Eigenvalue fluctuations for random regular graphs

Tobias Lee Johnson

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2014

Reading Committee:

Ioana Dumitriu, Chair

Soumik Pal, Chair

Sara Billey

Program Authorized to Offer Degree:
Department of Mathematics

University of Washington

Abstract

Eigenvalue fluctuations for random regular graphs

Tobias Lee Johnson

Co-Chairs of the Supervisory Committee:
Associate Professor Ioana Dumitriu
Department of Mathematics

Associate Professor Soumik Pal
Department of Mathematics

One of the major themes of random matrix theory is that many asymptotic properties of traditionally studied distributions of random matrices are *universal.* We probe the edges of universality by studying the spectral properties of random regular graphs. Specifically, we prove limit theorems for the fluctuations of linear spectral statistics of random regular graphs. We find both universal and non-universal behavior. Our most important tool is Stein's method for Poisson approximation, which we develop for use on random regular graphs.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS

I owe a debt to many people for their help and friendship. I thank Ioana for introducing me to research and to random regular graphs. I'm grateful to her for supporting me and for not taking my curmudgeonliness too seriously. I'm grateful to Soumik for teaching me probability and shaping my view of it. His suggestion that I learn how to use Stein's method and try to apply it to cycle counts of random regular graphs has shaped my mathematical career more than I would ever have imagined. I also thank Chris Hoffman, who has been a third advisor to me this year, introducing me to new areas of math and collaborating with me. I can't thank Elliot Paquette and Matt Junge enough. It's been wonderful to have them as friends and collaborators. Most of all, I thank my friends and my wife Lindsay.

Chapter 1

INTRODUCTION

1.1 How universal is universality?

Random matrix theory traditionally studies certain random matrices of interest to physicists and statisticians. The central question of classical random matrix theory is to prove that the eigenvalues of random matrices' show *universal* behavior as the size of the random matrices grow. Universality is not a precise concept. The classical central limit theorem gives an example of it: with only light conditions on a collection of random variables (being i.i.d. with finite variances), their centered and normalized sums converge in law to Gaussian.

The most basic symmetric random matrix model is the *Gaussian Orthogonal Ensemble*, abbreviated GOE. Let G be an $n \times n$ matrix whose entries are independent and distributed as $N(0, 2)$. Define X as $(G + G^T)/2$, a random symmetric matrix with independent entries on and above the diagonal. The random matrix X has centered Gaussian entries with variance 1 above the diagonal and variance 2 on the diagonal, and it is said to be drawn from the GOE. Any $n \times n$ random matrix with centered independent entries on and above the diagonal and variance 1 entries above the diagonal is called a *Wigner matrix*. (The word "ensemble" does not have any precise meaning, but it is usually refers to a collection of probability distributions on $n \times n$ matrices, as n ranges from 1 to infinity. Each distribution typically obeys some sort of invariance. For instance, if O is an arbitrary orthogonal matrix and X is drawn from the GOE, then OX has the same distribution as X.)

An example of universality for random matrices is that the eigenvalues of $n \times n$ Wigner matrices show the same limiting behavior as those of matrices from the GOE as $n \to \infty$. Most results along these lines were confirmed only recently, in a series of papers including [TV11, TV10, ESY09b, ESY09a, EPR$^+$10, ERSY10, ERS$^+$10].

The adjacency matrix of a random regular graph is similar to a Wigner matrix, but its entries are uncentered and lightly dependent. How does this affect the adjacency matrix's spectral properties? To put it another way, how universal is universality of random matrices? This is our main motivation for investigating properties of eigenvalues of random regular graphs from the perspective of random matrix theory.

1.2 Stein's method applied to random regular graphs

Graph eigenvalues have a close connection to the graph's structural properties (see [Chu97, Spi12]). We exploit this by determining spectral properties of random regular graphs by looking at the distribution of their cycle counts. The main novelty of our approach is the use of *Stein's method*, which to our knowledge had never been applied to random regular graphs before. Stein's method is a collection of techniques for distributional approximation. Stein's method naturally gives not just asymptotic results but also quantitative error bounds on the approximations. This was essential for the eigenvalue fluctuation results described in this thesis.

Stein's method was originally developed by Charles Stein for normal approximation; its first published use is [Ste72]. Louis Chen adapted the method for Poisson approximation [Che75]. Because of this, Stein's method is sometimes called the Stein-Chen or Chen-Stein method when applied to Poisson approximation. Now that Stein's method is understood in a more general and applied to a wide range of distributions, it is more typical to see it called just Stein's method, regardless of the type of approximation. The survey paper [Ros11] gives a broad introduction to Stein's method, and [CDM05] and [BHJ92] focus specifically on using it for Poisson approximation, as we do in this thesis.

The classical scenario for Poisson approximation is for sums of increasingly many, increasingly unlikely independent indicators: in other words, the convergence of $\mathrm{Bin}(n, \lambda/n)$ to $\mathrm{Poi}(\lambda)$ as $n \to \infty$. There are several approaches to Stein's method for Poisson approximation, each allowing this approximation to hold in the presence of some dependence. The most straightforward is the local approach: each indicator is independent of all others but a small "neighborhood". This was the original approach in [Che75], and it is generalized and put in a very usable form in [AGG89]. This approach does not seem to work in the context of random regular graphs, where nearly everything is lightly dependent on everything else. Another approach is size-bias coupling. This theory is developed at length for Poisson approximation in [BHJ92], though it is not viewed through the lens of size-biasing there. See[Ros11] and [AGK13] for how it fits into this framework. We use this method on the *permutation model* of random regular graph (see Section 1.4 for its definition). Another technique is the method of exchangeable pairs; see [CDM05] and [Ros11] for good expositions. This technique is perhaps the most flexible and the most finicky of the three. We use it for Poisson approximation in the *uniform model* of random regular graph, defined in Section 1.4. This technique has some clear similarities to a combinatorial technique called *the method of switchings*, and we make some rigorous connections between the two.

1.3 The results of this thesis

Consider an $n \times n$ Wigner random matrix X_n (a symmetric matrix with independent, mean zero, variance one entries above the diagonal). Choose an interval in the real line, and let N_n denote the number of eigenvalues of $n^{-1/2}X_n$ lying in this interval. A fundamental result in random matrix theory is that N_n/n converges in probability to a deterministic value as n tends to infinity. This value is the measure of the interval under Wigner's semicircle law, the measure on $[-2,2]$ given by the density $\frac{1}{2\pi}\sqrt{4-x^2}\,dx$. This measure is a universal limit, in the sense that it does not depend on the distributions of the individual matrix entries, besides their means and variances.

The analogue of this result for random regular graphs appears in [McK81]: Let $\lambda_1,\ldots,\lambda_n$ be the eigenvalues of a random d-regular graph on n vertices. If $f\colon \mathbb{R}\to\mathbb{R}$ is an indicator on an interval or is bounded and continuous, then as $n\to\infty$,

$$\frac{1}{n}\sum_{i=1}^{n} f(\lambda_i) \xrightarrow{pr} \int_{-2\sqrt{d-1}}^{2\sqrt{d-1}} f(x)p_d(x)\,dx.$$

The limiting measure $p_d(x)\,dx$ is not the semicircle law, but a different measure known now as the Kesten–McKay law. Its density is given on $|x|\le 2\sqrt{d-1}$ by

$$p_d(x) = \frac{d\sqrt{4(d-1)-x^2}}{2\pi(d^2-x^2)}. \tag{1.1}$$

The expression $\sum f(\lambda_i)$ is called a *linear eigenvalue statistic.*

The topic of this thesis is the second-order behavior of these linear statistics. We will show that when the degree of the random graphs is held fixed, their fluctuations converge to compound Poisson distributions, in contrast to the Gaussian limit known for Wigner matrices. If the degree grows with the size of the graph, however, the limit of the fluctuations is Gaussian, in line with the universal behavior. We show that this holds in two models of random regular graphs, defined in Section 1.4.

The path to these results is through an analysis of the distribution of cycle counts in these models by Stein's method. These results are interesting in their own right, and they make up Chapter 2. In Chapter 3, we apply them to prove the eigenvalue fluctuation results.

In Chapter 4, we consider a process of growing random regular graphs. The eigenvalue fluctuations are then a stochastic process whose marginals are given by the results of Chapter 3. This is analogous to a *corners process* in random matrix theory; see [BG13] for a good introduction. The idea is to think of a sequence of random matrices as the principal minors of an infinite random matrix. One can then consider not just the marginal distribution of the eigenvalues of each random matrix, but the joint distribution of eigenvalues of a matrix and its minors. The limiting fluctuations of some of these processes can be expressed in terms of the Gaussian free field [Bor10a, Bor10b, BG13]. We show that the same holds for the eigenvalues of the growing random regular graphs.

Most of this thesis is joint work. Chapters 2 and 3 are a synthesis of [DJPP13], [JP12], and [Joh12]. The results on the permutation model are from [DJPP13], which is joint with Ioana Dumitriu, Elliot Paquette, and Soumik Pal, and from [JP12], which is joint with Pal. The results on the uniform model are from [Joh12]. (See Section 1.4 for the definitions of these two models of random regular graphs). Theorem 2.4, a version of [JP12, Corollary 24i] with an improved rate, appears only in this thesis.

Chapter 4 is mostly taken from [Joh12], which is joint work with Pal. Section 4.5 is new to this thesis and was also done jointly with Pal. (The exception is Section 4.5.2, an extended introduction to the Gaussian free field. It and any errors contained in it are mine alone.) The main result here is Theorem 4.6, which shows the convergence of eigenvalue fluctuations to the Gaussian free field in a more explicit form than in [JP12].

1.4 Models of random regular graphs

In Chapters 2 and 3, we will present results on two models of random regular graphs, the *permutation mode* and the *uniform model*. Traditionally, combinatorialists were most concerned with the uniform model of random regular graphs. The permutation model is typically easier to work with, however, and it is the setting for many spectral results on random regular graphs (for example, [BS87, Fri91, Fri08]). There seems to have been a sense that that the two models had basically the same properties, besides the permutation model having loops and multiple edges. The contiguity result in [GJKW02] justifies this somewhat.

We now review the definitions of these two models and of our sequence of growing graphs.

1.4.1 The uniform model

A random d-regular graph on n vertices drawn from the uniform model is just a graph chosen uniformly from the set of all d-regular graphs (i.e., graphs where every vertex has degree exactly d) on n vertices without loops or multiple edges. Such graphs only exist when nd is even.

1.4.2 The permutation model

The permutation model is given by choosing $d/2$ independent, uniformly random permutations on n vertices, making a graph from the cycle structure of each permutation, and overlaying them. It exists only for even values of d. For a more formal definition, let $\pi_1, \ldots, \pi_{d/2}$ be independent, uniformly random permutations on n vertices. Define a graph on vertices $\{1, \ldots, n\}$ by making an edge between vertices x and y for every k such that $\pi_k(x) = y$. This model allows loops and multiple edges. We consider a loop at vertex x as counting as two edges when computing the degree of x, so that the graph really is d-regular. We also count a loop at vertex i as increasing the graph's adjacency matrix by 2 at position (i, i). The adjacency matrix of a graph from this model is then a sum of independent permutation matrices.

1.4.3 Growing random regular graphs

A *tower of random permutations* is a sequence of random permutations $(\pi^{(n)}, n \in \mathbb{N})$ such that

(i) $\pi^{(n)}$ is a uniformly distributed random permutation of $\{1, \dots, n\}$, and

(ii) for each n, if $\pi^{(n)}$ is written as a product of cycles then $\pi^{(n-1)}$ is derived from $\pi^{(n)}$ by deletion of the element n from its cycle.

The stochastic process that grows $\pi^{(n)}$ from $\pi^{(n-1)}$ by sequentially inserting an element n randomly is called the Chinese Restaurant Process. We give a further review of it in Section 4.1.1.

Now suppose we construct towers of random permutations $(\pi_d^{(n)}, n \geq 1)$, independent for each d. For any n and d, we can define a random $2d$-regular graph $G(n, 2d)$ from $\{\pi_j^{(n)}, 1 \leq j \leq d\}$ as in Section 1.4.2. Marginally, $G(n, 2d)$ is then a random graph from the permutation model. We will often keep d fixed and consider n as a growing parameter, referring to $G(n, 2d)$ as G_n. Here and later, G_0 will represent the empty graph.

We construct a continuous-time version of this by inserting new vertices into G_n with rate $n + 1$. Formally, define independent times $T_i \sim \mathrm{Exp}(i)$, and let

$$M_t = \max\left\{m\colon \sum_{i=1}^{m} T_i \leq t\right\},$$

and define the continuous-time Markov chain $G(t) = G_{M_t}$. When we vary d as well as n, we will also refer to this as $G(t, 2d)$.

Chapter 2

POISSON APPROXIMATION FOR CYCLE COUNTS IN RANDOM REGULAR GRAPHS

Let C_k denote the number of cycles of length k in a random graph G_n. The distribution of these random variables has been studied since [Bol80, Wor81], where it was proven that if G is a uniform d-random regular graph on n vertices, then $(C_3, \ldots, C_r)$ converges in law to a vector of independent Poisson random variables as n tends to infinity, with r held fixed.

The strongest results on the cycle counts of a random regular graph came in [MWW04], where the Poisson approximation was shown to hold even as $d = d(n)$ and $r = r(n)$ grow with n, so long as $(d-1)^{2r-1} = o(n)$. This is a natural boundary: in this asymptotic regime, all cycles in G_n of length r or less have disjoint edges, asymptotically almost surely. If $(d-1)^{2r-1}$ grows any faster, this fails. This led the authors in [MWW04] to speculate that the Poisson approximation failed beyond this threshold. Surprisingly, this is not the case. We will show that the Poisson approximation holds slightly beyond this threshold. We also give a quantitative bound on the accuracy of the approximation, which was our original motivation and is the necessary ingredient for our results on linear eigenvalue statistics.

We will give results on both the permutation model and the uniform model of random regular graphs. We use Stein's method in both cases, but we use different techniques for the two models: size-biased couplings for the permutation model and exchangeable pairs for the uniform model. We provide background and references on these techniques in the following section.

Before we go any further, we present the main results of this section. Rather than showing that the cycle counts are approximately Poisson, we will make a more general statement about process made up of the cycles themselves. To state our results, we must explain exactly what we mean by a cycle in a graph.

We start by discussing the permutation model. Let G_n be a random $2d$-regular graph on n vertices from the permutation model, formed from the independent permutations $\pi_1, \ldots, \pi_d$ as described in Section 1.4. This graph can be considered as a directed, edge-labeled graph in a natural way. If $\pi_l(i) = j$, then by definition G_n contains an edge between i to j. When convenient, we consider this edge to be directed from i to j and to be labeled by π_l.

Consider a walk on G_n, viewed in this way, and imagine writing down the label of each edge as it is traversed, putting π_i or π_i^{-1} according to the direction we walk over the edge. We call a walk *closed* if it starts and ends at the same vertex, and we call a closed walk a cycle it never visits a vertex twice (besides the first and last one), and it never traverses an edge more than once in either direction. Thus the word $w = w_1 \cdots w_k$ formed as a cycle is traversed is cyclically reduced, i.e., $w_i \neq w_{i+1}^{-1}$ for all i, considering i modulo k. For example, following an edge and then immediately backtracking does not form a 2-cycle, and the word formed by this walk is $\pi_i \pi_i^{-1}$ or $\pi_i^{-1}\pi_i$ for some i, which is not cyclically reduced. We consider two cycles equivalent if they are both walks on an identical set of edges; that is, we ignore the starting vertex and the direction of the walk. We will often denote the length of a cycle α by $|\alpha|$.

Definition 2.1. Let $\mathcal{I}_k$ be the set of all k-cycles in the complete graph on n vertices with edges labeled by $\pi_1^{\pm 1}, \ldots, \pi_d^{\pm 1}$, where the word formed as the cycle is traversed is cyclically reduced. Let $a(d,k)$ be number of cyclically reduced words of length k in this alphabet.

Observe that $|\mathcal{I}_k| = [n]_k a(d,k)/2k$, where $[n]_k = n(n-1)\cdots(n-k+1)$. By an inclusion-exclusion argument [DJPP13, Lemma 41],

$$a(d,k) = \begin{cases} (2d-1)^k - 1 + 2d & \text{if } k \text{ is even,} \\ (2d-1)^k + 1 & \text{if } k \text{ is odd.} \end{cases} \tag{2.1}$$

We are now ready to state the main Poisson approximation result for the permutation model.

Theorem 2.2 (Theorem 14 in [JP12])**.** *Let G_n be a random $2d$-regular graph on n vertices from the permutation model. Let $\mathcal{I} = \bigcup_{k=1}^{r} \mathcal{I}_k$ for some integer r. For any cycle $\alpha \in \mathcal{I}$, let $I_\alpha = 1\{G_n \text{ contains } \alpha\}$, and let $\mathbf{I} = (I_\alpha,\, \alpha \in \mathcal{I})$. Let $\mathbf{Z} = (Z_\alpha,\, \alpha \in \mathcal{I})$ be a vector whose* *coordinates are independent Poisson random variables with* $\mathbf{E}Z_\alpha = 1/[n]_k$ *for* $\alpha \in \mathcal{I}_k$*. Then for all* $d \geq 2$ *and* $n, r \geq 1$*,*

$$d_{TV}(\mathbf{I}, \mathbf{Z}) \leq \frac{c(2d-1)^{2r-1}}{n}$$

for some absolute constant c*.*

In the uniform model, there are no edge labels, and a cycle is simply a closed walk repeating no vertices. Again, we consider two walks equivalent if they are walks on the same set of edges.

Theorem 2.3 (Corollary 8 in [Joh12])**.** *Let* G_n *be a random* d*-regular graph on* n *vertices from the uniform model, and let* $\mathcal{I}$ *be the collection of all cycles of length* r *or less in the complete graph* K_n*. For any cycle* $\alpha \in \mathcal{I}$*, let* $I_\alpha = 1\{G_n \text{ contains } \alpha\}$*, and let* $\mathbf{I} = (I_\alpha,\, \alpha \in \mathcal{I})$*. Let* $\mathbf{Z} = (Z_\alpha,\, \alpha \in \mathcal{I})$ *be a vector whose coordinates are independent Poisson random variables with* $\mathbf{E}Z_\alpha = (d-1)^{|\alpha|}/[n]_{|\alpha|}$*. For some absolute constant* c*, for all* n *and* $d, r \geq 3$*,*

$$d_{TV}(\mathbf{I},\, \mathbf{Z}) \leq \frac{c(d-1)^{2r-1}}{n}.$$

These theorems immediately imply that the vectors of cycle counts of length r or less in the permutation and uniform models are also within $O((2d-1)^{2r-1}/n)$ and $O((d-1)^{2r-1}/n)$, respectively, of vectors of independent Poissons. In fact, we can do slightly better:

Theorem 2.4. *Let* G_n *be a random* $2d$*-regular graph on* n *vertices from the permutation model with cycle counts* $(C_k,\, k \geq 1)$*. Let* $Z_k,\, k \geq 1$ *be independent Poisson random variables with* $\mathbf{E}Z_k = a(d,k)/2k$*. For any* $d \geq 2$ *and* $n, r \geq 1$*,*

$$d_{TV}\big((C_1, \ldots, C_r),\, (Z_1, \ldots, Z_r)\big) \leq \frac{cr^2(2d-1)^r \log(2d-1)}{n},$$

for some absolute constant c*.*

Theorem 2.5 (Theorem 11 in [DJPP13])**.** *Let* G_n *be a random* d*-regular graph on* n *vertices from the uniform model with cycle counts* $(C_k,\, k \geq 3)$*. Let* $(Z_k,\, k \geq 3)$ *be independent Poisson random variables with* $\mathbf{E}Z_k = (d-1)^k/2k$*. For any* $n \geq 1$ *and* $r, d \geq 3$*,*

$$d_{TV}\big((C_3, \ldots, C_r),\, (Z_3, \ldots, Z_r)\big) \leq \frac{c\sqrt{r}(d-1)^{3r/2-1}}{n}$$

for some absolute constant c*.*

2.1 Background on Stein's method

2.1.1 Size-bias couplings

To give some intuition behind size-bias couplings, let us go to the archetypal setting for Poisson approximation. Let $I_1, \dots, I_n$ be independent Bernoulli random variables, equal to 1 with probability $1/n$ and 0 with probability $(n-1)/n$. Let X be the sum of these, which makes its distribution $\mathrm{Bin}(n, 1/n)$. Define X' to be $X - I_N + 1$, where N is uniformly chosen from $\{1, \dots, n\}$, independently of everything else. In other words, X' is given by taking one of the indicators at random and forcing it to be 1. It is not hard to show that X' is a size-biased version of X, meaning that

$$\mathbf{P}[X' = k] = \frac{k}{\mathbf{E}X}\mathbf{P}[X = k].$$

For large n, we have $X' \approx 1 + X$.

This definition of a size-biased version of X defined on the same probability space is an example of a more general construction; see [Ros11, Section 3.4.1]. In general, if a random variable X can be coupled with X', a size-biased version of itself, and X' is close to $X + 1$ in L^1, then X is approximately Poisson. A precise verison of this statement is [Ros11, Theorem 4.13].

