2005.

[63] Andrew Odlyzko. Asymptotic enumeration methods. In Ronald L. Graham, Martin Grötschel, and László Lovász, editors, *Handbook of Combinatorics*, volume 2, pages 1063–1229. North-Holland, 1995.

[64] Mathew D. Penrose. The longest edge of the random minimal spanning tree. *Annals of Applied Probability*, 7(2):340–361, 1997.

[65] Mathew D. Penrose. On k-connectivity for a geometric random graph. *Random Structures and Algorithms*, 15:145–164, 1999.

[66] Mathew D. Penrose. *Random Geometric Graphs*, volume 5 of *Oxford Studies in Probability*. Oxford University Press, 2003.

[67] Xavier Pérez. Empirical analysis of the connectivity threshold of mobile agents on the grid. In Sotiris E. Nikoletseas, editor, *4th International Workshop on Efficient and Experimental Algorithms* (WEA'05), volume 3503 of *Lecture Notes in Computer Science*, pages 253–264. Springer, 2005.

[68] Jordi Petit. *Layout Problems*. PhD thesis, Universitat Politècnica de Catalunya, 2001.

[69] Boris Pittel, Joel Spencer, and Nicholas Wormald. Sudden emergence of a giant k-core in a random graph. *Journal of Combinatorial Theory Series B*, 67(1):111–151, 1996.

[70] Robert W. Robinson and Nicholas C. Wormald. Almost all cubic graphs are hamiltonian. *Random Structures and Algorithms*, 3(2):117–126, 1992.

[71] Robert W. Robinson and Nicholas C. Wormald. Almost all regular graphs are hamiltonian. *Random Structures and Algorithms*, 5(2):363–374, 1994.

[72] Elizabeth M. Royer, P. Michael Melliar-Smith, and Louise E. Moser. An analysis of the optimum node density for ad hoc mobile networks. In *IEEE International Conference on Communications* (ICC'01), volume 3, pages 857–861, 2001.

[73] Paolo Santi and Doug M. Blough. An evaluation of connectivity in mobile wireless ad hoc networks. In *33rd International Conference on Dependable Systems and Networks*, pages 89–98, 2002.

[74] Lingsheng Shi and Nicholas C. Wormald. Colouring random 4-regular graphs. *Combinatorics, Probability and Computing*. (in press).

[75] Lingsheng Shi and Nicholas C. Wormald. Colouring random regular graphs. *Combinatorics, Probability and Computing*. (in press).

[76] J. Michael Steele. *Probability theory and combinatorial optimization*, volume 69 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics, 1997.

[77] Roderick Wong. *Asymptotic approximations of integrals*, volume 34 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics, 2001. Corrected reprint of the 1989 original.

[78] Nicholas C. Wormald. Models of random regular graphs. In J. D. Lamb and D. A. Preece, editors, *Surveys in Combinatorics, 1999*, volume 267 of *London Mathematical Society Lecture Note Series*, pages 239–298. Cambridge University Press, 1999.

[79] Nicholas C. Wormald. Random graphs and asymptotics. In Jonathan L. Gross and Jay Yellen, editors, *Handbook of Graph Theory*, Discrete Mathematics and its Applications, pages 817–836. CRC Press, 2004.

[80] Feng Xue and Panganamala R. Kumar. The number of neighbors needed for connectivity of wireless networks. *Wireless Networks*, 10(2):169–181, 2004.

[81] Feng Xue and Panganamala R. Kumar. *Scaling Laws for Ad Hoc Wireless Networks: An Information Theoretic Approach.* NOW Publishers, 2006.

[82] Joseph E. Yukich. *Probability Theory of Classical Euclidean Optimization Problems*, volume 1675 of *Lecture Notes in Mathematics*. Springer, 1998.