We will use a formulation of this idea from [BHJ92]. This formulation never explicitly make a size-biased version of the random variable to be approximated, but its idea is exactly the same. Recall the definition of $(I_\beta,\, \beta \in \mathcal{I})$ from Theorem 2.2. For each $\alpha \in \mathcal{I}$, let $(J_{\beta\alpha},\, \beta \in \mathcal{I})$ be distributed as $(I_\beta,\, \beta \in \mathcal{I})$ conditioned on $I_\alpha = 1$. The goal is to construct a coupling of $(I_\beta,\, \beta \in \mathcal{I})$ and $(J_{\beta\alpha},\, \beta \in \mathcal{I})$ so that the two random vectors are "close together". We hope that for each $\alpha \in \mathcal{I}$, the cycles in $\mathcal{I} \setminus \{\alpha\}$ can be partitioned into two sets $\mathcal{I}_\alpha^-$ and $\mathcal{I}_\alpha^+$ such that

$$J_{\beta\alpha} \le I_\beta \quad \text{if } \beta \in \mathcal{I}_\alpha^-, \tag{2.2}$$

$$J_{\beta\alpha} \ge I_\beta \quad \text{if } \beta \in \mathcal{I}_\alpha^+. \tag{2.3}$$

If this is the case, then one can approximate $(I_\beta,\, \beta \in \mathcal{I})$ by a Poisson process by calculating $\mathbf{Cov}(I_\alpha, I_\beta)$ for every $\alpha, \beta \in \mathcal{I}$, according to the following proposition.

Proposition 2.6 (Corollary 10.J.1 in [BHJ92]). *Suppose that* $\mathbf{I} = (I_\alpha,\ \alpha \in \mathfrak{I})$ *is a vector of 0-1 random variables with* $\mathbf{E}I_\alpha = p_\alpha$. *Suppose that* $(J_{\beta\alpha},\ \beta \in \mathfrak{I})$ *is distributed as described above, and that for each* α *there exists a partition and a coupling of* $(J_{\beta\alpha},\ \beta \in \mathfrak{I})$ *with* $(I_\beta,\ \beta \in \mathfrak{I})$ *such that* (2.2) *and* (2.3) *are satisfied.*

Let $\mathbf{Y} = (Y_\alpha,\ \alpha \in \mathfrak{I})$ *be a vector of independent Poisson random variables with* $\mathbf{E}Y_\alpha = p_\alpha$. *Then*

$$d_{TV}(\mathbf{I}, \mathbf{Y}) \leq \sum_{\alpha\in\mathfrak{I}} p_\alpha^2 + \sum_{\alpha\in\mathfrak{I}} \sum_{\beta\in\mathfrak{I}_\alpha^-} |\mathbf{Cov}(I_\alpha, I_\beta)| + \sum_{\alpha\in\mathfrak{I}} \sum_{\beta\in\mathfrak{I}_\alpha^+} \mathbf{Cov}(I_\alpha, I_\beta). \tag{2.4}$$

By bunching together indicators into bins, we can slightly improve the rates:

Proposition 2.7 (Theorem 10.K in [BHJ92]). *Assume all the conditions of the previous proposition. Suppose that we partition the index set as* $\mathfrak{I} = \bigcup_{k=1}^r \mathfrak{I}_k$, *and define*

$$W_k = \sum_{\alpha\in\mathfrak{I}_k} I_k, \qquad\qquad Y_k = \sum_{\alpha\in\mathfrak{I}_k} Y_\alpha.$$

Let $\lambda_k = \mathbf{E}Y_k$.

$$\begin{aligned} &d_{TV}\big((W_1, \ldots, W_r),\ (Y_1, \ldots, Y_r)\big) \\ &\qquad \leq 2(1 + e^{-1}\log^+ \max \lambda_j)\left(\sum_{k=1}^r \sum_{\alpha\in\mathfrak{I}_k} \frac{p_\alpha^2}{\lambda_k} + \sum_{j,k=1}^r \frac{A(j,k)}{\sqrt{\lambda_j\lambda_k}}\right), \end{aligned} \tag{2.5}$$

where

$$A(j,k) = \sum_{\alpha\in\mathfrak{I}_k}\left(\sum_{\beta\in\mathfrak{I}_\alpha^-\cap\mathfrak{I}_j} |\mathbf{Cov}(I_\alpha, I_\beta)| + \sum_{\beta\in\mathfrak{I}_\beta^+\cap\mathfrak{I}_j} \mathbf{Cov}(I_\alpha, I_\beta)\right)$$

2.1.2 Exchangeable pairs and switchings

For our Poisson approximation of cycle counts in the uniform model, we will use a different form of Stein's known as the method of exchangeable pairs. As we lay out the background necessary to apply Stein's method by exchangeable pairs, we will also explain a connection between this method and a combinatorial technique for asymptotic enumeration called the method of switchings.

The method of switchings, pioneered by Brendan McKay and Nicholas Wormald, has been applied to asymptotically enumerate combinatorial structures that defy exact counts, including Latin rectangles [GM90] and matrices with prescribed row and column sums [McK84, MW03, GMW06]. It has seen its biggest use in analyzing regular graphs; see [KSVW01], [MWW04], [KSV07], and [BSK09] for some examples. A good summary of switchings in random regular graphs can be found in Section 2.4 of [Wor99a].

The basic idea of the method is to choose two families of objects, A and B, and investigate only their relative sizes. To do this, one defines a set of switchings that connect elements of A to elements of B. If every element of A is connected to roughly p objects in B, and every element in B is connected to roughly q objects in A, then by a double-counting argument, $|A|/|B|$ is approximately q/p. When the objects in question are elements of a probability space, this gives an estimate of the relative probabilities of two events.

Stein's method (sometimes called the Stein-Chen method when used for Poisson approximation) is a powerful and elegant tool to compare two probability distributions. It was originally developed by Charles Stein for normal approximation; its first published use is [Ste72]. Louis Chen adapted the method for Poisson approximation [Che75]. Since then, Stein, Chen, and a score of others have adapted Stein's method to a wide variety of circumstances. The survey paper [Ros11] gives a broad introduction to Stein's method, and [CDM05] and [BHJ92] focus specifically on using it for Poisson approximation.

We will use the technique of exchangeable pairs, following the treatment in [CDM05]. Suppose we want to bound the distance of the law of X from the Poisson distribution. The technique is to introduce an auxiliary randomization to X to get a new random variable X' so that X and X' are exchangeable (that is, (X, X') and (X', X) have the same law). If X and X' have the right relationship—specifically, if they behave like two steps in an immigration-death process whose stationary distribution is Poisson—then Stein's method gives an easy proof that X is approximately Poisson.

Switchings and Stein's method have bumped into each other several times. For instance, both techniques have been used to study Latin rectangles [Ste78, GM90], and the analysis of random contingency tables in [DS98] is similar to combinatorial work like [GM08]. Nevertheless, we believe that this is the first explicit connection between the two techniques. The essential idea is to use a random switching as the auxiliary randomization in constructing an exchangeable pair.

We believe the connection between switchings and Stein's method may prove profitable to users of both techniques. Using Stein's method in conjunction with a switchings argument allows for a quantitative bound on the accuracy of the approximation. Stein's method can also be used for approximation by other distributions besides Poisson, and for proving concentration bounds (see [Cha07]). On the other hand, Stein's method cannot prove results as sharp as [MWW04, Theorem 2], which gives an extremely accurate bound on the probability that a random graph has no cycles of length r or less. The bare-hands switching arguments used there might be useful to anyone who needs a particularly sharp bound on a Poisson approximation at a single point (see [JP13, Proposition 1.7]).

Now, we give the background we need on Stein's method of exchangeable pairs. Recall that the main idea of Stein's method of exchangeable pairs is to perturb a random variable X to get a new random variable X', and then to examine the relationship between the two. The basic heuristic is that if (X, X') is exchangeable and

$$\mathbf{P}[X' = X + 1 \mid X] \approx \frac{\lambda}{c},$$
$$\mathbf{P}[X' = X - 1 \mid X] \approx \frac{X}{c},$$

for some constant c, then X is approximately Poisson with mean λ. (When X and X' are exactly Poisson with mean λ and are two steps in an immigration-death chain whose stationary distribution is that, these equations hold exactly.) The following proposition gives a precise, multivariate version of this heuristic.

Proposition 2.8 ([CDM05, Proposition 10])**.** *Let $W = (W_1, \dots, W_r)$ be a random vector taking values in $\mathbb{N}^r$, and let the coordinates of $Z = (Z_1, \dots, Z_r)$ be independent Poisson random variables with $\mathbf{E}Z_k = \lambda_k$. Let $W' = (W'_1, \dots, W'_r)$ be defined on the same space as W, with (W, W') an exchangeable pair.*

For any choice of σ-algebra $\mathcal{F}$ with respect to which W is measurable and any choice of constants c_k,

$$d_{TV}(W, Z) \leq \sum_{k=1}^{r} \xi_k \Big(\mathbf{E}\big|\lambda_k - c_k \mathbf{P}[\Delta_k^+ \mid \mathcal{F}]\big| + \mathbf{E}\big|W_k - c_k \mathbf{P}[\Delta_k^- \mid \mathcal{F}]\big| \Big),$$

with $\xi_k = \min(1, 1.4\lambda_k^{-1/2})$ *and*

$$\Delta_k^+ = \{W_k' = W_k + 1,\ W_j = W_j' \text{ for } k < j \le r\},$$
$$\Delta_k^- = \{W_k' = W_k - 1,\ W_j = W_j' \text{ for } k < j \le r\}.$$

Remark 2.9. We have changed the statement of the proposition from [CDM05] in two small ways: we condition our probabilities on $\mathcal{F}$, rather than on W, and we do not require that $\mathbf{E}W_k = \lambda_k$ (though the approximation will fail if this is far from true). Neither change invalidates the proof of the proposition.

Remark 2.10. There is a direct connection between switchings and a certain bare-hands version of Stein's method. Though this is not what we use in this paper, it is helpful in understanding why Stein's method and the method of switchings are so similar. If (X, X') is exchangeable, then as explained in [Ste92, Section 2], one can directly investigate ratios of probabilities of different values of X using the equation

$$\frac{\mathbf{P}[X = x_1]}{\mathbf{P}[X = x_2]} = \frac{\mathbf{P}[X' = x_1 \mid X = x_2]}{\mathbf{P}[X' = x_2 \mid X = x_1]}.$$

This technique bears a strong resemblance to the method of switchings: if we think of X as some property of a random graph (for example, number of cycles) and X' as that property after a random switching has been applied, then this formula instructs us to count how many switchings change X from x_1 to x_2 and vice versa, just as one does when using switchings for asymptotic enumeration.

2.2 Poisson approximation in the permutation model

We introduce two lemmas. The first gives a bound on the distance between Poisson random variables with almost the same means, and the second provides a technical bound that we need.

Lemma 2.11. *Let* $\mathbf{Y} = (Y_\alpha,\, \alpha \in \mathcal{I})$ *and* $\mathbf{Z} = (Z_\alpha,\, \alpha \in \mathcal{I})$ *be vectors of independent Poisson random variables. Then*

$$d_{TV}(\mathbf{Y},\, \mathbf{Z}) \le \sum_{\alpha \in \mathcal{I}} |\mathbf{E}Y_\alpha - \mathbf{E}Z_\alpha|.$$

Proof. We will apply the Stein-Chen method directly. Define the operator $\mathcal{A}$ by

$$\mathcal{A}h(x) = \sum_{\alpha\in\mathcal{I}} \mathbf{E}[Z_\alpha]\big(h(x+e_\alpha)-h(x)\big) + \sum_{\alpha\in\mathcal{I}} x_\alpha\big(h(x-e_\alpha)-h(x)\big)$$

for any $h\colon \mathbb{Z}_+^{|\mathcal{I}|}\to\mathbb{R}$ and $x\in\mathbb{Z}_+^{|\mathcal{I}|}$. This is the Stein operator for the law of $\mathbf{Z}$, and $\mathbf{E}\mathcal{A}h(\mathbf{Z})=0$ for any bounded function h. By Proposition 10.1.2 and Lemma 10.1.3 in [BHJ92], for any set $A\subseteq\mathbb{Z}_+^{|\mathcal{I}|}$, there is a function h such that

$$\mathcal{A}h(x) = 1\{x\in A\} - \mathbf{P}[\mathbf{Z}\in A],$$

and this function has the property that

$$\sup_{\substack{x\in\mathbb{Z}_+^{|\mathcal{I}|}\\ \alpha\in\mathcal{I}}} |h(x+e_\alpha)-h(x)| \le 1. \tag{2.6}$$

Thus we can bound the total variation distance between the laws of $\mathbf{Y}$ and $\mathbf{Z}$ by bounding $|\mathbf{E}\mathcal{A}h(\mathbf{Y})|$ over all such functions h.

We write $\mathcal{A}h(\mathbf{Y})$ as

$$\begin{aligned}\mathcal{A}h(\mathbf{Y}) = \sum_{\alpha\in\mathcal{I}} \mathbf{E}[Y_\alpha]\big(h(\mathbf{Y}+e_\alpha)-h(\mathbf{Y})\big) + \sum_{\alpha\in\mathcal{I}} Y_\alpha\big(h(\mathbf{Y}-e_\alpha)-h(x)\big)\\ + \sum_{\alpha\in\mathcal{I}}\big(\mathbf{E}Z_\alpha-\mathbf{E}Y_\alpha\big)\big(h(\mathbf{Y}+e_\alpha)-h(\mathbf{Y})\big).\end{aligned}$$

The first two of these sums have expectation zero, so

$$|\mathbf{E}\mathcal{A}h(\mathbf{Y})| \le \sum_{\alpha\in\mathcal{I}} |\mathbf{E}Z_\alpha - \mathbf{E}Y_\alpha|\mathbf{E}|h(\mathbf{Y}+e_\alpha)-h(\mathbf{Y})|.$$

By (2.6), $|h(\mathbf{Y}+e_\alpha)-h(\mathbf{Y})|\le 1$, which proves the lemma. □

Lemma 2.12. *Let a and b be d-dimensional vectors with nonnegative integer components, and let $\langle a,b\rangle$ denote the standard Euclidean inner product.*

$$\prod_{i=1}^d \frac{1}{[n]_{a_i+b_i}} - \prod_{i=1}^d \frac{1}{[n]_{a_i}[n]_{b_i}} \le \frac{\langle a,b\rangle}{n}\prod_{i=1}^d\frac{1}{[n]_{a_i+b_i}}$$

Proof. We define a family of independent random maps σ_i and τ_i for $1\le i\le d$. Choose σ_i uniformly from all injective maps from $[a_i]$ to $[n]$, and choose τ_i uniformly from all injective maps from $[b_i]$ to $[n]$. Effectively, σ_i and τ_i are random ordered subsets of $[n]$. We say that σ_i and τ_i *clash* if their images overlap.

$$\mathbf{P}[\sigma_i \text{ and } \tau_i \text{ clash for some } i] = 1 - \prod_{i=1}^{d} \frac{[n]_{a_i+b_i}}{[n]_{a_i}[n]_{b_i}}.$$

For any $1 \leq i \leq d$, $1 \leq j \leq a_i$, and $1 \leq k \leq b_i$, the probability that $\sigma_i(j) = \tau_i(k)$ is $1/n$. By a union bound,

$$\mathbf{P}[\sigma_i \text{ and } \tau_i \text{ clash for some } i] \leq \sum_{i=1}^{d} \frac{a_i b_i}{n} = \frac{\langle a, b\rangle}{n}.$$

We finish the proof by dividing both sides of this inequality by $\prod_{i=1}^{d}[n]_{a_i+b_i}$. □

Proof of Theorem 2.2. We will give the proof in three sections: First, we make the coupling and show that it satisfies (2.2) and (2.3). Next, we apply Proposition 2.6 to approximate $\mathbf{I}$ by $\mathbf{Y}$, a vector of independent Poissons with $\mathbf{E}Y_\alpha = \mathbf{E}I_\alpha$. Last, we approximate $\mathbf{Y}$ by $\mathbf{Z}$ to prove the theorem.

If $d > n^{1/2}$ or $r > n^{1/10}$, then $c(2d-1)^{2r-1}/n > 1$ for a sufficiently large choice of c, and the theorem holds trivially. Thus we will assume throughout that $d \leq n^{1/2}$ and $r \leq n^{1/10}$ (the choice of $1/10$ here is completely arbitrary). The expression $O(f(d,r,n))$ should be interpreted as a function of d, r, and n whose absolute value is bounded by $Cf(d,r,n)$ for some absolute constant C, for all d, r, and n satisfying $2 \leq d \leq n^{1/2}$ and $r \leq n^{1/10}$.

Step 1. *Constructing the coupling.*

Fix some $\alpha \in \mathcal{I}$. We will construct a random vector $(J_{\beta\alpha},\, \beta \in \mathcal{I})$ distributed as $(I_\beta,\, \beta \in \mathcal{I})$ conditioned on $I_\alpha = 1$. We do this by constructing a random graph G'_n distributed as G_n conditioned to contain the cycle α. Once this is done, we will define $J_{\beta\alpha} = 1\{G'_n \text{ contains cycle } \beta\}$.

Let $\pi_1, \ldots, \pi_d$ be the random permutations that give rise to G_n. We will alter them to form permutations $\pi'_1, \ldots, \pi'_d$, and we will construct G'_n from these. Let us first consider what distributions $\pi'_1, \ldots, \pi'_d$ should have. For example, suppose that α is the cycle

$$1 \xrightarrow{\pi_3} 2 \xleftarrow{\pi_1} 3 \xrightarrow{\pi_3} 4 \xrightarrow{\pi_1} 1.$$

Then π'_1 should be distributed as a uniform random n-permutation conditioned to make $\pi'_1(3) = 2$ and $\pi'_1(4) = 1$, and π'_3 should be distributed as a uniform random n-permutation conditioned to make $\pi'_3(1) = 2$ and $\pi'_3(3) = 4$, while π'_2 should just be a uniform random n-permutation. A random graph constructed from π'_1, π'_2, and π'_3 will be distributed as G_n conditioned to contain α.

We now describe the construction of $\pi'_1, \ldots, \pi'_d$. Suppose α is the cycle

$$s_0 \xrightarrow{\;w_1\;} s_1 \xrightarrow{\;w_2\;} s_2 \xrightarrow{\;w_3\;} \cdots \xrightarrow{\;w_k\;} s_k = s_0, \tag{2.7}$$

with each edge directed according to whether $w_i(s_{i-1}) = s_i$ or $w_i(s_i) = s_{i-1}$. Fix some $1 \leq l \leq d$, and suppose that the edge-label π_l appears M times in the cycle α. Let (a_m, b_m) for $1 \leq m \leq M$ be these directed edges. We must construct π'_l to have the uniform distribution conditioned on $\pi'_l(a_m) = b_m$ for $1 \leq m \leq M$.

We define a sequence of random transpositions by the following algorithm: Let τ_1 swap $\pi_l(a_1)$ with b_1. Let τ_2 swap $\tau_1\pi_l(a_2)$ with b_2, and so on. We then define $\pi'_l = \tau_M \cdots \tau_1\pi_l$. This permutation satisfies $\pi'_l(a_m) = b_m$ for $1 \leq m \leq M$, and it is distributed uniformly, subject to the given constraints, which can be proven by induction on each swap. We now define G'_n from the permutations $\pi'_1, \ldots, \pi'_d$ in the usual way. It is defined on the same probability space as G_n, and it is distributed as G_n conditioned to contain α, giving us a random vector $(J_{\beta\alpha},\, \beta \in \mathcal{I})$ coupled with $(I_\beta,\, \beta \in \mathcal{I})$.

Now, we will give a partition $\mathcal{I}^- \cup \mathcal{I}^+ = \mathcal{I} \setminus \{\alpha\}$ satisfying (2.2) and (2.3). Suppose that G_n contains an edge $s_i \xrightarrow{w_{i+1}} v$ with $v \neq s_{i+1}$, or an edge $v \xrightarrow{w_{i+1}} s_{i+1}$ with $v \neq s_i$. The graph G'_n cannot contain this edge, since it contains α. In fact, edges of this form are the *only* ones found in G_n but not G'_n:

Lemma 2.13. *Suppose there is an edge $i \xrightarrow{\pi_l} j$ contained in G_n but not in G'_n. Then α contains either an edge $i \xrightarrow{\pi_l} v$ with $v \neq j$, or α contains an edge $v \xrightarrow{\pi_l} j$ with $v \neq i$.*

Proof. Suppose $\pi_l(i) = j$, but $\pi'_l(i) \neq j$. Then j must have been swapped when making π'_l, which can happen only if $\pi_l(a_m) = j$ or $b_m = j$ for some m. In the first case, $a_m = i$

and α contains the edge $i \xrightarrow{\pi_l} b_m$ with $b_m \neq j$, and in the second α contains the edge $a_m \xrightarrow{\pi_l} j$ with $a_m \neq i$. □

Define $\mathcal{I}_\alpha^-$ as all cycles in $\mathcal{I}$ that contain an edge $s_i \xrightarrow{w_{i+1}} v$ with $v \neq s_{i+1}$ or an edge $v \xrightarrow{w_{i+1}} s_{i+1}$ with $v \neq s_i$, and define $\mathcal{I}_\alpha^+$ to be the rest of $\mathcal{I} \setminus \{\alpha\}$. Since G_n' cannot contain any cycle in $\mathcal{I}_\alpha^-$, we have $J_{\beta\alpha} = 0$ for all $\beta \in \mathcal{I}_\alpha^-$, satisfying (2.2). For any $\beta \in \mathcal{I}_\alpha^+$, Lemma 2.13 shows that if β appears in G_n, it must also appear in G_n'. Hence $J_{\beta\alpha} \geq I_\beta$, and (2.3) is satisfied.

Step 2. *Approximation of* $\mathbf{I}$ *by* $\mathbf{Y}$.

The conditions of Proposition 2.6 are satisfied, and we need only bound the sums in (2.4). Let $p_\alpha = \mathbf{E} I_\alpha$, the probability that cycle α appears in G_n. Recall that this equals $\prod_{i=1}^d 1/[n]_{e_i}$, where e_i is the number of times π_i and π_i^{-1} appear in the word of α. This means that

$$\frac{1}{n^k} \leq p_\alpha \leq \frac{1}{[n]_k}, \tag{2.8}$$

where $k = |\alpha|$, the length of cycle α.

We bound the first sum in (2.4) by

$$\begin{aligned} \sum_{\alpha \in \mathcal{I}} p_\alpha^2 = \sum_{k=1}^r \sum_{\alpha \in \mathcal{I}_k} p_\alpha^2 &\leq \sum_{k=1}^r \sum_{\alpha \in \mathcal{I}_k} \frac{1}{[n]_k^2} \\ &= \sum_{k=1}^r \left(\frac{[n]_k a(d,k)}{2k} \right) \left(\frac{1}{[n]_k^2} \right) \\ &\leq \sum_{k=1}^r \frac{2d(2d-1)^{k-1}}{2k[n]_k} = O\left(\frac{d}{n}\right). \end{aligned} \tag{2.9}$$

To bound the second sum in (2.4), we investigate the size of $\mathcal{I}_\alpha^-$. Suppose that $\alpha \in \mathcal{I}_k$, and α has the form given in (2.7). Any $\beta \in \mathcal{I}_\alpha^-$ must contain an edge $s_i \xrightarrow{w_{i+1}} v$ with $v \neq s_{i+1}$, or an edge $v \xrightarrow{w_{i+1}} s_{i+1}$ with $v \neq s_i$, and there are at most $2k(n-1)$ edges of this form. For any given edge, there are at most $[n-2]_{j-2}(2d-1)^{j-1}$ cycles in $\mathcal{I}_j$ that contain that edge, for any $j \geq 2$. Thus for any $\alpha \in \mathcal{I}_k$, the number of cycles of length $j \geq 2$ in $\mathcal{I}_\alpha^-$ is at most $2k[n-1]_{j-1}(2d-1)^{j-1}$, and this bound also holds for $j = 1$.

For any $\beta \in \mathcal{I}_\alpha^-$, it holds that $\mathbf{E}[I_\alpha I_\beta] = 0$, so that $\mathbf{Cov}(I_\alpha, I_\beta) = -p_\alpha p_\beta$. Putting this all together and applying (2.8), we have

$$\begin{aligned}\sum_{\alpha\in\mathcal{I}}\sum_{\beta\in\mathcal{I}_\alpha^-}|\mathbf{Cov}(I_\alpha,I_\beta)| &= \sum_{k=1}^{r}\sum_{\alpha\in\mathcal{I}_k}\sum_{j=1}^{r}\sum_{\beta\in\mathcal{I}_\alpha^-\cap\mathcal{I}_j} p_\alpha p_\beta \\ &\leq \sum_{k=1}^{r}|\mathcal{I}_k|\frac{1}{[n]_k}\sum_{j=1}^{r}|\mathcal{I}_\alpha^-\cap\mathcal{I}_j|\frac{1}{[n]_j} \\ &\leq \sum_{k=1}^{r}\frac{a(d,k)}{2k}\sum_{j=1}^{r}\frac{2k(2d-1)^{j-1}}{n} \\ &= \sum_{k=1}^{r}a(d,k)O\left(\frac{(2d-1)^{r-1}}{n}\right) = O\left(\frac{(2d-1)^{2r-1}}{n}\right). \end{aligned} \tag{2.10}$$

The final sum in (2.4) is the most difficult to bound. We partition $\mathcal{I}_\alpha^+$ into sets $\mathcal{I}_\alpha^+ = \mathcal{I}_\alpha^0 \cup \cdots \cup \mathcal{I}_\alpha^{|\alpha|-1}$, where $\mathcal{I}_\alpha^l$ is all cycles in $\mathcal{I}_\alpha^+$ that share exactly l labeled edges with α. For any $\beta \in \mathcal{I}_\alpha^+$,

$$\mathbf{E}[I_\alpha I_\beta] = \mathbf{P}[G \text{ contains } \alpha \text{ and } \beta] = \prod_{i=1}^{d}\frac{1}{[n]_{e_i}},$$

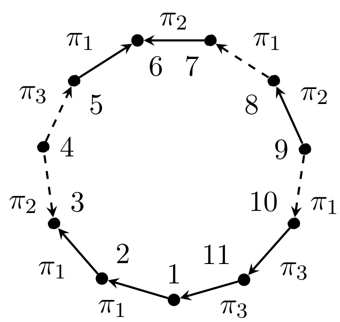
where e_i is the number of π_i-labeled edges in $\alpha \cup \beta$. Thus for $\beta \in \mathcal{I}_\alpha^l$,

$$\frac{1}{n^{|\alpha|+|\beta|-l}} \leq \mathbf{E}[I_\alpha I_\beta] \leq \frac{1}{[n]_{|\alpha|+|\beta|-l}}. \tag{2.11}$$

We start by seeking estimates on the size of $\mathcal{I}_\alpha^l$ for $l \geq 1$. Fix some choice of l edges of α. We start by counting the cycles in $\mathcal{I}_\alpha^l$ that share exactly these edges with α. We illustrate this in Figure 2.1. Call the graph consisting of these edges H, and suppose that H has p components. Since it is a forest, H has $l + p$ vertices.

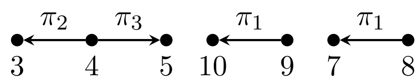
Let $A_1, \ldots, A_p$ be the components of H. We can assemble any element $\beta \in \mathcal{I}_\alpha^l$ that overlaps with α in H by stringing together these components in some order, with other edges in between. Each component can appear in β in one of two orientations. Since the vertices in β have no fixed ordering, we can assume without loss of generality that β begins with component A_1 with a fixed orientation. This leaves $(p-1)!2^{p-1}$ choices for the order and orientation of $A_2, \ldots, A_p$ in β.

Imagine now the components laid out in a line, with gaps between them, and count the number of ways to fill the gaps. Suppose that β is to have length j. Each of the p gaps

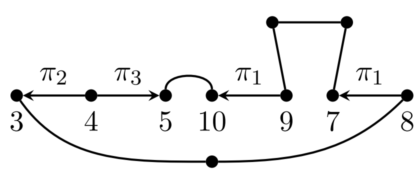


The cycle α, with H dashed. The subgraph H has components $A_1, \ldots, A_p$. In this example, the number of components of H is $p = 3$, the size of α is $k = 11$, and the number of edges in H is $l = 4$.

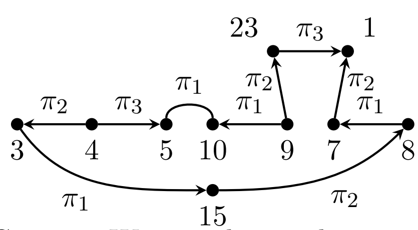
In this example, we will construct a cycle β of length $j = 10$ that overlaps with α at H.



Step 1. We lay out the components $A_1, \ldots, A_p$. We can order and orient $A_2, \ldots, A_p$ however we would like, for a total of $(p-1)!2^{p-1}$ choices. Here, we have ordered the components A_1, A_3, A_2, and we have reversed the orientation of A_3.



Step 2. Next, we choose how many edges will go in each gap between components. Each gap must contain at least one edge, and we must add a total of $j - l$ edges, giving us $\binom{j-l-1}{p-1}$ choices. In this example, we have added one edge after A_1, three after A_3, and two after A_2.



Step 3. We can choose the new vertices in $[n-p-l]_{j-p-l}$ ways, and we can direct and give labels to the new edges in at most $(2d-1)^{j-l}$ ways.

Figure 2.1: Assembling an element $\beta \in \mathcal{I}^l_\alpha$ that overlaps with α at a given subgraph H.

must contain at least one edge, and the total number of edges in all the gaps is $j-l$. Thus the total number of possible gap sizes is the number of compositions of $j-l$ into p parts, or $\binom{j-l-1}{p-1}$.

Now that we have chosen the number of edges to appear in each gap, we choose the edges themselves. We can do this by giving an ordered list $j-p-l$ vertices to go in the gaps, along with a label and an orientation for each of the $j-l$ edges this gives. There are $[n-p-l]_{j-p-l}$ ways to choose the vertices. We can give each new edge any orientation and label subject to the constraint that the word of the cycle we construct must be reduced. This means we have at most $2d-1$ choices for the orientation and label of each new edge, for a total of at most $(2d-1)^{j-i}$.

All together, there are at most $(p-1)!2^{p-1}\binom{j-l-1}{p-1}[n-p-l]_{j-p-l}(2d-1)^{j-l}$ elements of $\mathcal{I}_j$ that overlap with the cycle α at the subgraph H. We now calculate the number of different ways to choose a subgraph H of α with l edges and p components. Suppose α is given as in (2.7). We first choose a vertex s_{i_0}. Then, we can specify which edges to include in H by giving a sequence $a_1, b_1, \ldots, a_p, b_p$ instructing us to include in H the first a_1 edges after s_{i_0}, then to exclude the next b_1, then to include the next a_2, and so on. Any sequence for which a_i and b_i are positive integers, $a_1 + \cdots + a_p = l$, and $b_1 + \cdots + b_p = k - l$ gives us a valid choice of l edges of α making up p components. This counts each subgraph H a total of p times, since we could begin with any component of H. Hence the number of subgraphs H with l edges and p components is $(k/p)\binom{l-1}{p-1}\binom{k-l-1}{p-1}$. This gives us the bound

$$|\mathcal{I}_\alpha^l \cap \mathcal{I}_j| \leq \sum_{p=1}^{l\wedge(j-l)} (k/p)\binom{l-1}{p-1}\binom{k-l-1}{p-1}(p-1)! \times 2^{p-1}\binom{j-l-1}{p-1}[n-p-l]_{j-p-l}(2d-1)^{j-l}.$$

We apply the bounds

$$\binom{l-1}{p-1} \leq \frac{r^{p-1}}{(p-1)!},$$

$$\binom{k-l-1}{p-1}, \binom{j-l-1}{p-1} \leq (er/(p-1))^{p-1},$$

to get

$$|\mathcal{I}_\alpha^l \cap \mathcal{I}_j| \leq k(2d-1)^{j-l}[n-1-l]_{j-1-l}\left(1+\sum_{p=2}^{i\wedge(k-i)}\frac{1}{p}\left(\frac{2e^2r^3}{(p-1)^2}\right)^{p-1}\frac{1}{[n-1-l]_{p-1}}\right).$$

Since $r \leq n^{1/10}$, the sum in the above equation is bounded by an absolute constant. Applying this bound and (2.11), for any $\alpha \in \mathcal{I}_k$ and $l \geq 1$,

$$\begin{aligned}\sum_{\beta\in\mathcal{I}_\alpha^l}\mathbf{Cov}(I_\alpha, I_\beta) &\leq \sum_{j=l+1}^{r}\sum_{\beta\in\mathcal{I}_\alpha^l\cap\mathcal{I}_j}\frac{1}{[n]_{k+j-l}} \\ &\leq \sum_{j=l+1}^{r} O\left(\frac{k(2d-1)^{j-l}}{n^{k+1}}\right) \\ &= O\left(\frac{k(2d-1)^{r-l}}{n^{k+1}}\right).\end{aligned} \tag{2.12}$$

Therefore

$$\begin{aligned}\sum_{\alpha\in\mathcal{I}}\sum_{l\geq 1}\sum_{\beta\in\mathcal{I}_\alpha^l}\mathbf{Cov}(I_\alpha, I_\beta) &= \sum_{k=1}^{r}\sum_{\alpha\in\mathcal{I}_k}\sum_{l=1}^{k-1}\sum_{\beta\in\mathcal{I}_\alpha^l}\mathbf{Cov}(I_\alpha, I_\beta) \\ &\leq \sum_{k=1}^{r}\sum_{\alpha\in\mathcal{I}_k}\sum_{l=1}^{k-1} O\left(\frac{k(2d-1)^{r-l}}{n^{k+1}}\right) \\ &= \sum_{k=1}^{r}\frac{[n]_k a(d,k)}{2k} O\left(\frac{k(2d-1)^{r-1}}{n^{k+1}}\right) \\ &= \sum_{k=1}^{r} O\left(\frac{(2d-1)^{r+k-1}}{n}\right) \\ &= O\left(\frac{(2d-1)^{2r-1}}{n}\right).\end{aligned} \tag{2.13}$$

Last, we must bound $\sum_{\alpha\in\mathcal{I}}\sum_{\beta\in\mathcal{I}_\alpha^0}\mathbf{Cov}(I_\alpha, I_\beta)$. For any word w, let e_i^w be the number of appearances of π_i and π_i^{-1} in w. Let α and β be cycles with words w and u respectively, and let $k = |\alpha|$ and $j = |\beta|$. Suppose that $\beta \in \mathcal{I}_\alpha^0$. Then

$$\begin{aligned}\mathbf{Cov}(I_\alpha, I_\beta) &= \prod_{i=1}^{d}\frac{1}{[n]_{e_i^w+e_i^u}} - \prod_{i=1}^{d}\frac{1}{[n]_{e_i^w}[n]_{e_i^u}} \\ &\leq \frac{\langle e^w, e^u\rangle}{n}\prod_{i=1}^{d}\frac{1}{[n]_{e_i^w+e_i^u}} \leq \frac{\langle e^w, e^u\rangle}{n[n]_{k+j}}\end{aligned}$$

by Lemma 2.12. For any pair of words $w \in \mathcal{W}_k$ and $u \in \mathcal{W}_j$, there are at most $[n]_k[n]_j$ pairs of cycles $\alpha, \beta \in \mathcal{I}$ with words w and u, respectively. Enumerating over all $w \in \mathcal{W}_k$ and $u \in \mathcal{W}_j$, we count each pair of cycles α, β exactly $4kj$ times. Thus

$$\sum_{\alpha \in \mathcal{I}_k} \sum_{\beta \in \mathcal{I}^0_\alpha \cap \mathcal{I}_j} \mathbf{Cov}(I_\alpha, I_\beta) \le \frac{[n]_k[n]_j}{4kjn[n]_{k+j}} \sum_{w \in \mathcal{W}_k} \sum_{u \in \mathcal{W}_j} \langle e^w, e^u \rangle$$
$$\le \frac{1 + O(r^2/n)}{4kjn} \left\langle \sum_{w \in \mathcal{W}_k} e^w, \sum_{u \in \mathcal{W}_j} e^u \right\rangle.$$

The vector $\sum_{w \in \mathcal{W}_k} e^w$ has every entry equal by symmetry, as does $\sum_{u \in \mathcal{W}_j} e^u$. Thus each entry of $\sum_{w \in \mathcal{W}_k} e^w$ is $ka(d,k)/d$, and each entry of $\sum_{u \in \mathcal{W}_j} e^u$ is $ja(d,j)/d$. The inner product in the above equation comes to $kja(d,k)a(d,j)/d$, giving us

$$\sum_{\alpha \in \mathcal{I}_k} \sum_{\beta \in \mathcal{I}^0_\alpha \cap \mathcal{I}_j} \mathbf{Cov}(I_\alpha, I_\beta) \le \frac{a(d,k)a(d,j)(1 + O(r^2/n))}{4dn}$$
$$= O\left(\frac{(2d-1)^{j+k-1}}{n} \right). \tag{2.14}$$

Summing over all $1 \le k, j \le r$,

$$\sum_{\alpha \in \mathcal{I}} \sum_{\beta \in \mathcal{I}^0_\alpha} \mathbf{Cov}(I_\alpha, I_\beta) = \left(\frac{(2d-1)^{2r-1}}{n} \right). \tag{2.15}$$

We can now combine equations (2.9), (2.10), (2.13), and (2.15) with Proposition 2.6 to show that

$$d_{TV}(\mathbf{I}, \mathbf{Y}) = O\left(\frac{(2d-1)^{2r-1}}{n} \right). \tag{2.16}$$

Step 3. *Approximation of* $\mathbf{Y}$ *by* $\mathbf{Z}$.

By Lemma 2.11 and (2.8),

$$d_{TV}(\mathbf{Y}, \mathbf{Z}) \le \sum_{\alpha \in \mathcal{I}} |\mathbf{E}Y_\alpha - \mathbf{E}Z_\alpha| \le \sum_{k=1}^{r} \sum_{\alpha \in \mathcal{I}_k} \left(\frac{1}{[n]_k} - \frac{1}{n^k} \right)$$
$$= \sum_{k=1}^{r} \frac{a(d,k)}{2k} \left(1 - \frac{[n]_k}{n^k} \right).$$

Since $[n]_k \ge n^k(1 - k^2/2n)$,

$$d_{TV}(\mathbf{Y}, \mathbf{Z}) \le \sum_{k=1}^{r} \frac{a(d,k)k}{4n} = O\left(\frac{r(2d-1)^r}{n} \right). \tag{2.17}$$

Together with (2.16), this bounds the total variation distance between the laws of **I** and **Z** and proves the theorem. □

Proof of. Consider the partition $\mathcal{I} = \bigcup_{k=1}^{r} \mathcal{I}_k$, and define W_k and Y_k as in the statement of Proposition 2.7. As in the proof of Theorem 2.2, we may assume that $d \leq n^{1/2}$ and $n \leq n^{1/10}$. With these restrictions, we have

$$\log^+ \max \lambda_j = O\big(r \log(2d-1)\big),$$
$$\lambda_k^{-1} = O\bigg(\frac{k}{(2d-1)^k}\bigg),$$
$$(\lambda_j \lambda_k)^{-1/2} = O\bigg(\frac{\sqrt{jk}}{(2d-1)^{(j+k)/2}}\bigg).$$

We have already bounded all the terms in (2.5) in the previous proof. From (2.9),

$$\sum_{k=1}^{r} \sum_{\alpha \in \mathcal{I}_k} \frac{p_\alpha^2}{\lambda_k} = O\Big(\frac{d}{n}\Big).$$

From (2.10),

$$\sum_{\alpha \in \mathcal{I}_k} \sum_{\beta \in \mathcal{I}_\alpha^- \cap \mathcal{I}_j} |\mathbf{Cov}(I_\alpha, I_\beta)| = O\bigg(\frac{(2d-1)^{j+k-1}}{n}\bigg). \tag{2.18}$$

Recalling the partition of $\mathcal{I}_\alpha^+$ on p. 19, and following (2.12), for any $\alpha \in \mathcal{I}_k$ and $l \geq 1$,

$$\sum_{\beta \in \mathcal{I}_\alpha^l \cap \mathcal{I}_j} \mathbf{Cov}(I_\alpha, I_\beta) = O\bigg(\frac{k(2d-1)^{j-l}}{n^{k+1}}\bigg),$$

and

$$\sum_{\alpha \in \mathcal{I}_k} \sum_{l \geq 1} \sum_{\beta \in \mathcal{I}_\alpha^l \cap \mathcal{I}_j} \mathbf{Cov}(I_\alpha, I_\beta) = \sum_{\alpha \in \mathcal{I}_k} O\bigg(\frac{k(2d-1)^{j-1}}{n^{k+1}}\bigg) = O\bigg(\frac{(2d-1)^{j+k-1}}{n}\bigg).$$

Together with (2.14), this shows that

$$\sum_{\alpha \in \mathcal{I}_k} \sum_{\beta \in \mathcal{I}_\alpha^+ \cap \mathcal{I}_j} \mathbf{Cov}(I_\alpha, I_\beta) = O\bigg(\frac{(2d-1)^{j+k-1}}{n}\bigg).$$

This and (2.18) prove that

$$A(j,k) = O\bigg(\frac{(2d-1)^{j+k-1}}{n}\bigg).$$

Now, we apply Proposition 2.7:

$$d_{TV}\big((W_1,\dots,W_r),\ (Y_1,\dots,Y_r)\big) = O\bigg(\frac{r^2(2d-1)^{r-1}\log(2d-1)}{n}\bigg).$$

Last, we apply (2.17) to bound the distance between $(Y_1,\dots,Y_r)$ and $(Z_1,\dots,Z_r)$ and complete the proof. □

2.3 Poisson approximation in the uniform model

2.3.1 Preliminaries

For vertices u and v in a graph, we will use the notation $u \sim v$ to denote that the edge uv exists. The distance between two vertices is the length of the shortest path between them, and the distance between two edges or sets of vertices is the shortest distance between a vertex in one set and a vertex in the other.

Here and throughout, we will use $c_1, c_2, \dots$ to denote absolute constants whose values are unimportant to us.

Proposition 2.14. *Let G be a random d-regular graph on n vertices, with $d \le n^{1/3}$.*

(a) Let α be a cycle of length $k \le n^{1/10}$ in the complete graph K_n. Then

$$\mathbf{P}[\alpha \subseteq G] \le \frac{c_1(d-1)^k}{n^k}.$$

(b) Let β be another cycle in K_n of length $j \le n^{1/10}$, and suppose that α and β share f edges. Then

$$\mathbf{P}[\alpha \cup \beta \subseteq G] \le \frac{c_2(d-1)^{j+k-f}}{n^{j+k-f}}.$$

(c) Let H be a subgraph of K_n consisting of a j-cycle and a k-cycle joined by path of length l, as in Figure 2.2. Suppose that $j,k,l \le n^{1/10}$. Then

$$\mathbf{P}[H \subseteq G] \le \frac{c_3(d-1)^{j+k+l}}{n^{j+k+l}}.$$

Proof. These statements all follow directly from Theorem 3a in [MWW04]. □

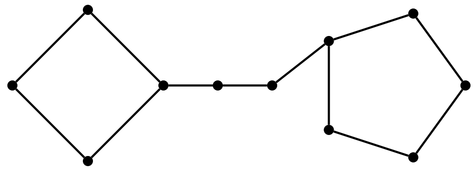


Figure 2.2: A 4-cycle and a 5-cycle, connected by a path of length 3.

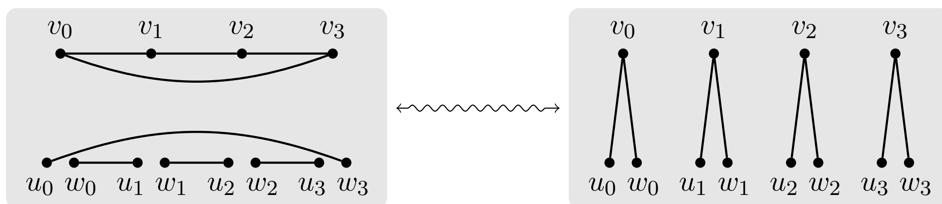


Figure 2.3: The change from left to right is a *forward switching*, and from right to left is a *backward switching*.

2.3.2 Counting switchings

We will follow [MWW04], defining and counting switchings. After this, we will break with that paper by using the switchings to apply Stein's method. Besides some small notational differences, the definitions will be the same as those in [MWW04]. To avoid repetition of the phrase "cycles of length r or less," we will refer to such cycles as *short*.

Let G be a d-regular graph. Suppose that $\alpha = v_0 \cdots v_{k-1}$ is a cycle in G, and let $e_i = v_i v_{i+1}$, interpreting all indices modulo k from now on. Let $e_i' = w_i u_{i+1}$ for $0 \leq i \leq k-1$ be oriented edges such that neither u_i nor w_i is adjacent to v_i. Consider the act of deleting these $2k$ edges and replacing them with the edges $v_i u_i$ and $v_i w_i$ for $0 \leq i \leq k-1$ to obtain a new d-regular graph G' with the cycle α deleted (see Figure 2.3). We call this action induced given by the sequences (v_i), (u_i), and (w_i) a *forward α-switching*. We will consider forward α-switchings only up to cyclic rotation of indices; that is, we identify the $2k$ different α-switchings obtained by cyclically rotating all sequences v_i, u_i, and w_i.

To go the opposite direction, suppose G contains oriented paths $u_i v_i w_i$ for $0 \leq i \leq k-1$

such that $v_i \not\sim v_{i+1}$ and $w_i \not\sim u_{i+1}$. Consider the act of deleting all edges $u_i v_i$ and $v_i w_i$ and replacing them with $v_i v_{i+1}$ and $w_i u_{i+1}$ for all $0 \leq i \leq k-1$ to create a new graph G' that contains the cycle $\alpha = v_0 \cdots v_{k-1}$. We call this a *backwards α-switching.* Again, we consider switchings only up to cyclic rotation of all indices.

We call an α-switching *valid* if α is the only short cycle created or destroyed by the switching. For each valid forward α-switching taking G to G', there is a corresponding valid backwards α-switching taking G' to G. Let F_α and B_α be the number of valid forward and backwards α-switchings, respectively, on some graph G. Using arguments drawn from [MWW04, Lemma 3], we give some estimates on them.

Lemma 2.15. *Let G be a deterministic d-regular graph on n vertices with cycle counts $\{C_k,\, k \geq 3\}$. For any short cycle $\alpha \subseteq G$ of length k,*

$$F_\alpha \leq [n]_k d^k. \tag{2.19}$$

If α does not share an edge with another short cycle,

$$F_\alpha \geq [n]_k d^k \left(1 - \frac{2k\sum_{j=3}^{r} jC_j + c_4 k(d-1)^r}{nd}\right). \tag{2.20}$$

Proof. The question is, with $\alpha = v_0 \cdots v_{k-1}$ and $e_i = v_i v_{i+1}$ given, how many ways are there to choose $e'_0, \ldots, e'_{k-1}$ that give a valid switching? There are at most $[n]_k d^k$ choices of oriented edges $e'_0, \ldots, e'_{k-1}$, which proves the upper bound (2.19). For the lower bound, we demonstrate a procedure to choose these edges that is guaranteed to give us a valid forward α-switching. Suppose that $e'_0, \ldots, e'_{k-1}$ satisfy

(a) e'_i is not contained in any short cycle;

(b) the distance from e_i to e'_i is at least r;

(c) the distance from e'_i to $e'_{i'}$ is at least $r/2$;

(d) the distance from w_i to u_i is at least r.

Then the switching is valid by an argument identical to the one in [MWW04], which we will reproduce for convenience. By (b), for all i, neither u_i nor w_i is adjacent to v_i (or to $v_{i'}$ for any i'), as required in the definition of a switching. Let G' be the graph obtained by applying the switching. We need to check now that the switching is valid; that is, the only short cycle it creates or destroys is α.

Since α shares no edges with other short cycles, its deletion does not destroy any other short cycles. Condition (a) ensures that no short cycles are destroyed by removing $e'_0, \ldots, e'_{k-1}$. The switching does not create any short cycles either: Suppose otherwise, and let β be the new cycle in G'. It consists of paths in $G \cap G'$, separated by new edges in G'. Any such path in $G \cap G'$ must have length at least $r/2$, because

- if it starts and ends in α and has length less than $r/2$, then combining this path with a path in α gives an short cycle in G that overlaps with α;

- if it starts in α and finishes in $W = \{u_0, w_0, \ldots, u_{k-1}, w_{k-1}\}$ and has length less than $r/2$, then combining this path with a path in α gives a path violating condition (b);

- if it starts at some e'_i and ends at $e'_{i'}$ then it must have length $r/2$ by (c) if $i' \neq i$, and by (a) if $i' = i$.

Thus β contains exactly one path in $G \cap G'$. The remainder of β must be an edge $u_i v_i$ or $w_i v_i$, impossible by (b), or a path $u_i v_i w_i$, impossible by (d).

Now, we find the number of switchings that satisfy conditions (a)–(d) to get a lower bound on F_α. We will do this by bounding from above the number of switchings out of the $[n]_k d^k$ counted in (2.19) that fail each condition (a)–(d).

- There are a total of $\sum_{j=3}^{r} jC_j$ edges in short cycles in G. Choosing one of the edges $e'_0, \ldots, e'_{k-1}$ from these and the rest arbitrarily, there are at most $[n-1]_{k-1} d^{k-1} k \sum_{j=3}^{r} 2jC_j$ switchings that fail condition (a).

- The number of edges of distance less than r from some edge is at most $2\sum_{j=0}^{r}(d-1)^j - 1 = O((d-1)^r)$. At most $[n-1]_{k-1} d^{k-1} k O\big((d-1)^r\big)$ switchings then fail condition (b).

- By a similar argument, at most $[n]_{k-1}d^{k-1}k^2O\big((d-1)^{r/2}\big)$ switchings fail condition (c).

- By a similar argument, at most $[n]_{k-1}d^{k-1}kO\big((d-1)^r\big)$ switchings fail condition (d).

Adding these up and combining $O(\cdot)$ terms, we find that at most

$$[n-1]_{k-1}d^{k-1}k\left(\sum_{j=3}^{r}2jC_j+O\big((d-1)^r\big)\right)$$

switchings out of the original $[n]_k d^k$ fail conditions by (a)–(d), establishing (2.20). □

For backwards switchings, we give a similar upper bound, but we only give our lower bound in expectation.

Lemma 2.16. *Let G be a random d-regular graph on n vertices, and let α be a cycle of length $k \le r$ in the complete graph K_n. Then*

$$B_\alpha \le \big(d(d-1)\big)^k \tag{2.21}$$

and

$$\mathbf{E}B_\alpha \ge \big(d(d-1)\big)^k\left(1-\frac{c_5k(d-1)^{r-1}}{n}\right). \tag{2.22}$$

Proof. The question this time is given α, how many choices of oriented paths yield a valid switching? For any fixed α, there are at most $(d(d-1))^k$ choices of oriented paths, proving (2.21). For the lower bound, let $B=\sum_\beta B_\beta$, where β runs over all cycles of length k in the complete graph. We will first show that

$$B \ge \frac{[n]_k\big(d(d-1)\big)^k}{2k}\left(1-\frac{4k\sum_{j=3}^{r}jC_j+O\big(k(d-1)^r\big)}{nd}\right). \tag{2.23}$$

As in Lemma 2.15, we give conditions that ensure a valid switching. Let $\beta=v_0\cdots v_{k-1}$, and suppose that the paths $u_iv_iw_i$ in G for $0\le i\le k-1$ satisfy

(a) the edges v_iu_i and v_iw_i are not contained in any short cycles;

(b) for all $1 \le j \le r/2$, the distance between the paths $u_i v_i w_i$ and $u_{i+j} v_{i+j} w_{i+j}$ is at least $r - j + 1$.

Any choice of edges satisfying these conditions gives a valid backwards switching: Condition (b) ensures that $v_i \not\sim v_{i+1}$ and $w_i \not\sim u_{i+1}$, as required in the definition of a switching. Let G' be the graph obtained by applying the switching. We need to check that no short cycles besides β are created or destroyed by the switching. By (a), none are destroyed. Suppose a short cycle β' other than β is created in G'. It consists of paths in $G \cap G'$, portions of β, and edges $w_i u_{i+1}$. Any such path in $G \cap G'$ must have length at least $r/2$ because

- if it starts at u_i, v_i, or w_i and ends at u_{i+j}, v_{i+j}, or w_{i+j} for $1 \le j \le r/2$, then (b) implies this;
- if it starts and ends at one of u_i, v_i, and w_i, then (a) implies this.

Hence β' must contain exactly one such path. The remainder of β' must either be an edge $w_i u_{i+1}$, or a portion of β', both of which are impossible by (b).

There are $[n]_k d^k/2k$ choices for β, and at most $(d(d-1))^k$ choices for u_i, w_i, $0 \le i < k$. As before, we count how many of these potential switchings satisfy conditions (a) and (b) to get a lower bound on B. By similar arguments as in the proof of Lemma 2.15, we find that at most

$$2[n-1]_{k-1}\big(d(d-1)\big)^{k-1}(d-1)\sum_{j=3}^{r} jC_j$$

of the switchings violate condition (a), and at most $[n]_{k-1}\big(d(d-1)\big)^{k-1}O\big((d-1)^{r+1}\big)$ violate condition (b), which proves (2.23).

By Proposition 2.14a (or by [MWW04, eq. 2.2]),

$$\mathbf{E}C_k \le \frac{c_1(d-1)^k}{2k}.$$

Applying this to (2.23) gives

$$\mathbf{E}B \ge \frac{[n]_k\big(d(d-1)\big)^k}{2k}\left(1 - O\left(\frac{k(d-1)^{r-1}}{n}\right)\right)$$

By the exchangeability of the vertex labels of G, the law of B_β is the same for all k-cycles β. It follows that $\mathbf{E}B = ([n]_k/2k)\mathbf{E}B_\alpha$, proving (2.22). □

2.3.3 Applying Stein's method

We will prove a generalization of Theorem 2.3, allowing the process of cycles to be indexed by any collection of cycles, rather than just all cycles of length r or less.

Theorem 2.17. *Let G be a random d-regular graph on n vertices. For some collection $\mathfrak{I}$ of cycles in the complete graph K_n of maximum length r, we define $\mathbf{I} = (I_\alpha,\, \alpha \in \mathfrak{I})$, with $I_\alpha = \mathbf{1}\{G \text{ contains } \alpha\}$. Let $\mathbf{Z} = (Z_\alpha,\, \alpha \in \mathfrak{I})$ be a vector of independent Poisson random variables, with $\mathbf{E}Z_\alpha = (d-1)^{|\alpha|}/[n]_{|\alpha|}$, where $|\alpha|$ denotes the length of the cycle α.*

For some absolute constant c_6, for all n and $d, r \geq 3$ satisfying $r \leq n^{1/10}$ and $d \leq n^{1/3}$,

$$d_{TV}(\mathbf{I},\, \mathbf{Z}) \leq \sum_{\alpha \in \mathfrak{I}} \frac{c_6 |\alpha| (d-1)^{|\alpha|+r-1}}{n^{|\alpha|+1}}.$$

Proof. We will construct an exchangeable pair by taking a step in a reversible Markov chain. To make this chain, define a graph $\mathfrak{G}$ whose vertices consist of all d-regular graphs on n vertices. For every valid forward α-switching with $\alpha \in \mathfrak{I}$ from a graph G_0 to G_1, make an undirected edge in $\mathfrak{G}$ between G_0 and G_1. Place a weight of $1/[n]_{|\alpha|} d^{|\alpha|}$ on each of these edges. The essential fact that will make our arguments work is that valid forward α-switchings from G_0 to G_1 are in bijective correspondence with valid backwards α-switchings from G_1 to G_0. Thus, we could have equivalently defined $\mathfrak{G}$ by forming an edge for every valid backwards switching.

Define the degree of a vertex in a graph with weighted edges to be the sum of the adjacent edge weights. Let d_0 be the maximum degree of $\mathfrak{G}$ as defined so far. To make $\mathfrak{G}$ regular, add a weighted loop to each vertex that brings its degree up to d_0. Now, consider a random walk on $\mathfrak{G}$ that moves with probability proportional to the edge weights. This random walk is a Markov chain reversible with respect to the uniform distribution on d-regular graphs on n vertices. Thus, if G has this distribution, and we obtain G' by advancing one step in the random walk, the pair of graphs (G, G') is exchangeable.

Let I'_α be an indicator on G' containing the cycle α, and define $\mathbf{I}' = (I'_\alpha,\, \alpha \in \mathfrak{I})$. It follows from the exchangeability of G and G' that $\mathbf{I}$ and $\mathbf{I}'$ are exchangeable, and we can apply Proposition 2.8 on this pair. Define the events Δ^+_α and Δ^-_α as in that proposition. By our construction,

$$\mathbf{P}[\Delta_\alpha^+ \mid G] = \frac{B_\alpha}{d_0[n]_{|\alpha|}d^{|\alpha|}}, \qquad \mathbf{P}[\Delta_\alpha^- \mid G] = \frac{F_\alpha}{d_0[n]_{|\alpha|}d^{|\alpha|}}.$$

Thus by Proposition 2.8 with all constants set to d_0,

$$d_{TV}(\mathbf{I}, \mathbf{Z}) \leq \sum_{\alpha\in\mathcal{I}} \mathbf{E}\left|\frac{(d-1)^{|\alpha|}}{[n]_{|\alpha|}} - \frac{B_\alpha}{[n]_{|\alpha|}d^{|\alpha|}}\right| + \sum_{\alpha\in\mathcal{I}} \mathbf{E}\left|I_\alpha - \frac{F_\alpha}{[n]_{|\alpha|}d^{|\alpha|}}\right|. \tag{2.24}$$

We will bound these two sums. Fix some $\alpha \in \mathcal{I}$, and let $|\alpha| = k$. By Lemma 2.16,

$$\frac{B_\alpha}{[n]_k d^k} \leq \frac{(d-1)^k}{[n]_k}.$$

Thus

$$\mathbf{E}\left|\frac{(d-1)^k}{[n]_k} - \frac{B_\alpha}{[n]_k d^k}\right| = \mathbf{E}\left[\frac{(d-1)^k}{[n]_k} - \frac{B_\alpha}{[n]_k d^k}\right].$$

Applying the lower bound on $\mathbf{E}B_\alpha$ from Lemma 2.16 then gives

$$\mathbf{E}\left|\frac{(d-1)^k}{[n]_k} - \frac{B_\alpha}{[n]_k d^k}\right| \leq \frac{c_5 k (d-1)^{k+r-1}}{n[n]_k}. \tag{2.25}$$

In bounding the other sum, we partition our state space of random regular graphs into three events:

$A_1 = \{G \text{ does not contain } \alpha\}$,

$A_2 = \{G \text{ contains } \alpha, \text{ which does not share an edge with another short cycle in } G\}$,

$A_3 = \{G \text{ contains } \alpha, \text{ which shares an edge with another short cycle in } G\}$.

On A_1, we have $I_\alpha = F_\alpha = 0$. On A_2, both bounds from Lemma 2.15 apply, giving us

$$\left|I_\alpha - \frac{F_\alpha}{[n]_k d^k}\right| \leq \frac{2k\sum_{j=3}^r jC_j + c_4 k(d-1)^r}{nd}.$$

On A_3, we have $I_\alpha = 1$ and $F_\alpha = 0$. In all,

$$\begin{aligned}\mathbf{E}\left|I_\alpha - \frac{F_\alpha}{[n]_k d^k}\right| &\leq \mathbf{E}\left[\mathbf{1}_{A_2}\frac{2k\sum_{j=3}^r jC_j + c_4 k(d-1)^r}{nd} + \mathbf{1}_{A_3}\right] \\ &= \frac{2k}{nd}\mathbf{E}\left[\mathbf{1}_{A_2}\sum_{j=3}^r jC_j\right] + \frac{c_4 k(d-1)^r}{nd}\mathbf{P}[A_2] + \mathbf{P}[A_3].\end{aligned}$$

Let $\mathcal{J}$ be the set of all cycles of length r or less in K_n that share no edges with α. On the set A_2, the graph G contains no cycles outside of this set (except for α), and $\sum_{j=3}^{r} jC_j = k + \sum_{\beta\in\mathcal{J}}|\beta|I_\beta$. Thus

$$\begin{aligned}\mathbf{E}\left|I_\alpha - \frac{F_\alpha}{[n]_k d^k}\right| &\leq \frac{2k^2}{nd}\mathbf{E}\mathbf{1}_{A_2} + \frac{2k}{nd}\sum_{\beta\in\mathcal{J}}|\beta|\mathbf{E}\mathbf{1}_{A_2}I_\beta + \frac{c_4 k(d-1)^r}{nd}\mathbf{P}[A_2] + \mathbf{P}[A_3] \\ &\leq \frac{2k^2}{nd}\mathbf{E}I_\alpha + \frac{2k}{nd}\sum_{\beta\in\mathcal{J}}|\beta|\mathbf{E}I_\alpha I_\beta + \frac{c_4 k(d-1)^r}{nd}\mathbf{E}I_\alpha + \mathbf{P}[A_3]. \qquad (2.26)\end{aligned}$$

By Proposition 2.14a,

$$\frac{2k^2}{nd}\mathbf{E}I_\alpha = O\Big(\frac{k^2(d-1)^k}{n^{k+1}}\Big) \qquad (2.27)$$

and

$$\frac{c_4 k(d-1)^r}{nd}\mathbf{E}I_\alpha = O\Big(\frac{k(d-1)^{k+r-1}}{n^{k+1}}\Big). \qquad (2.28)$$

By Proposition 2.14b, for any $\beta \in \mathcal{J}$, we have $\mathbf{E}I_\alpha I_\beta \leq c_2(d-1)^{j+k}/n^{j+k}$. For each $3 \leq j \leq r$, there are at most $[n]_j/2j$ cycles in $\mathcal{J}$ of length j. Therefore

$$\begin{aligned}\frac{2k}{nd}\sum_{\beta\in\mathcal{J}}|\beta|\mathbf{E}I_\alpha I_\beta &\leq \frac{2k}{nd}\sum_{j=3}^{r}\frac{[n]_j}{2j}\Big(\frac{jc_2(d-1)^{j+k}}{n^{j+k}}\Big) \\ &\leq \frac{k}{nd}\sum_{j=3}^{r}\frac{c_2(d-1)^{j+k}}{n^k} = O\Big(\frac{k(d-1)^{k+r-1}}{n^{k+1}}\Big). \qquad (2.29)\end{aligned}$$

The last term of (2.26) is the most difficult to bound. Let $\mathcal{K}$ be the set of short cycles in K_n that share an edge with α, not including α itself. By a union bound,

$$\mathbf{P}[A_3] \leq \sum_{\beta\in\mathcal{K}}\mathbf{E}I_\alpha I_\beta. \qquad (2.30)$$

Now, we classify and count the cycles $\beta \in \mathcal{K}$ according to the structure of $\alpha\cup\beta$. Suppose that β has length j, and consider the intersection of α and β (the graph consisting of all vertices and edges contained in both α and β). Suppose this intersection graph has p components and f edges. As computed on [MWW04, p. 5], the number of possible isomorphism types of $\alpha \cup \beta$ given p and f is at most $(2r^3)^{p-1}/(p-1)!^2$. For each possible isomorphism type of $\alpha \cup \beta$, there are no more than $2kn^{j-p-f}$ possible choices of β such that $\alpha \cup \beta$ falls into this isomorphism class. This is because $\alpha \cup \beta$ has $j+k-p-f$ vertices, k of which are determined by α. In defining β, the remaining $j-p-f$ vertices can be chosen to be anything, and the intersection of α and β can be rotated around α in $2k$ ways, all without changing the isomorphism class of $\alpha \cup \beta$. In all, we have shown that the number of j-cycles whose overlap with α has p components and f edges is at most

$$\frac{(2r^3)^{p-1}}{(p-1)!^2} 2kn^{j-p-f}.$$

For any such choice of β, we have $\mathbf{E} I_\alpha I_\beta \le c_2(d-1)^{j+k-f}/n^{j+k-f}$ by Proposition 2.14b. Applying this to (2.30),

$$\begin{aligned}
\mathbf{P}[A_3] &\le \sum_{j=3}^{r} \sum_{p,f\ge 1} \frac{(2r^3)^{p-1}}{(p-1)!^2} 2kn^{j-p-f} \frac{c_2(d-1)^{j+k-f}}{n^{j+k-f}} \\
&= \sum_{j=3}^{r} \sum_{p,f\ge 1} \frac{(2r^3)^{p-1}}{(p-1)!^2} \frac{2kc_2(d-1)^{j+k-f}}{n^{k+p}} \\
&= \sum_{j=3}^{r} O\Big(\frac{k(d-1)^{j+k-1}}{n^{k+1}}\Big) = O\Big(\frac{k(d-1)^{k+r-1}}{n^{k+1}}\Big). \qquad (2.31)
\end{aligned}$$

Combining (2.27), (2.28), (2.29), and (2.31), we have

$$\mathbf{E}\left| I_\alpha - \frac{F_\alpha}{[n]_k d^k}\right| = O\Big(\frac{k(d-1)^{k+r-1}}{n^{k+1}}\Big).$$

Applying this and (2.25) to (2.24) establishes the theorem. □

Proof of Theorem 2.3. If $r > n^{1/10}$ or $d > n^{1/3}$, then $c(d-1)^{2r-1}/n > 1$ for a sufficiently large choice of c, and the total variation bound is trivial. Thus we can assume that this is not the case and apply the previous theorem:

$$\begin{aligned}
d_{TV}(\mathbf{I}, \mathbf{Z}) &\le \sum_{\alpha\in\mathcal{I}} \frac{c_6|\alpha|(d-1)^{|\alpha|+r-1}}{n^{|\alpha|+1}} \\
&= \sum_{k=3}^{r} \frac{[n]_k}{2k}\Big(\frac{c_6 k(d-1)^{k+r-1}}{n^{k+1}}\Big) \\
&= O\Big(\frac{(d-1)^{2r-1}}{n}\Big).
\end{aligned}$$

□

Since the cycle counts $(C_3, \dots, C_r)$ are a functional of $\mathbf{I}$, this corollary implies that

$$d_{TV}\big((C_3,\dots,C_r), (Z_3,\dots,Z_r)\big) \le \frac{c(d-1)^{2r-1}}{n},$$

where $(Z_3, \ldots, Z_r)$ is a vector of independent Poisson random variables with $\mathbf{E}Z_k = (d-1)^k/2k$. This bound is often less than optimal, since this theorem fails to exploit the $\lambda_k^{-1/2}$ factors in Proposition 2.8. We will take advantage of these factors in the following proposition, and then apply this to prove Theorem 2.5.

Proposition 2.18. *With the set-up of Theorem 2.17, divide up the collection of cycles* $\mathcal{I}$ *into bins* $\mathcal{B}_1, \ldots, \mathcal{B}_s$. *Let*

$$I_k = \sum_{\alpha \in \mathcal{B}_k} I_\alpha, \qquad Z_k = \sum_{\alpha \in \mathcal{B}_k} Z_\alpha,$$

and let $\lambda_k = \mathbf{E}Z_k$. *Then*

$$d_{TV}\big((I_1, \ldots, I_s),\, (Z_1, \ldots, Z_s)\big) \leq c_6 \sum_{k=1}^{s} \xi_k \sum_{\alpha \in \mathcal{B}_k} \frac{|\alpha|(d-1)^{|\alpha|+r-1}}{n^{|\alpha|+1}},$$

where $\xi_k = \min\big(1, 1.4\lambda_k^{-1/2}\big)$.

Proof. Define the exchangeable pair (G, G') as in Theorem 2.17, and define $I_1', \ldots, I_s'$ as the analogous quantities in G'. Define Δ_k^+ and Δ_k^- as in Proposition 2.8, noting that

$$\mathbf{P}[\Delta_k^+ \mid G] = \sum_{\alpha \in \mathcal{B}_k} \frac{B_\alpha}{d_0 [n]_{|\alpha|} d^{|\alpha|}}, \qquad \mathbf{P}[\Delta_k^- \mid G] = \sum_{\alpha \in \mathcal{B}_k} \frac{F_\alpha}{d_0 [n]_{|\alpha|} d^{|\alpha|}}.$$

By Proposition 2.8,

$$\begin{aligned} d_{TV}&\big((I_1, \ldots, I_s),\, (Z_1, \ldots, Z_s)\big) \\ &\leq \sum_{k=1}^{s} \xi_k \left(\mathbf{E}\left|\lambda_k - d_0 \mathbf{P}[\Delta_k^+ \mid G]\right| + \mathbf{E}\left|I_k - d_0 \mathbf{P}[\Delta_k^- \mid G]\right|\right) \\ &= \sum_{k=1}^{s} \xi_k \mathbf{E}\left|\sum_{\alpha \in \mathcal{B}_k} \left(\frac{(d-1)^{|\alpha|}}{[n]_{|\alpha|}} - \frac{B_\alpha}{[n]_{|\alpha|} d^{|\alpha|}}\right)\right| \\ &\qquad + \sum_{k=1}^{s} \xi_k \mathbf{E}\left|\sum_{\alpha \in \mathcal{B}_k} \left(I_\alpha - \frac{F_\alpha}{[n]_{|\alpha|} d^{|\alpha|}}\right)\right|. \end{aligned}$$

These summands were already bounded in expectation in Theorem 2.17, and applying these bounds proves the proposition. □

Proof of Theorem 2.5. If $d > n^{1/3}$ or $r > n^{1/10}$, then $c\sqrt{r}(d-1)^{3r/2-1}/n > 1$ for a sufficiently large choice of c, and the theorem holds trivially. Thus we can assume that $d \le n^{1/3}$ and $r \le n^{1/10}$.

Let $\lambda_k = (d-1)^k/2k$. With $\mathcal{I}_k$ defined as the set of all cycles in K_n of length k, we apply the previous proposition with bins $\mathcal{I}_3, \dots, \mathcal{I}_r$ to get

$$\begin{aligned} d_{TV}\big((C_3,\dots,C_r),\,(Z_3,\dots,Z_r)\big) &\le c_6 \sum_{k=3}^{r} 1.4\lambda_k^{-1/2} \sum_{\alpha\in\mathcal{I}_k} \frac{k(d-1)^{k+r-1}}{n^{k+1}} \\ &= \sum_{k=3}^{r} O\Big(\frac{\sqrt{k}(d-1)^{k/2+r-1}}{n}\Big) \\ &= O\Big(\frac{\sqrt{r}(d-1)^{3r/2-1}}{n}\Big). \end{aligned}$$

□

Chapter 3

FLUCTUATIONS OF LINEAR EIGENVALUE STATISTICS

3.1 Fluctuations for random regular graphs: main results

Let $\lambda_1 \geq \cdots \geq \lambda_n$ be the eigenvalues of $(d-1)^{-1/2}A_n$, where A_n is the adjacency matrix of a random d-regular graph. The main result is that the fluctuations of $\sum f(\lambda_i)$ for a sufficiently smooth function f converge either in law either to compound Poisson or to Gaussian, depending on whether d is held fixed or grows. The exact limiting distribution depends on f; it can be written in terms of the decomposition of f as a sum of modified Chebyshev polynomials, which we define now:

$$\begin{aligned}
\Gamma_0(x) &= 1, \\
\Gamma_{2k}(x) &= 2T_{2k}\left(\frac{x}{2}\right) + \frac{d-2}{(d-1)^k} && \text{for } k \geq 1, \\
\Gamma_{2k+1}(x) &= 2T_{2k+1}\left(\frac{x}{2}\right) && \text{for } k \geq 0,
\end{aligned}$$

with $\{T_n(x)\}_{n\in\mathbb{N}}$ the Chebyshev polynomials of the first kind on the interval $[-1,1]$.

Let $\rho > 1$, and consider the image of the circle of radius ρ, centered at the origin, under the map $f(z) = \frac{z+z^{-1}}{2}$. We call this the Bernstein ellipse of radius ρ. The ellipse has foci at ± 1, and the sum of the major semiaxis and the minor semiaxis is exactly ρ. Analyticity on a Bernstein ellipse implies a decomposition as a sum of Chebyshev polynomials. We can now give our main result on eigenvalue fluctuations:

Theorem 3.1. *Fix $d \geq 3$, and let G_n be a random d-regular graph on n vertices from the permutation or uniform model, with adjacency matrix A_n. Let $\lambda_1 \geq \cdots \geq \lambda_n$ be the eigenvalues of $(d-1)^{-1/2}A_n$.*

Let $\alpha_0 = 1$ in the case of the permutation model and $\alpha_0 = 3/2$ for the permutation model. Suppose that f is a function defined on $\mathbb{C}$, analytic inside a Bernstein ellipse of radius 2ρ, where $\rho = (d-1)^{\alpha}$ for some $\alpha > \alpha_0$, and such that $|f(z)|$ is bounded inside this ellipse.

Then $f(x)$ can be expanded on $[-2,2]$ as

$$f(x) = \sum_{k=0}^{\infty} a_k \Gamma_k(x),$$

and $Y_f^{(n)} \triangleq \sum_{i=1}^{n} f(\lambda_i) - na_0$ converges in law as $n \to \infty$ to the infinitely divisible random variable

$$Y_f \triangleq \sum_{k=1}^{\infty} \frac{a_k}{(d-1)^{k/2}} \mathrm{CNBW}_k^{(\infty)},$$

with $\mathrm{CNBW}_k^{(\infty)}$ as defined on p. 39 for the permutation or uniform model of random graph.

We can also prove that the limiting distribution of linear eigenvalue functionals is normal when the degree of G_n grows with n. The conditions of the theorem are messy, and more needs to be defined before we can even state it. The result is found in Theorem 3.5.

3.2 Proof of eigenvalue fluctuation results

We will use Theorems 2.4 and 2.5 to estimate the distribution of *cyclically non-backtracking walks* in a random regular graph. As we will see in Proposition 3.3, counts of these walks can be written in terms of the graph's eigenvalues, which allows us to compute the limiting fluctuations of linear eigenvalue statistics.

If a walk on a graph begins and ends at the same vertex, we call it *closed.* We call a walk on a graph *non-backtracking* if it never follows an edge and immediately follows that same edge backwards. Non-backtracking walks are also known as irreducible.

Consider a closed non-backtracking walk, and suppose that its last step is anything other than the reverse of its first step (i.e., the walk does not look like the one given in Figure 3.1). Then we call it a *cyclically non-backtracking walk.* These walks occasionally go by the name strongly irreducible.

Let G_n be a random d-regular graph on n vertices, with the exact model to be specified later. To allow for more consistent statements between the permutation and uniform models, we talk about d-regular graphs rather than $2d$-regular graphs from the permutation model, with the understanding that d is even. Let $C_k^{(n)}$ be the number of cycles of length k in G_n. We define the random variable $\mathrm{CNBW}_k^{(n)}$ to be the number of cyclically non-backtracking

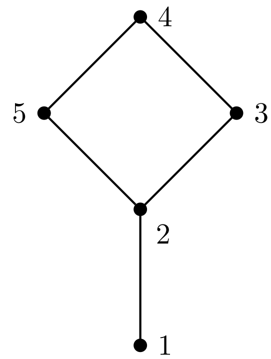


Figure 3.1: The walk $1 \to 2 \to 3 \to 4 \to 5 \to 2 \to 1$ is non-backtracking, but not cyclically non-backtracking. Such walks have a "lollipop" shape.

walks of length k in G_n. Define $(C_k^{(\infty)},\, k \geq 1)$ to be independent Poisson random variables. When we discuss the permutation model, take $\mathbf{E}C_k^{(\infty)} = a(d,k)/2k$. When we work with the uniform model, take $\mathbf{E}C_k^{(\infty)} = (d-1)^k/2k$ for $k \geq 3$, and define $C_1^{(\infty)}$, $C_2^{(\infty)}$, $C_1^{(n)}$, and $C_2^{(n)}$ as zero.

Define

$$\mathrm{CNBW}_k^{(\infty)} = \sum_{j|k} 2jC_j^{(\infty)}.$$

For any cycle in G_n of length j, where j divides k, we obtain $2j$ cyclically non-backtracking walks of length k by choosing a starting point and direction and then walking around the cycle repeatedly. In fact, if d and k are small compared to n, then these are likely to be the only cyclically non-backtracking walks of length k in G_n, as we will prove in the course of the following theorems.

Theorem 3.2. *For some absolute constant c, it holds in the permutation model of random d-regular graph that*

$$d_{TV}\Big(\big(\mathrm{CNBW}_k^{(n)},\, 1 \leq k \leq r\big), \big(\mathrm{CNBW}_k^{(\infty)},\, 1 \leq k \leq r\big)\Big) \leq \frac{cr^4(d-1)^r}{n},$$

and in the uniform model of random d-regular graph that

$$d_{TV}\Big(\big(\mathrm{CNBW}_k^{(n)},\ 1\le k\le r\big),\big(\mathrm{CNBW}_k^{(\infty)},\ 1\le k\le r\big)\Big)\le \frac{c\sqrt{r}(d-1)^{3r/2}}{n}.$$

Proof. For any measurable function f and random variables X and Y, we have $d_{TV}(f(X),\ f(Y))\le d_{TV}(X,Y)$. It follows by Theorem 2.4 that in the permutation model,

$$d_{TV}\left(\Big(\sum_{j|k}2jC_j^{(n)},\ 1\le k\le r\Big),\big(\mathrm{CNBW}_k^{(\infty)},\ 1\le k\le r\big)\right)\le O\bigg(\frac{r^2(d-1)^r\log(d-1)}{n}\bigg), \tag{3.1}$$

and it follows by Theorem 2.5 that in the uniform model,

$$d_{TV}\left(\Big(\sum_{j|k}2jC_j^{(n)},\ 1\le k\le r\Big),\big(\mathrm{CNBW}_k^{(\infty)},\ 1\le k\le r\big)\right)\le O\bigg(\frac{\sqrt{r}(d-1)^{3r/2-1}}{n}\bigg). \tag{3.2}$$

To finish the proof, we will show that

$$\Big(\sum_{j|k}2jC_j^{(n)},\ 1\le k\le r\Big)=\big(\mathrm{CNBW}_k^{(n)},\ 1\le k\le r\big) \tag{3.3}$$

with high probability, in both models. We go out of order and consider the uniform model first. These two vectors differ exactly when either of the following occur:

Event E_1: G_n contains a j-cycle and a k-cycle with a vertex in common, with $j+k\le r$.

Event E_2: G_n contains a j-cycle and a k-cycle whose distance is l, with $l\ge 1$ and $j+k+2l\le r$ (see Figure 2.2).

We have already done most of the work in bounding the probability of event E_1. Let α be some arbitrary k-cycle. In (2.31), we bounded the probability that G_n contained α and another cycle sharing an *edge* with α. With the same notation and nearly the same analysis (the only real change is allowing f to be zero),

$$\begin{aligned}\mathbf{P}[E_1]&\le\sum_{k=3}^{r-3}\frac{[n]_k}{2k}\sum_{j=3}^{r-3-k}\sum_{\substack{p\ge1,\\ f\ge0}}\frac{(2r^3)^{p-1}}{(p-1)!^2}2kn^{j-p-f}\frac{c_2(d-1)^{j+k-f}}{n^{j+k-f}}\\&=O\left(\frac{(d-1)^r}{n}\right).\end{aligned}$$

To bound the probability of E_2, first observe that the number of subgraphs of K_n consisting of a j-cycle and a k-cycle (which do not overlap) connected by a path of length l is $[n]_{j+k+l-1}/4$. By Proposition 2.14c, each of these is contained in G_n with probability at $O\big((d-1)^{j+k+l}/n^{j+k+l}\big)$. By a union bound,

$$\begin{aligned}\mathbf{P}[E_2] &\leq \sum_{j+k+2l\leq r} \frac{[n]_{j+k+l-1}}{4} O\left(\frac{(d-1)^{j+k+l}}{n^{j+k+l}}\right) \\ &= O\left(\frac{(d-1)^{r-1}}{n}\right).\end{aligned}$$

Thus (3.3) holds with probability $1 - O\big((d-1)^r/n\big)$. If two random variables are equal with probability $1-\epsilon$, then the total variation distance between their laws is at most ϵ. Thus the two random vectors in (3.3) have total variation distance $O\big((d-1)^r/n\big)$. This fact and (3.2) prove the theorem for the uniform model.

A similar argument in the permutation model would work. Instead, we will just cite [DJPP13, Corollary 16], which says that (3.3) holds with probability $O(r^4(d-1)^r/n)$ in the permutation model, using an argument based on [LP10]. This together with (3.1) completes the proof. □

Next, we will relate Theorem 3.2 to the eigenvalues of the adjacency matrix of G_n. Recall the modified Chebyshev polynomials $\Gamma_k(x)$ defined in Section 3.1. The following proposition is folkloric, following a long tradition of linking up counts of walks on graphs with polynomial traces of their adjacency matrices.

Proposition 3.3 ([DJPP13, Proposition 32]). *Let A_n be the adjacency matrix of G_n, and let $\lambda_1 \geq \cdots \geq \lambda_n$ be the eigenvalues of $(d-1)^{-1/2}A_n$. Then*

$$\sum_{i=1}^{n} \Gamma_k(\lambda_i) = (d-1)^{-k/2}\mathrm{CNBW}_k^{(n)}.$$

By Theorem 3.2, we know the limiting distribution of $\sum_{i=1}^{n} f(\lambda_i)$ when $f(x) = \Gamma_k(x)$. The plan now is to extend this to a more general class of functions by approximating by this polynomial basis. We will need the following bounds on the eigenvalues of random regular graphs.

Proposition 3.4. *Let G_n be a random d-regular graph on n vertices, in either the permutation or uniform models.*

(a) Suppose that $d \geq 3$ is fixed. For any $\epsilon > 0$, asymptotically almost surely, all but the highest eigenvalue of G_n is bounded by $2\sqrt{d-1}+\epsilon$.

(b) Suppose that $d = d(n)$ satisfies $d = o(n^{1/2})$. Then for some absolute constant c_7, asymptotically almost surely, all but the highest eigenvalue of G_n is bounded by $c_7\sqrt{d}$.

Proof. In the permutation model, [Fri08, Theorem 1.1] proves (a) and [DJPP13, Theorem 24] proves (b).

In the uniform model, it is well known that (a) follows from the results in [Fri08] by various contiguity results, but we cannot find an argument written down anywhere and will give one here. When d is even, it follows from [Fri08, Theorem 1.1] and the fact that for fixed d, permutation random graphs have no loops or multiple edges with probability bounded away from zero. This implies that the eigenvalue bound holds for permutation random graphs conditioned to be simple, and [GJKW02, Corollary 1.1] transfers the result to the uniform model. When d is odd (and n even, as it has to be), we apply [Fri08, Theorem 1.3], which gives the eigenvalue bound for graphs formed by superimposing d random perfect matchings of the n vertices. These are simple with probability bounded away from zero, and [Wor99a, Corollary 4.17] transfers the result to the uniform model.

Fact (b) in the uniform model is proven in a more general context in [BFSU99, Lemma 18]. □

Proof of Theorem 3.1. The following facts about the Chebyshev approximation follow exactly as in Lemma 34 of [DJPP13]:

(i) The Chebyshev series approximation for $f(x)$ converges pointwise on the interval $\left[2, d/\sqrt{d-1}\right]$.

(ii) The series converges uniformly on $[-2-\epsilon, 2+\epsilon]$, for some $\epsilon > 0$. In fact, defining the partial sum $f_m(x) = \sum_{k=0}^{m} a_k \Gamma_k(x)$,

$$\sup_{|x|\leq 2+\epsilon} |f(x) - f_m(x)| \leq M(d-1)^{-\alpha' m},$$

where M is a constant depending on f and d, and $\alpha_0 < \alpha' < \alpha$.

(iii) The coefficients obey the bound

$$|a_k| \leq M(d-1)^{-\alpha k}.$$

The sum defining Y_f converges almost surely, since it can be rewritten as

$$Y_f = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} \frac{a_{ij}}{(d-1)^{ij/2}} 2jC_j^{(\infty)},$$

and this is a sum of independent random variables, bounded in L^2 by fact (iii). For some $\beta < 1/\alpha$, define

$$r_n = \left\lfloor \frac{\beta \log n}{\log(d-1)} \right\rfloor,$$
$$X_f^{(n)} = \sum_{k=1}^{r_n} \frac{a_k}{(d-1)^{k/2}} \mathrm{CNBW}_k^{(n)}.$$

We will use $X_f^{(n)}$ to approximate $Y_f^{(n)}$, noting that $X_f^{(n)} = \sum_{i=1}^{n} f_{r_n}(\lambda_i) - na_0$. By Theorem 3.2 and our choice of r_n,

$$\lim_{n\to\infty} d_{TV}\left(X_f^{(n)}, \sum_{k=1}^{r_n} \frac{a_k}{(d-1)^{k/2}} \mathrm{CNBW}_k^{(\infty)} \right) = 0.$$

This sum converges almost surely to Y_f as n tends to infinity, so $X_f^{(n)}$ converges in law to Y_f. By Slutsky's Theorem, we need only show that $Y_f^{(n)} - X_f^{(n)}$ converges to zero in probability.

Fix $\delta > 0$. We need to show that

$$\lim_{n\to\infty} \mathbf{P}\left[\left| Y_f^{(n)} - X_f^{(n)} \right| > \delta \right] = 0.$$

We have

$$\left| Y_f^{(n)} - X_f^{(n)} \right| \leq \sum_{i=1}^{n} |f(\lambda_i) - f_{r_n}(\lambda_i)|.$$

The top eigenvalue λ_1 is always equal to $d/2\sqrt{d-1}$, and by fact (i), we have the deterministic limit $f_{r_n}(\lambda_1) \to f(\lambda_1)$. Thus $f(\lambda_i) - f_{r_n}(\lambda_i) < \delta/2$ for all sufficiently large n.

Suppose that the remaining eigenvalues are contained in $[-2-\epsilon, 2+\epsilon]$. By fact (ii),

$$\sum_{i=2}^{n} |f(\lambda_i) - f_{r_n}(\lambda_i)| \leq M(n-1)(d-1)^{-\alpha' r_n} \leq M n^{-\alpha'\beta+1}, \tag{3.4}$$

and this tends to zero since $\alpha'\beta < 1$. For sufficiently large n, this sum is thus bounded by $\delta/2$. We can conclude that for all large enough n,

$$\mathbf{P}\left[\left|Y_f^{(n)} - X_f^{(n)}\right| > \delta\right] \leq \mathbf{P}\left[\sup_{2\leq i\leq n} |\lambda_i| \leq 2+\epsilon\right],$$

and this tends to zero by Proposition 3.4a. □

The following theorem can be applied only when the degree of the graph grows more slowly than any positive power of n. This does not appear explicitly in the statement of the theorem, but its conditions cannot be satisfied otherwise.

To remove dependence on d from our polynomial basis, define

$$\Phi_0(x) = 1$$
$$\Phi_k(x) = 2T_k\left(\frac{x}{2}\right) \quad \text{for } k \geq 1.$$

Theorem 3.5. *Let G_n be a random d_n-regular graph on n vertices from the permutation or uniform models, with $d_n \to \infty$ as $n \to \infty$. Let $\lambda_1 \geq \cdots \geq \lambda_n$ be the eigenvalues of $(d_n-1)^{-1/2}A_n$. Suppose f is an entire function on $\mathbb{C}$, and recall c_7 from Proposition 3.4. The function f admits the absolutely convergent expansion $f(x) = \sum_{i=0}^{\infty} a_i\Phi_i(x)$ on $[-c_7, c_7]$. Denote the kth truncation of this series by $f_k \triangleq \sum_{i=0}^{k} a_i\Phi_i$. Let*

$$r_n = \left\lfloor \frac{\beta \log n}{\log(d_n-1)} \right\rfloor,$$

with β to be specified later. Suppose that the following conditions on f hold:

(i) Let $\alpha_0 = 1$ in the case of the permutation model and $\alpha_0 = 3/2$ for the uniform model. For some $\alpha > \alpha_0$ and $M > 0$,

$$\sup_{|x|\leq c_7} |f(x) - f_k(x)| \leq M \exp(-\alpha k h(k)),$$

where h is some function such that $h(r_n) \geq \log(d_n - 1)$ for some choice of $\beta < 1/\alpha$, for sufficiently large n.

(ii)

$$\lim_{n\to\infty} \left| f_{r_n}\left(d_n(d_n-1)^{-1/2}\right) - f\left(d_n(d_n-1)^{-1/2}\right)\right| = 0.$$

Let $\mu_k(d) = \mathbf{E}\mathrm{CNBW}_k^{(\infty)}$, noting that $\mathrm{CNBW}_k^{(\infty)}$ depends on d. We define the following array of constants, which we will use to recenter the random variable $\sum_{i=1}^n f(\lambda_i)$:

$$m_f(n) \triangleq na_0 + \sum_{k=1}^{r_n} \frac{a_k}{(d_n-1)^{k/2}}\big(\mu_k(d_n) - (d_n-2)n\mathbf{1}\{k \text{ is even}\}\big)$$

Then, as $n \to \infty$, the random variable

$$Y_f^{(n)} \triangleq \sum_{i=1}^n f(\lambda_i) - m_f(n)$$

converges in law to a normal random variable with mean zero and variance $\sigma_f = \sum_{k=1}^\infty 2ka_k^2$ for the permutation model case and $\sigma_f = \sum_{k=3}^\infty 2ka_k^2$ for the uniform model case.

Proof. As n tends tends to infinity, so does r_n, since assumption (i) could not be satisfied otherwise. Let $k_0 = 1$ in the permutation model case and $k_0 = 3$ in the uniform model case. Define

$$X_f^{(n)} = \sum_{k=k_0}^{r_n} \frac{a_k}{(d_n-1)^{k/2}}\mathrm{CNBW}_k^{(n)} - \mathbf{E}\sum_{k=k_0}^{r_n} \frac{a_k}{(d_n-1)^{k/2}}\mathrm{CNBW}_k^{(\infty)}$$

and

$$\widetilde{X}_f^{(n)} = \sum_{k=k_0}^{r_n} \frac{a_k}{(d_n-1)^{k/2}}\mathrm{CNBW}_k^{(\infty)} - \mathbf{E}\sum_{k=k_0}^{r_n} \frac{a_k}{(d_n-1)^{k/2}}\mathrm{CNBW}_k^{(\infty)},$$

noting that $X_f^{(n)} = \sum_{i=1}^n f_{r_n}(\lambda_i) - m_f(n)$. Also, note that $\mathrm{CNBW}_k^{(\infty)}$ depends on d_n.

Let

$$N_k^{(n)} = \begin{cases} (d_n-1)^{-k/2}\big(\mathrm{CNBW}_k^{(\infty)} - \mathbf{E}\mathrm{CNBW}_k^{(\infty)}\big) & \text{if } k \leq r_n, \\ 0 & \text{otherwise,} \end{cases}$$

and let $Z_1, Z_2, \ldots$ be independent normals with $\mathbf{E}Z_k = 0$ and $\mathbf{E}Z_k^2 = 2k$. We will show that $(N_k^{(n)},\, k \geq k_0)$ converges in law to $(Z_k,\, k \geq k_0)$ as $n \to \infty$. Rewrite $N_k^{(n)}$ as

$$N_k^{(n)} = \frac{1}{(d_n-1)^{k/2}}\bigl(2kC_k^{(\infty)} - (d_n-1)^k\bigr) + \frac{1}{(d_n-1)^{k/2}}\sum_{\substack{j|k\\ j<k}}\bigl(2jC_j^{(\infty)} - (d_n-1)^j\bigr).$$

The first term converges to a centered normal with variance $2k$ as $n \to \infty$, by the normal approximation of the Poisson distribution. The random variables $C_1^{(\infty)}, C_2^{(\infty)}, \ldots$ are independent, so the convergence of $(N_k^{(n)},\, k \geq k_0)$ follows if we show that the remaining terms converge to zero in probability. This holds by Chebyshev's inequality, since

$$\mathbf{Var}\left[\frac{1}{(d_n-1)^{k/2}}\sum_{\substack{j|k\\ j<k}}\bigl(2jC_j^{(\infty)} - (d_n-1)^j\bigr)\right] = \sum_{\substack{j|k\\ j<k}} O\bigl(2j(d_n-1)^{j-k}\bigr),$$

and this vanishes as d_n grows.

It follows by the continuous mapping theorem that $\widetilde{X}_f^{(n)}$ converges to normal with variance σ_f. By Theorem 3.2, the total variation distance between $X_f^{(n)}$ and $\widetilde{X}_f^{(n)}$ approaches zero as $n \to \infty$, so $X_f^{(n)}$ converges in law to the same limit.

All that remains is to show that $Y_f^{(n)} - X_f^{(n)}$ converges to zero in probability. This follows exactly as in Theorem 3.1, using assumptions (i) and (ii) and Proposition 3.4b. $\square$

Chapter 4

MINOR PROCESSES AND THE GAUSSIAN FREE FIELD

The typical approach to random matrices is to consider a sequence of random matrices X_n of increasing size. Each matrix X_n is considered in isolation; the different matrices are not considered on a common probability space, so they have no joint distribution. Some recent work has instead looked at the matrices together on a single probability space. For example, suppose that X is an infinite random Hermitian matrix with independent real standard Gaussians along the diagonal and independent complex standard Gaussians above the diagonal. Let X_n be the first n rows and columns of X. Then X_n is drawn from the *Gaussian Unitary Ensemble*, and the joint distribution of the eigenvalues of these matrices is called the *GUE-corners process* or *GUE-minors process*. This process was studied in [Bar01] and [JN06]. One can also form general β-Hermite corners processes [GS14, Definition 1.1] and β-Jacobi corners processes [BG13]. These processes are closely related to interacting particle systems; see [Fer14] for a survey. There are also many connections with the KPZ universality class of random surfaces [BF14]. Minors of Dyson's Brownian motion have also been studied [ANvM12] and can be put into a common framework with corners processes [War07, GS14].

The connection to the Gaussian free field (to be called the GFF from now on) comes from [Bor10a]. We describe a particular but important case of that paper's main result, given by considering only the single sequence $\{1, 2, \ldots\}$. Let W be an infinite symmetric matrix whose entries have all moments finite. Suppose the the entries above the diagonal are i.i.d. and match the standard Gaussian to four moments, and the diagonal entries have variance 2. Let W_n be the matrix consisting of the first n rows and columns of W. Borodin then considered the joint eigenvalue fluctuations of these random matrices.

Let z be a complex number in the upper half plane $\mathbb{H}$. Define $y = |z|^2$ and $x = 2\Re(z)$. Consider the minor $W(\lfloor ny \rfloor)$, and let $N(z)$ be the number of its eigenvalues that are greater than or equal to $\sqrt{n}x$. Define the *height function*

$$H_n(z) \triangleq \sqrt{\frac{\pi}{2}} N(z). \tag{4.1}$$

Then Borodin shows that $\{H_n(z) - EH_n(z),\ z \in \mathbb{H}\}$, converges in a certain sense to the GFF on $\mathbb{H}$, a random generalized function that we describe in more detail in Sections 4.5.1 and 4.5.2.

We will prove a similar result for the eigenvalue fluctuations of the growing random regular graphs described in Section 1.4.3. Our first result is about the process of *short cycles* in the graph process $G(t)$. By a cycle of length k in a graph, we mean what is sometimes called a simple cycle: a walk in the graph that begins and ends at the same vertex, and that otherwise repeats no vertices. We will give a more formal definition in Section 4.1.2. Let $(C_k^{(s)}(t),\ k \in \mathbb{N})$ denote the number of cycles of various lengths k that are present in $G(s+t)$. This process is not Markov, but nonetheless it converges to a Markov process (indexed by t) as s tends to infinity.

To describe the limit, recall the value of $a(d,k)$, given in (2.1). Consider the set of natural numbers $\mathbb{N} = \{1, 2, \ldots\}$ with the measure

$$\mu(k) = \frac{1}{2}\left[a(d,k) - a(d,k-1)\right], \quad k \in \mathbb{N}, \quad a(d,0) \triangleq 0.$$

Consider a Poisson point process χ on $\mathbb{N} \times [0,\infty)$ with an intensity measure given on $\mathbb{N} \times (0,\infty)$ by the product measure $\mu \otimes \mathrm{Leb}$, where Leb is the Lebesgue measure, and with additional masses of $a(d,k)/2k$ on $(k,0)$ for $k \in \mathbb{N}$.

Let $\widetilde{P}_x$ denote the law of an one-dimensional pure-birth process on $\mathbb{N}$ given by the generator:

$$Lf(k) = k\left(f(k+1) - f(k)\right), \quad k \in \mathbb{N},$$

starting from $x \in \mathbb{N}$. This is also known as the *Yule process*.

Suppose we are given a realization of χ. For any atom (k,y) of the countably many atoms of χ, we start an independent process $(X_{k,y}(t),\ t \geq 0)$ with law $\widetilde{P}_k$. Define the random sequence

$$N_k(t) \triangleq \sum_{(j,y) \in \chi \cap \{[k] \times [0,t]\}} 1\left\{X_{j,y}(t-y) = k\right\}.$$

In other words, at time t, for every site k, we count how many of the processes that started at time $y \leq t$ at site $j \leq k$ are currently at k. Note that both $(N_k(\cdot),\ k \in \mathbb{N})$ and $(N_k(\cdot),\ k \in [K])$, for some $K \in \mathbb{N}$, are Markov processes, while $N_k(\cdot)$ for fixed k is not.

Theorem 4.1. *As $s \to \infty$, the process $(C_k^{(s)}(t),\ k \in \mathbb{N},\ 0 \leq t < \infty)$ converges in law in the Skorokhod space $D_{\mathbb{R}^\infty}[0,\infty)$ to the Markov process $(N_k(t),\ k \in \mathbb{N},\ 0 \leq t < \infty)$. The limiting process is stationary.*

Remark 4.2. In fact, the same argument used to prove Theorem 4.1 shows that the process $(C_k^{(s)}(t),\ -\infty < t < \infty)$ converges in law to the Markov process $(N_k(t),\ -\infty < t < \infty)$ running in stationarity. The same conclusion holds for all the following theorems in this section.

We now focus on eigenvalues of $G(t)$. Note that there is no easy exact relationship between the eigenvalues of G_n for various n since the eigenvectors play a role in determining any such identity. In fact, the eigenvalues of G_n and G_{n+1} need not be interlaced. We will follow the approach of the previous sections and consider linear eigenvalue statistics. For any $2d$-regular graph on n vertices G and function $f\colon \mathbb{R} \to \mathbb{R}$, we will define the random variable

$$\operatorname{tr} f(G) \triangleq \sum_{i=1}^{n} \hat{f}(\lambda_i)$$

where $\lambda_1 \geq \ldots \geq \lambda_n$ are the eigenvalues of adjacency matrix of G divided by $2\sqrt{2d-1}$, and $\hat{f}$ is f with its constant term adjusted (see Definition 4.20 and Remark 4.21 for an explanation). The scaling is necessary to take a limit with respect to d. Let $[n] = \{1, \ldots, n\}$, and let $[\infty] = \mathbb{N}$.

Theorem 4.3. *For each d, there exists a set of polynomials $f_1, f_2, \ldots$ with f_i of degree i such that for any $K \in \mathbb{N} \cup \{\infty\}$, the process $(\operatorname{tr} f_k(G(s+t)),\ k \in [K],\ t \geq 0)$ converges in law, as s tends to infinity, to the Markov process $(N_k(t),\ k \in [K],\ t \geq 0)$ of Theorem 4.1. (The polynomials are given explicitly in (4.12).) For any polynomial f, the process $\big(\operatorname{tr} f(G(s+t))\big)$ converges to a linear combination of the coordinate processes of $(N_k(t),\ k \in \mathbb{N})$.*

Next, we take d to infinity. We will make the following notational convention: for any polynomial f, we will denote the limiting process of $(\operatorname{tr} f(G(s+t)),\ t \geq 0)$ by

$(\operatorname{tr} f\,(G(\infty+t))\,,\ t \geq 0)$. Recall that this process is a linear combination of $(N_k(t),\, k \in \mathbb{N},\, t \geq 0)$.

Theorem 4.4. *Let $\{T_k,\ k \in \mathbb{N}\}$ denote the Chebyshev orthogonal polynomials of the first kind on $[-1,1]$. As d tends to infinity, the collection of processes*

$$(\operatorname{tr} T_k\,(G(\infty+t)) - \mathbf{E}\operatorname{tr} T_k\,(G(\infty+t))\,,\ t \geq 0,\ k \in \mathbb{N})$$

converges weakly in $D_{\mathbb{R}^\infty}[0,\infty)$ to a collection of independent Ornstein-Uhlenbeck processes $(U_k(t),\ t \geq 0,\ k \in \mathbb{N})$, running in equilibrium. Here the equilibrium distribution of U_k is $N(0,k/2)$ and U_k satisfies the stochastic differential equation

$$dU_k(t) = -kU_k(t)\,dt + k\,dW_k(t), \quad t \geq 0,$$

and $(W_k,\ k \in \mathbb{N})$ are i.i.d. standard one-dimensional Brownian motions.

Thus, the collection of random variables $\big(\operatorname{tr} T_k\,(G(\infty+t)) - \mathbf{E}\operatorname{tr} T_k\,(G(\infty+t))\big)$, indexed by k and t, converges as d tends to infinity to a centered Gaussian process with covariance kernel given by

$$\lim_{d\to\infty} \mathbf{Cov}\,(\operatorname{tr} T_i\,(G(\infty+t))\,, \operatorname{tr} T_k\,(G(\infty+s))) = \delta_{ik}\frac{k}{2}e^{k(s-t)}. \tag{4.2}$$

for $s \leq t$.

This covariance structure is intimately linked to the GFF; we will make this more apparent in Theorem 4.6. For the moment, this is best illustrated by a comparison to Borodin's result. We specialize [Bor10a, Proposition 3] for the case of GOE ($\beta = 1$). Fix m positive real numbers $t_1 < t_2 < \ldots < t_m$. In the notation of [Bor10a], we take $L = n$ and $B_i(n) = [\lfloor t_i n \rfloor]$. The matrix $W(n)$ is defined as the first n rows and columns of an infinite Wigner matrix. Then, for any positive integers $j_1, j_2, \ldots, j_m$, the random vector

$$\big(\operatorname{tr} T_{j_i}\,\big(W(\lfloor t_i n \rfloor)/2\sqrt{t_i n}\big) - \mathbf{E}\operatorname{tr} T_{j_i}\,\big(W(\lfloor t_i n \rfloor)/2\sqrt{t_i n}\big)\,,\ i \in [m]\big)$$

converges in law as n tends to infinity to a centered Gaussian vector. For $s \leq t$,

$$\lim_{n\to\infty} \mathbf{Cov}\left(\operatorname{tr} T_i\left(W(\lfloor tn \rfloor)/2\sqrt{tn}\right),\ \operatorname{tr} T_k\left(W(\lfloor sn \rfloor)/2\sqrt{sn}\right)\right) = \delta_{ik}\frac{k}{2}\left(\frac{s}{t}\right)^{k/2},$$

nearly the same as (4.4). The appearance of the exponential in (4.4) comes from the time-change we introduced when we made our graph process $G(t)$ run in continuous time.

Here, we have taken a limit in t followed by a limit in d. When we take the limit in t, we get an abstract limiting object. In order to give a direct connection between the eigenvalue fluctuations and the GFF, we need to take the two limits simultaneously. As we now vary both t and d, recall the notation $G(t,d)$ from Section 1.4.3. Let $N(t)$ be the number of vertices in $G(t,d)$, which does not depend on d.

Proposition 4.5. *There exists an increasing, right-continuous $d(t)$ taking integer values and growing to infinity such that as $s \to \infty$, the process*

$$\Big(\operatorname{tr} T_k\big(G(s+t, 2d(s+t))\big) - \mathbf{E}\big[\operatorname{tr} T_k\big(G(s+t, 2d(s+t))\big) \,\big|\, N(t)\big],\ k \in \mathbb{N},\ t \geq 0\Big)$$

converges weakly in $D_{\mathbb{R}^\infty}[0,\infty)$ to the same limit of Ornstein-Uhlenbeck processes $(U_k(\cdot),\ k \geq 1)$ as in Theorem 4.4.

Now, we define a height function $F_t(x)$ as the number of eigenvalues of the adjacency matrix of $G(t, 2d(t))$ that are less than or equal to $2\sqrt{2d(t)-1}x$, taking $d(t)$ from the previous proposition. Let $\overline{F}_t(x)$ denote the centered height function

$$\overline{F}_t(x) = F_t(x) - \mathbf{E}[F_t(x) \mid N(t)].$$

(We need to subtract off this conditional expectation, not just the expectation, because otherwise the fluctuations of $N(t)$ swamp the eigenvalue fluctuations that we are interested in.) Define

$$H_s(x,t) = \sqrt{\frac{\pi}{2}}\,\overline{F}_{s+t}(x).$$

As $s \to \infty$, these functions converge to the GFF in the following sense:

Theorem 4.6. *Let $\Omega(x,t) = e^t\big(x + i\sqrt{1-x^2}\big)$ for $-1 \leq x \leq 1$. Let h denote the GFF on $\mathbb{H}$ with vanishing Dirichlet boundary conditions. For any polynomials $p_1(x), \ldots, p_n(x) \in \mathbb{C}[x]$ and times $t_1, \ldots, t_n$,*

$$\left(\int_{-\infty}^{\infty} p_i(x) H_s(x, t_i))\,dx,\ 1 \leq i \leq n\right) \xrightarrow{\mathcal{L}} \left(\int_{-1}^{1} p_i(x) h(\Omega(x, t_i))\,dx,\ 1 \leq i \leq n\right)$$

as $s \to \infty$.

Remark 4.7. A common model for random regular graphs is the *configuration model* or *pairing model* (see [Wor99b] for more information). The model is defined as follows: Start with n buckets, each containing d prevertices. Then, separate these dn prevertices into pairs, choosing uniformly from every possible pairing. Finally, collapse each bucket into a single vertex, making an edge between one vertex and another if a prevertex in one bucket is paired with a prevertex in the other bucket. This model has the advantage that choosing a graph from it conditional on it containing no loops or parallel edges is the same as choosing a graph uniformly from the set of graphs without loops and parallel edges. The model also allows for graphs of odd degrees, unlike the permutation model.

It is possible to construct a process of growing random regular graphs similar to the one in this paper using a dynamic version of this model. Given some initial pairing of prevertices labeled $\{1, \ldots, dn\}$, extend it to a random pairing of $\{1, \ldots, dn + 2\}$ by the following procedure: Choose X uniformly from $\{1, \ldots, dn + 1\}$. Pair $dn + 2$ with X. If $X = dn + 1$, leave the other pairs unchanged; if not, pair the previous partner of X with $dn + 1$. This is an analogue of the Chinese Restaurant Process in the setting of random pairings, in that if the initial pairing is uniformly chosen, then so is the extended one.

If d is odd, we repeat this procedure a total of d times to extend a random d-regular graph on n vertices to have $n + 2$ vertices (when d is odd, the number of vertices in the graph must be even). When d is even, repeat $d/2$ times to add one new vertex to a random graph. In this way, we can construct a sequence of growing random regular graphs. We believe that all the results of this paper hold in this model with minor changes, with similar proofs.

4.1 Preliminaries

4.1.1 A primer on the Chinese Restaurant Process

The Chinese Restaurant Process, introduced by Dubins and Pitman, is a particular example of a two parameter family of stochastic processes that constructs sequentially random exchangeable partitions of the positive integers via the cyclic decomposition of a random permutation. Our short description is taken from [Pit06, Section 3.1].

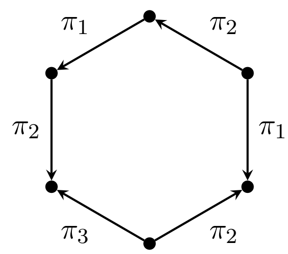


Figure 4.1: A cycle whose word is the equivalence class of $\pi_2\pi_1^{-1}\pi_2\pi_1\pi_2\pi_3^{-1}$ in $\mathcal{W}_6/D_{12}$.

An initially empty restaurant has an unlimited number of circular tables numbered $1, 2, \ldots$, each capable of seating an unlimited number of customers. Customers numbered $1, 2, \ldots$ arrive one by one and are seated at the tables according to the following plan. Person 1 sits at table 1. For $n \geq 1$ suppose that n customers have already entered the restaurant, and are seated in some arrangement, with at least one customer at each of the tables j for $1 \leq j \leq k$ (say), where k is the number of tables occupied by the first n customers to arrive. Let customer $n+1$ choose with equal probability to sit at any of the following $n+1$ places: to the left of customer j for some $1 \leq j \leq n$, or alone at table $k+1$. Define $\pi^{(n)} \colon [n] \to [n]$ as the permutation whose cyclic decomposition is given by the tables; that is, if after n customers have entered the restaurant, customers i and j are seated at the same table, with i to the left of j, then $\pi^{(n)}(i) = j$, and if customer i is seated alone at some table then $\pi^{(n)}(i) = i$. The sequence $(\pi^{(n)})$ is then a tower of random permutations as defined in Section 1.4.3.

4.1.2 Combinatorics on words

Recall the discussion on p. 8 on viewing the graph formed from independent permutations $\pi_1^{(n)}, \ldots, \pi_d^{(n)}$ as a directed, edge-labeled graph. As we did there, we drop the subscripts and let $\pi_l = \pi_l^{(n)}$. We previously discussed the word formed as we walked around a cycle by writing down the label of each edge as it is traversed, putting π_i or π_i^{-1} according to the direction we walk over the edge. Now, we will treat this more rigorously.

Let $\mathcal{W}_k$ denote the set of cyclically reduced words of length k. We would like to associate each k-cycle in G_n with the word in $\mathcal{W}_k$ formed by the above procedure, but since we can start the walk at any point in the cycle and walk in either of two directions, there are actually up to $2k$ different words that could be formed by it. Thus we identify elements of $\mathcal{W}_k$ that differ only by rotation and inversion (for example, $\pi_1\pi_2^{-1}\pi_1\pi_2$ and $\pi_1^{-1}\pi_2\pi_1^{-1}\pi_2^{-1}$) and denote the resulting set by $\mathcal{W}_k/D_{2k}$, where D_{2k} is the dihedral group acting on the set $\mathcal{W}_k$ in the natural way.

Definition 4.8 (Properties of words)**.** For any k-cycle in G_n, the element of $\mathcal{W}_k/D_{2k}$ given by walking around the cycle is called the *word* of the cycle (see Figure 4.1). For any word w, let $|w|$ denote the length of w. Let $h(w)$ be the largest number m such that $w = u^m$ for some word u. If $h(w) = 1$, we call w *primitive.* For any $w \in \mathcal{W}_k$, the orbit of w under the action of D_{2k} contains $2k/h(w)$ elements, a fact which we will frequently use. Let $c(w)$ denote the number of pairs of double letters in w, i.e., the number of integers i modulo $|w|$ such that $w_i = w_{i+1}$. For example, $c(\pi_1\pi_1\pi_2^{-1}\pi_2^{-1}\pi_1) = 3$. If $|w| = 1$, we take $c(w) = 0$. We will also consider $|\cdot|$, $h(\cdot)$, and $c(\cdot)$ as functions on $\mathcal{W}_k/D_{2k}$, since they are invariant under cyclic rotation and inversion.

To more easily refer to words in $\mathcal{W}_k/D_{2k}$, choose some canonical representative $w_1 \cdots w_k \in \mathcal{W}_k$ for every $w \in \mathcal{W}_k/D_{2k}$. Based on this, we will often think of elements of $\mathcal{W}_k/D_{2k}$ as words instead of equivalence classes, and we will make statements about the ith letter of a word in $\mathcal{W}_k/D_{2k}$. For $w = w_1 \cdots w_k \in \mathcal{W}_k/D_{2k}$, let $w^{(i)}$ refer to the word in $\mathcal{W}_{k+1}/D_{2k+2}$ given by $w_1 \cdots w_i w_i w_{i+1} \cdots w_k$. We refer to this operation as *doubling* the ith letter of w. A related operation is to *halve* a pair of double letters, for example producing $\pi_1\pi_2\pi_3\pi_4$ from $\pi_1\pi_2\pi_3\pi_4\pi_1$. (Since we apply these operations to words identified with their rotations, we do not need to be specific about which letter of the pair is deleted.) The following technical lemma underpins most of our combinatorial calculations.

Lemma 4.9. *Let $u \in \mathcal{W}_k/D_{2k}$ and $w \in \mathcal{W}_{k+1}/D_{2k+2}$. Suppose that a letters in u can be doubled to form w, and b pairs of double letters in w can be halved to form u. Then*

$$\frac{a}{h(u)} = \frac{b}{h(w)}.$$

Remark 4.10. At first glance, one might expect that $a = b$. The example $u = \pi_1\pi_2\pi_1\pi_1\pi_2$ and $w = \pi_1\pi_1\pi_2\pi_1\pi_1\pi_1\pi_2$ shows that this is wrong, since only one letter in u can be doubled to give w, but two different pairs in w can be halved to give u.

Proof. Let $\mathrm{Orb}(u)$ and $\mathrm{Orb}(w)$ denote the orbits of u and w under the action of the dihedral group in $\mathcal{W}_k$ and $\mathcal{W}_{k+1}$, respectively. When we speak of halving a pair of letters in a word in $\mathrm{Orb}(w)$, always delete the second of the two letters (for example, $\pi_1\pi_2\pi_1$ becomes $\pi_1\pi_2$, not $\pi_2\pi_1$). When we double a letter in a word in $\mathrm{Orb}(u)$, put the new letter after the doubled letter (for example, doubling the second letter of $\pi_1\pi_2^{-1}$ gives $\pi_1\pi_2^{-1}\pi_2^{-1}$, not $\pi_2^{-1}\pi_1\pi_2^{-1}$.)

For each of the $2k/h(u)$ words in $\mathrm{Orb}(u)$, there are a doubling operations yielding a word in $\mathrm{Orb}(w)$. For each of the $(2k+2)/h(w)$ words in $\mathrm{Orb}(w)$, there are b halving operations yielding a word in $\mathrm{Orb}(u)$. For every halving operation on a word in $\mathrm{Orb}(w)$, there is a corresponding doubling operation on a word in $\mathrm{Orb}(u)$ and vice versa, except for halving operations that straddle the ends of the word, as in $\pi_1\pi_2\pi_1$. There are $2b/h(w)$ of these, giving us

$$\begin{aligned}\frac{2ka}{h(u)} &= \frac{(2k+2)b}{h(w)} - \frac{2b}{h(w)} \\ &= \frac{2kb}{h(w)},\end{aligned}$$

and the lemma follows from this. □

Let $\mathcal{W}' = \bigcup_{k=1}^{\infty} \mathcal{W}_k/D_{2k}$, and let $\mathcal{W}'_K = \bigcup_{k=1}^{K} \mathcal{W}_k/D_{2k}$. We will use the previous lemma to prove the following technical property of the $c(\cdot)$ statistic.

Lemma 4.11. *In the vector space with basis* $\{q_w\}_{w\in\mathcal{W}'_K}$,

$$\sum_{w\in\mathcal{W}'_{K-1}} \sum_{i=1}^{|w|} \frac{1}{h(w)} q_{w^{(i)}} = \sum_{w\in\mathcal{W}'_K} \frac{c(w)}{h(w)} q_w.$$

Proof. Fix some $w \in \mathcal{W}_k/D_{2k}$, and let $a(u)$ denote the number of letters of u that can be doubled to give w, for any $u \in \mathcal{W}_{k-1}/D_{2k-2}$. We need to prove that

$$\sum_{u\in\mathcal{W}_{k-1}/D_{2k-2}} \frac{a(u)}{h(u)} = \frac{c(w)}{h(w)}.$$

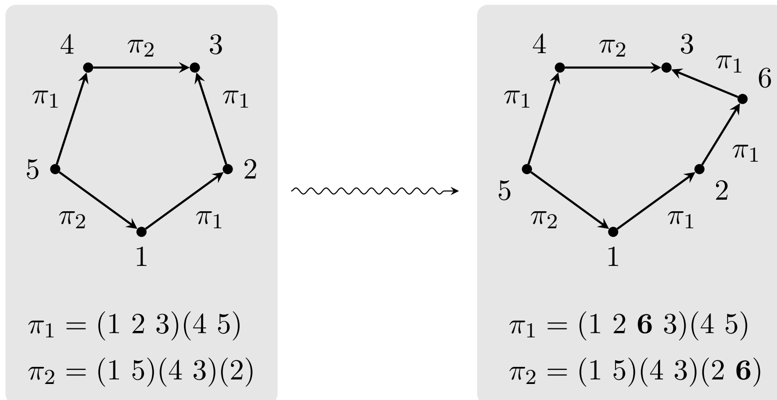


Figure 4.2: The vertex 6 is inserted between vertices 2 and 3 in π_1, causing the above cycle to grow.

Let $b(u)$ be the number of pairs in w that can be halved to give u. By Lemma 4.9,

$$\sum_{u \in \mathcal{W}_{k-1}/D_{2k-2}} \frac{a(u)}{h(u)} = \sum_{u \in \mathcal{W}_{k-1}/D_{2k-2}} \frac{b(u)}{h(w)},$$

and $\sum_{u \in \mathcal{W}_{k-1}/D_{2k-2}} b(u) = c(w)$. □

4.2 The process limit of the cycle structure

As the graph $G(t)$ grows, new cycles form, which we can classify into two types. Suppose a new vertex numbered n is inserted at time t, and this insertion creates a new cycle. If the edges entering and leaving vertex n in the new cycle have the same edge label, then the new cycle has "grown" from a cycle with one fewer vertex, as in Figure 4.2. If the edges entering and leaving n in the cycle have different labels, then the cycle has formed "spontaneously" as in Figure 4.3, rather than growing from a smaller cycle. This classification will prove essential in understanding the evolution of cycles in $G(t)$.

Once a cycle comes into existence in $G(t)$, it remains until a new vertex is inserted into one of its edges. Typically, this results in the cycle growing to a larger cycle, as in Figure 4.2. If a new vertex is simultaneously inserted into multiple edges of the same cycle, the cycle is instead split into smaller cycles as in Figure 4.4. These new cycles are spontaneously formed,

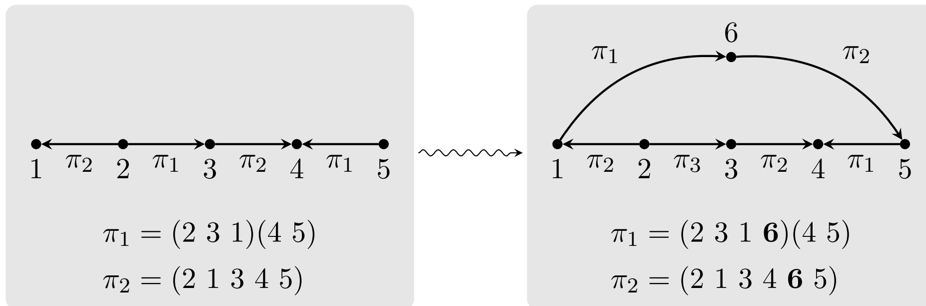


Figure 4.3: A cycle forms "spontaneously" when the vertex 6 is inserted into the graph.

according to the classification of new cycles given in the previous paragraph. Tracking the evolution of these smaller cycles in turn, we see that as the graph evolves, a cycle grows into a cluster of overlapping cycles. However, it will follow from Proposition 4.19 that for short cycles, this behavior is not typical. Thus in our limiting object, cycles will grow only into larger cycles.

4.2.1 Heuristics for the limiting process

We give some estimates that will motivate the definition of the limiting process in Section 4.2.2. This section is entirely motivational, and we will not attempt to make anything rigorous.

Suppose that vertex n is inserted into $G(t)$ at some time t. First, we consider the rate that cycles form spontaneously with some word $w \in \mathcal{W}_k/D_{2k}$. There are $2k/h(w)$ words in the orbit of w under the action of D_{2k}, and out of these, $2(k - c(w))/h(w)$ have nonequal first and last letters. For each such word $u = u_1 \cdots u_k$, we can give a walk on the graph by starting at vertex n and following the edges indicated by u, going from n to $u_1(n)$ to $u_2(u_1(n))$ and so on. If this walk happens to be a cycle, the condition $u_1 \neq u_k$ implies that it would be spontaneously formed.

In a short interval Δt when $G(t)$ has $n - 1$ vertices, the probability that vertex n is inserted is about $n\,\Delta t$. For any word u, the walk from vertex n generated by u is a cycle with probability approximately $1/n$, since after applying the random permutations $u_1, \ldots, u_k$ in

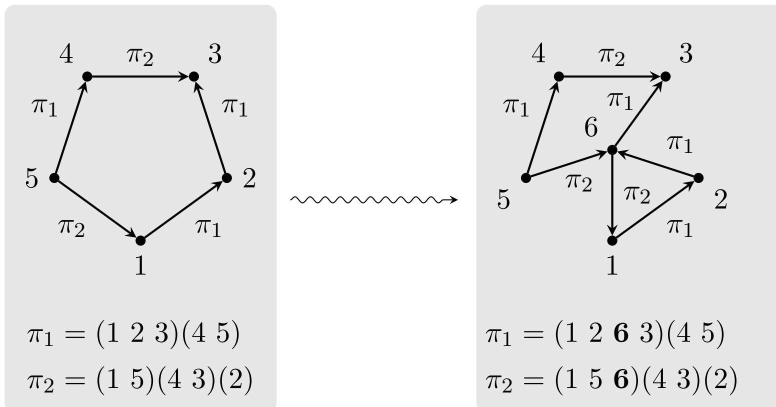


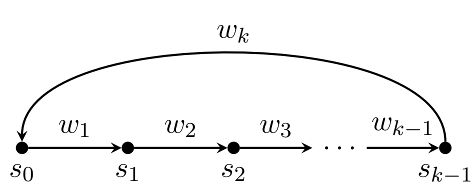
Figure 4.4: The vertex 6 is inserted into the cycle in two different places in the same step, causing the cycle to split in two. Note that each new cycle would be classified as spontaneously formed.

turn, we will be left at an approximately uniform random vertex. Any new spontaneous cycle formed with word w will be counted by one of these walks, with u in the orbit of w, and it will be counted again by the walk generated by $u_k^{-1} \cdots u_1^{-1}$. The expected number of spontaneous cycles formed in a short interval Δt is then approximately

$$\frac{1}{h(w)}(k - c(w))\frac{n\,\Delta t}{n} = \frac{1}{h(w)}(k - c(w))\,\Delta t.$$

Thus we will model the spontaneous formation of cycles with word w by a Poisson process with rate $(k - c(w))/h(w)$.

Next, we consider how often a cycle with word $w \in \mathcal{W}_k$ grows into a larger cycle. Suppose that $G(t)$ has $n - 1$ vertices, and that it contains a cycle of the form



When vertex n is inserted into the graph, the probability that it is inserted after s_{i-1} in permutation w_i is $1/n$. Thus, after a spontaneous cycle with word w has formed, we can model the evolution of its word as a continuous-time Markov chain where each letter is doubled with rate one.

4.2.2 Formal definition of the limiting process

Consider the measure μ on $\mathcal{W}'$ given by

$$\mu(w) = \frac{|w| - c(w)}{h(w)}.$$

Consider a Poisson point process χ on $\mathcal{W}' \times [0, \infty)$ with an intensity measure given by the product measure $\mu \otimes \mathrm{Leb}$, where Leb refers to the Lebesgue measure. Each atom (w, t) of χ represents a new spontaneous cycle with word w formed at time t.

Now, we define a continuous-time Markov chain on the countable space $\mathcal{W}'$ governed by the following rates: From state $w \in \mathcal{W}_k/D_{2k}$, jump with rate one to each of the k words in $\mathcal{W}_{k+1}/D_{2k+2}$ obtained by doubling a letter of w. If a word can be formed in more than one way by doubling a letter in w, then it receives a correspondingly higher rate. For example, from $w = \pi_1\pi_1\pi_2$, the chain jumps to $\pi_1\pi_1\pi_1\pi_2$ with rate two and to $\pi_1\pi_1\pi_2\pi_2$ with rate one. Let $\widetilde{P}_w$ denote the law of this process started from $w \in \mathcal{W}'$.

Suppose we are given a realization of χ. For any atom (w, s) of the countably many atoms of χ, we start an independent process $(X_{w,s}(t),\ t \geq 0)$ with law $\widetilde{P}_w$. Define the stochastic process

$$N_w(t) \triangleq \sum_{\substack{(u,s)\in\chi \\ s\leq t}} 1\left\{X_{u,s}(t-s) = w\right\}.$$

Interpreting these processes as in the previous section, $N_w(t)$ counts the number of cycles formed spontaneously at time s that have grown to have word w at time t.

The fact that the process exists is obvious since one can define the countably many independent Markov chains on a suitable product space. The following lemma establishes some of its key properties.

Lemma 4.12. *Recall that $\mathcal{W}'_L = \bigcup_{k=1}^{L} \mathcal{W}_k/D_{2k}$. We have the following conclusions:*

(i) For any $L \in \mathbb{N}$, the stochastic process $\{(N_w(t),\ w \in \mathcal{W}'_L),\ t \geq 0\}$ is a time-homogeneous Markov process with respect to its natural filtration, with RCLL paths.

(ii) Recall that for $w \in \mathcal{W}_k/D_{2k}$*, the element* $w^{(i)} \in \mathcal{W}_{k+1}/D_{2k+2}$ *is the word formed by doubling the ith letter of* w*. The generator for the Markov process* $\{(N_w(t),\, w \in \mathcal{W}'_L),\ t \geq 0\}$ *acts on* f *at* $x = (x_w,\, w \in \mathcal{W}'_L)$ *by*

$$\begin{aligned}\mathcal{L}f(x) = &\sum_{w \in \mathcal{W}'_L} \sum_{i=1}^{|w|} x_w \left[f(x - e_w + e_{w^{(i)}}) - f(x)\right] \\ &+ \sum_{w \in \mathcal{W}'_L} \frac{|w| - c(w)}{h(w)} \left[f(x + e_w) - f(x)\right],\end{aligned}$$

where e_w *is the canonical basis vector equal to one at entry* w *and equal to zero everywhere else. For a word* u *of length greater than* L*, take* $e_u = 0$*.*

(iii) The product measure of $\mathrm{Poi}(1/h(w))$ *over all* $w \in \mathcal{W}'_L$ *is the unique invariant measure for this Markov process.*

Proof. Conclusion (i) follows from construction, as does conclusion (ii). To prove conclusion (iii), we start by the fundamental identity of the Poisson distribution: if $X \sim \mathrm{Poi}(\lambda)$, then for any function f, we have

$$\mathbf{E}Xg(X) = \lambda \mathbf{E}g(X+1). \tag{4.3}$$

We need to show that if the coordinates of $X = (X_w,\, w \in \mathcal{W}'_L)$ are independent Poisson random variables with $\mathbf{E}X_w = 1/h(w)$, then

$$\mathbf{E}\mathcal{L}f(X) = 0. \tag{4.4}$$

Since the process is an irreducible Markov chain on countable state space, the existence of one invariant distribution shows that the chain is positive recurrent and that the invariant distribution is unique.

To argue (4.4) we will repeatedly apply identity (4.3) to functions g constructed from f by keeping all but one coordinate fixed. Thus, for any $w \in \mathcal{W}'_L$ and $1 \leq i \leq |w|$, we condition on all X_u with $u \neq w$ and hold those coordinates of f fixed to obtain,

$$\mathbf{E}X_w f\left(X - e_w + e_{w^{(i)}}\right) = \frac{1}{h(w)} \mathbf{E}f\left(X + e_{w^{(i)}}\right)$$

taking $e_{w^{(i)}} = 0$ when $|w| = L$. In the same way,

$$\mathbf{E} X_w f(X) = \frac{1}{h(w)} \mathbf{E} f(X + e_w).$$

By these two equalities,

$$\begin{aligned}
\mathbf{E} \sum_{w \in \mathcal{W}'_L} \sum_{i=1}^{|w|} & X_w [f(X - e_w + e_{w^{(i)}}) - f(X)] \\
&= \sum_{w \in \mathcal{W}'_L} \sum_{i=1}^{|w|} \frac{1}{h(w)} \mathbf{E}\big[f(X + e_{w^{(i)}}) - f(X + e_w) \big] \\
&= \sum_{w \in \mathcal{W}'_{L-1}} \sum_{i=1}^{|w|} \frac{1}{h(w)} \mathbf{E} f(X + e_{w^{(i)}}) + \sum_{w \in \mathcal{W}_L / D_{2L}} \frac{|w|}{h(w)} \mathbf{E} f(X) \\
&\quad - \sum_{w \in \mathcal{W}'_L} \frac{|w|}{h(w)} \mathbf{E} f(X + e_w).
\end{aligned}$$

Specializing Lemma 4.11 to $q_w = \mathbf{E} f(X + e_w)$, the first sum is

$$\sum_{w \in \mathcal{W}'_{L-1}} \sum_{i=1}^{|w|} \frac{1}{h(w)} \mathbf{E} f(X + e_{w^{(i)}}) = \sum_{w \in \mathcal{W}'_L} \frac{c(w)}{h(w)} \mathbf{E} f(X + e_w),$$

which gives us

$$\begin{aligned}
\mathbf{E} \sum_{w \in \mathcal{W}'_L} \sum_{i=1}^{|w|} & X_w [f(X - e_w + e_{w^{(i)}}) - f(X)] \\
&= \sum_{w \in \mathcal{W}'_L} \frac{c(w) - |w|}{h(w)} \mathbf{E} f(X + e_w) + \sum_{w \in \mathcal{W}_L / D_{2L}} \frac{|w|}{h(w)} \mathbf{E} f(X).
\end{aligned}$$

All that remains in proving (4.4) is to show that

$$\sum_{w \in \mathcal{W}'_L} \frac{|w| - c(w)}{h(w)} = \sum_{w \in \mathcal{W}_L / D_{2L}} \frac{|w|}{h(w)}.$$

Specializing Lemma 4.11 to $q_w = 1$ shows that $\sum_{w \in \mathcal{W}'_L} c(w)/h(w) = \sum_{w \in \mathcal{W}'_{L-1}} |w|/h(w)$. Thus

$$\begin{aligned}
\sum_{w \in \mathcal{W}'_L} \frac{|w| - c(w)}{h(w)} &= \sum_{w \in \mathcal{W}'_L} \frac{|w|}{h(w)} - \sum_{w \in \mathcal{W}'_{L-1}} \frac{|w|}{h(w)} \\
&= \sum_{w \in \mathcal{W}_L / D_{2L}} \frac{|w|}{h(w)},
\end{aligned}$$

establishing (4.4) and completing the proof. $\square$

From now on, we will consider the process $(N_w(t),\, k \in \mathbb{N},\, t \geq 0)$ to be running under stationarity, i.e., with marginal distributions given by conclusion (iii) of the last lemma. This process is easily constructed as described above, but with additional point masses of weight $1/h(w)$ for each $w \in \mathcal{W}'$ at $(w, 0)$ added to the intensity measure of χ, thus giving us the correct distribution at time zero.

4.2.3 Time-reversed processes

Fix some time $T > 0$. We define the time-reversal $\overleftarrow{N}_w(t) \triangleq N_w(T-t)$ for $0 \leq t \leq T$.

Lemma 4.13. *For any fixed $L \in \mathbb{N}$, the process $\{(\overleftarrow{N}_w(t),\, w \in \mathcal{W}'_L),\, 0 \leq t \leq T\}$ is a time-homogenous Markov process with respect to the natural filtration. A trivial modification at jump times renders RCLL paths. The transition rates of this chain are given as follows. Let $u \in \mathcal{W}_{k-1}/D_{2k-1}$ and $w \in \mathcal{W}_k/D_{2k}$, and suppose that u can be obtained from w by halving b different pairs. Let $x = (x_w,\, w \in \mathcal{W}'_L)$.*

(i) The chain jumps from x to $x + e_u - e_w$ with rate bx_w.

(ii) The chain jumps from x to $x - e_w$ with rate $(k - c(w))x_w$.

(iii) If $w \in \mathcal{W}_L/D_{2L}$, then the chain jumps from x to $x + e_w$ with rate $L/h(w)$.

Proof. Any Markov process run backwards under stationarity is Markov. If the chain has transition rate $r(x, y)$ from states x to y, then the transition rate of the backwards chain from x to y is $r(y, x)\nu(y)/\nu(x)$, where ν is the stationary distribution. We will let ν be the stationary distribution from Lemma 4.12iii and calculate the transition rates of the backwards chain, using the rates given in Lemma 4.12ii.

Let a denote the number of letters in u that give w when doubled. The transition rate of the original chain from $x + e_u - e_w$ to x is $a(x_u + 1)$, so the transition rate of the backwards chain from x to $x + e_u - e_w$ is

$$a(x_u + 1)\frac{\nu(x + e_{k-1,c-1} - e_{k,c})}{\nu(x)} = \frac{ah(w)x_w}{h(u)},$$

and this is equal to bx_w by Lemma 4.9. A similar calculation shows that the transition rate from x to $x - e_w$ is

$$\frac{(k - c(w))\nu(x - e_w)}{h(w)\nu(x)} = (k - c(w))x_w,$$

proving (ii). The transition rate from x to $x + e_w$ for $w \in \mathcal{W}_L / D_{2L}$ is

$$\frac{\nu(x + e_w)}{\nu(x)}(x_w + 1)L = \frac{L}{h(w)},$$

which completes the proof. □

By definition,

$$\overleftarrow{N}_w(t) = \sum_{\substack{(u,s) \in \chi \\ s \le T-t}} 1\left\{X_{u,s}(T - t - s) = w\right\}.$$

We will modify this slightly to define the process

$$\overleftarrow{M}_w(t) \triangleq \sum_{\substack{(u,s) \in \chi \\ s \le T-t}} 1\left\{X_{u,s}(T - t - s) = w \text{ and } |X_{u,s}(T - s)| \le L\right\}.$$

The idea is that $\overleftarrow{M}_w(t)$ is the same as $\overleftarrow{N}_w(t)$, except that it does not count cycles at time t that had more than L vertices at time zero. The process $(\overleftarrow{M}_w(t),\, w \in \mathcal{W}'_L)$ is a Markov chain with the same transition rates as $(\overleftarrow{N}_w(t),\, w \in \mathcal{W}'_L)$, except that it does not jump from x to $x + e_w$ for $w \in \mathcal{W}_L / D_{2L}$. These two chains also have the same initial distribution, but $(\overleftarrow{M}_w(t),\, w \in \mathcal{W}'_L)$ is not stationary (in fact, it is eventually absorbed at zero).

4.3 Process convergence of the cycle structure

Recall that $C_k^{(s)}(t)$ is the number of cycles of length k in the graph $G(s + t)$, defined on p. 6. For $w \in \mathcal{W}'$, let $C_w^{(s)}(t)$ be the number of cycles in $G(s + t)$ with word w. We will prove that $\left(C_w^{(s)}(\cdot),\, w \in \mathcal{W}'\right)$ converges to a distributional limit, from which the convergence of $\left(C_k^{(s)}(\cdot),\, k \in \mathbb{N}\right)$ will follow. The proof depends on knowing the limiting *marginal* distribution of $C_w^{(s)}(t)$. The following corollary of Theorem 2.2 gives the facts we need:

Corollary 4.14. *Let $\{Z_w,\, w \in \mathcal{W}'_K\}$ be a family of independent Poisson random variables with $\mathbf{E}Z_w = 1/h(w)$. For any fixed integer K and $d \ge 1$,* *(i) as* $t \to \infty$,

$$(C_w(t),\, w \in \mathcal{W}'_K) \xrightarrow{\mathcal{L}} (Z_w,\, w \in \mathcal{W}'_K);$$

(ii) as $t \to \infty$, *the probability that there exist two cycles of length* K *or less sharing a vertex in* $G(t)$ *approaches zero.*

Proof. When $d = 1$, there is only one word of each length in $\mathcal{W}'_K$, and statement (i) reduces to the well-known fact that the cycle counts of a random permutation converge to independent Poisson random variables (see [AT92] for much more on this subject). In this case, $G(t)$ is made up of disjoint cycles for all times t, so that statement (ii) is trivially satisfied.

When $d \geq 2$, let $C_w^{(n)}$ be the number of cycles with word w in G_n. Observe that $C_w^{(n)} = \sum_\alpha I_\alpha$, with I_α as in the statement of Theorem 2.2 and the sum over all cycles in $\mathcal{I}$ with word w. The random variable Z_w is the analogous sum over Z_α, since the number of cycles in $\mathcal{I}$ with word w is $[n]_k/h(w)$. By Theorem 2.2,

$$(C_w^{(n)},\, w \in \mathcal{W}'_K) \xrightarrow{\mathcal{L}} (Z_w,\, w \in \mathcal{W}'_K). \tag{4.5}$$

Now, we just extend this to continuous time. The random vector $(C_w(t),\, w \in \mathcal{W}'_K)$ is a mixture of the random vectors $(C_w^{(n)},\, w \in \mathcal{W}'_K)$ over different values of n. That is,

$$\mathbf{P}\left[\left(C_w(t),\, w \in \mathcal{W}'_K\right) \in A\right] = \sum_{n=1}^{\infty} \mathbf{P}[M_t = n]\mathbf{P}\left[\left(C_w^{(n)},\, w \in \mathcal{W}'_K\right) \in A\right]$$

for any set A, recalling that $G(t) = G_{M_t}$. Equation (4.5) together with the fact that $\mathbf{P}[M_t > N] \to 1$ as $t \to \infty$ for any N imply that $(C_w(t),\, w \in \mathcal{W}'_K)$ converges in law to $(Z_w,\, w \in \mathcal{W}'_K)$, establishing statement (i).

The discrete time version of statement (ii) is given by [DJPP13, Corollary 16]. Statement (ii) follows from it in the same way. □

Now, we turn to the convergence of the processes. We will often need to transfer the convergence of a process to its limit to the convergence of a functional of the process. The following criterion, which we present without proof, lets us apply the continuous mapping theorem to do so.

Lemma 4.15 ([EK86, Section 3.11, Exercise 14]). *Let E and F be metric spaces, and let $f\colon E \to F$ be continuous. Then the mapping $x \mapsto f \circ x$ from $D_E[0,\infty) \to D_F[0,\infty)$ is continuous.*

Theorem 4.16. *The process $\big(C_w^{(s)}(\cdot),\, w \in \mathcal{W}'\big)$ converges in law as $s \to \infty$ to $(N_w(\cdot),\, w \in \mathcal{W}')$ in the space $D_{\mathbb{R}^\infty}[0,\infty)$.*

Proof. The main difficulty in turning the intuitive ideas of Section 4.2.1 into an actual proof is that $\big(C_w^{(s)}(t),\, w \in \mathcal{W}'\big)$ is not Markov. We now sketch how we evade this problem. We will run our chain backwards, defining $\overleftarrow{G}_s(t) = G(s+T-t)$ for some fixed $T > 0$. Then, we ignore all of $\overleftarrow{G}_s(0)$ except for the subgraph consisting of cycles of size L and smaller, which we will call $\overleftarrow{\Gamma}_s(0)$. The graph $\overleftarrow{\Gamma}_s(t)$ is the evolution of this subgraph as time runs backward, ignoring the rest of $\overleftarrow{G}_s(t)$. Then, we consider the number of cycles with word w in $\overleftarrow{\Gamma}_s(t)$, which we call $\phi_w(\overleftarrow{\Gamma}_s(t))$. Choose $K \ll L$. Then $\phi_w(\overleftarrow{\Gamma}_s(t))$ is likely to be the same as $C_w^{(s)}(T-t)$ for any word w with $|w| \leq K$. The remarkable fact that makes $\phi_w(\overleftarrow{\Gamma}_s(t))$ possible to analyze is that if $\overleftarrow{\Gamma}_s(0)$ consists of disjoint cycles, then $\big(\phi_w(\overleftarrow{\Gamma}_s(t)),\, w \in \mathcal{W}'_L\big)$ is a Markov chain governed by the same transition rates as $\big(\overleftarrow{M}_w(t),\, w \in \mathcal{W}'_L\big)$.

Another important idea of the proof is to ignore the vertex labels in $\overleftarrow{G}_s(t)$, so that we do not know in what order the vertices will be removed. Thus we can view $\overleftarrow{G}_s(t)$ as a Markov chain with the following description: Assign each vertex an independent Exp(1) clock. When the clock of vertex v goes off, remove it from the graph, and patch together the π_i-labeled edges entering and leaving v for each $1 \leq i \leq d$.

Step 1. *Definitions of $\overleftarrow{\Gamma}_s(t)$ and ϕ_w and analysis of $\big(\phi_w(\overleftarrow{\Gamma}_s(t)),\, w \in \mathcal{W}'_L\big)$.*

Fix $T > 0$ and define $\overleftarrow{G}_s(t) = G(s+T-t)$. As mentioned above, we will consider $\overleftarrow{G}_s(t)$ only up to relabeling of vertices, which makes it a process on the countable state space consisting of all edge-labeled graphs on finitely many unlabeled vertices. With respect to its natural filtration, it is a Markov chain in which each vertex is removed with rate one, as described above.

To formally define $\overleftarrow{\Gamma}_s(t)$, fix integers $L > K$ and let $\overleftarrow{\Gamma}_s(0)$ be the subgraph of $\overleftarrow{G}_s(0)$ made up of all cycles of length L or less. We then evolve $\overleftarrow{\Gamma}_s(t)$ in parallel with $\overleftarrow{G}_s(t)$. When a vertex v is deleted from $\overleftarrow{G}_s(t)$, the corresponding vertex v in $\overleftarrow{\Gamma}_s(t)$ is deleted if it is present. If v has a π_i-labeled edge entering and leaving it in $\overleftarrow{\Gamma}_s(t)$, then these two edges are patched together. Other edges in $\overleftarrow{\Gamma}_s(t)$ adjacent to v are deleted. This makes $\overleftarrow{\Gamma}_s(t)$ a subgraph of $\overleftarrow{G}_s(t)$, as well as a continuous-time Markov chain on the countable state space consisting of all edge-labeled graphs on finitely many unlabeled vertices. The transition probabilities of $\overleftarrow{\Gamma}_s(t)$ do not depend on s.

From Corollary 4.14, we can find the limiting distribution of $\overleftarrow{\Gamma}_s(0)$. Suppose that γ is a graph in the process's state space that is not a disjoint union of cycles. By Corollary 4.14ii,

$$\lim_{s\to\infty} \mathbf{P}[\overleftarrow{\Gamma}_s(0) = \gamma] = 0.$$

Suppose instead that γ is made up of disjoint cycles, with z_w cycles of word w for each $w \in \mathcal{W}'_L$. By Corollary 4.14i,

$$\lim_{s\to\infty} \mathbf{P}[\overleftarrow{\Gamma}_s(0) = \gamma] = \prod_{w\in\mathcal{W}'_L} \mathbf{P}[Z_w = z_w], \tag{4.6}$$

where $(Z_w,\, w \in \mathcal{W}'_L)$ are independent Poisson random variables with $\mathbf{E}Z_w = 1/h(w)$. Thus $\overleftarrow{\Gamma}_s(0)$ converges in law as $s \to \infty$ to a limiting distribution supported on the graphs made up of disjoint unions of cycles. For different values of s, the chains $\overleftarrow{\Gamma}_s(t)$ differ only in their initial distributions, and the convergence in law of $\overleftarrow{\Gamma}_s(0)$ as $s \to \infty$ induces the process convergence of $\{\overleftarrow{\Gamma}_s(t),\, 0 \le t \le T\}$ to a Markov chain $\{\overleftarrow{\Gamma}(t),\, 0 \le t \le T\}$ with the same transition rates whose initial distribution is the limit of $\overleftarrow{\Gamma}_s(0)$.

For any finite edge-labeled graph G, let $\phi_w(G)$ be the number of cycles in G with word w. By Lemma 4.15 and the continuous mapping theorem, the process $(\phi_w(\overleftarrow{\Gamma}_s(t)),\, w \in \mathcal{W}'_L)$ converges in law to $(\phi_w(\overleftarrow{\Gamma}(t)),\, w \in \mathcal{W}'_L)$ as $s \to \infty$.

We will now demonstrate that this process has the same law as $(\overleftarrow{M}_w(t),\, w \in \mathcal{W}'_L)$. The graph $\overleftarrow{\Gamma}(t)$ consists of disjoint cycles at time $t = 0$, and as it evolves, these cycles shrink or are destroyed. The process $(\phi_w(\overleftarrow{\Gamma}(t)),\, w \in \mathcal{W}'_L)$ jumps exactly when a vertex in a cycle in $\overleftarrow{\Gamma}(t)$ is deleted. If the deleted vertex lies in a cycle between two edges with the same label, the cycle shrinks. If the deleted vertex lies in a cycle between two edges with different labels, the cycle is destroyed. The only relevant consideration in where the process will jump at time t is the number of vertices of these two types in $\overleftarrow{\Gamma}(t)$, which can be deduced from $(\phi_w(\overleftarrow{\Gamma}(t)),\, w \in \mathcal{W}'_L)$. Thus this process is a Markov chain.

Consider two words $u, w \in \mathcal{W}'_K$ such that w can be obtained from u by doubling a letter. Suppose that u can be obtained from w by halving any of b pairs of letters. Suppose that the chain is at state $x = (x_v,\, v \in \mathcal{W}'_L)$. There are bx_w vertices that when deleted cause the chain to jump from x to $x - e_w + e_u$, each of which is removed with rate one. Thus the chain jumps from x to $x - e_w + e_u$ with rate bx_w. Similarly, it jumps to $x - e_w$ with rate $(|w| - c(w))x_w$. These are the same rates as the chain $(\overleftarrow{M}_w(t),\, w \in \mathcal{W}'_L)$ from Section 4.2.3. The initial distribution given by (4.6) is also the same as that of $(\overleftarrow{M}_w(t),\, w \in \mathcal{W}'_L)$, demonstrating that the two processes $\big(\phi_w(\overleftarrow{\Gamma}(t)),\, w \in \mathcal{W}'_L\big)$ and $(\overleftarrow{M}_w(t),\, w \in \mathcal{W}'_L)$ have the same law.

Step 2. *Approximation of $\overleftarrow{C}^{(s)}_w(t)$ by $\phi_w(\overleftarrow{\Gamma}_s(t))$.*

We will compare the two processes $\big\{(\overleftarrow{C}^{(s)}_w(t),\, w \in \mathcal{W}'_K),\, 0 \le t \le T\big\}$ and $\{(\phi_w(\overleftarrow{\Gamma}_s(t)),\, w \in \mathcal{W}'_K),\, 0 \le t \le T\}$ and show that for sufficiently large L, they are identical with probability arbitrarily close to one.

Consider some cycle in $\overleftarrow{G}_s(t)$; we can divide its vertices into those that lie between two edges of the cycle with different labels, and those that lie between two edges with the same label. We call this second class the *shrinking vertices* of the cycle, because if one is deleted from $\overleftarrow{G}_s(t)$ as it evolves, the cycle shrinks. We define $E_s(L)$ to be the event that for some cycle in $\overleftarrow{G}_s(0)$ of size $l > L$, at least $l - K$ of its shrinking vertices are deleted by time T.

We claim that outside of the event $E_s(L)$, the two processes $\{(\overleftarrow{C}^{(s)}_w(t),\, w \in \mathcal{W}'_K),\, 0 \le t \le T\}$ and $\{(\phi_w(\overleftarrow{\Gamma}_s(t)),\, w \in \mathcal{W}'_K),\, 0 \le t \le T\}$ are identical. Suppose that these two processes are not identical. Then there is some cycle α of size K or less present in $\overleftarrow{G}_s(t)$ but not in $\overleftarrow{\Gamma}_s(t)$ for $0 < t \le T$. As explained in Section 4.2, as a cycle evolves (in forward time), it grows into an overlapping cluster of cycles. Thus $\overleftarrow{G}_s(0)$ contains some cluster of overlapping cycles that shrinks to α at time t. One of the cycles in this cluster has length greater than L, or the cluster would be contained in $\overleftarrow{\Gamma}_s(0)$ and α would have been contained in $\overleftarrow{\Gamma}_s(t)$.

To see that $l - K$ shrinking vertices must be deleted from this cycle, consider the evolution of α into the cluster of cycles in both forward and reverse time. If a vertex is inserted into a single edge of a cycle in forward time, we see in reverse time the deletion of a shrinking vertex. If a vertex is simultaneously inserted into two edges of a cycle, causing the cycle to split, we see in reverse time the deletion of a non-shrinking vertex of a cycle. As α grows, a cycle of size greater than L can form only by single-insertion of at least $l-K$ vertices into the eventual cycle. In reverse time, this is seen as deletion of $l-K$ shrinking vertices. This demonstrates that $E_s(L)$ holds.

We will now show that for any $\epsilon > 0$, there is an L sufficiently large that $\mathbf{P}[E_s(L)] < \epsilon$ for any s. Let $w \in \mathcal{W}_l/D_{2l}$ with $l > L$, and let $I \subseteq [l]$ such that $|I| = l-K$ and $w_i = w_{i-1}$ for all $i \in I$, considering indices modulo l. For any cycle in $\overleftarrow{G}_s(0)$ with word l, the set I corresponds to a set of $l-K$ shrinking vertices of the cycle.

We define $F(w,I)$ to be the event that $\overleftarrow{G}_s(0)$ contains one or more cycles with word w, and that the vertices corresponding to I in one of these cycles are all deleted within time T. By a union bound,

$$\mathbf{P}[E_s(L)] \le \sum_{w,I} \mathbf{P}[F(w,I)]. \tag{4.7}$$

We proceed by enumerating all pairs of w and I. For any pair w, I, deleting the letters in w at positions given by I results in a word $u \in \mathcal{W}_K/D_{2K}$. For any given $u = u_1 \cdots u_K \in \mathcal{W}_K/D_{2K}$, the word $w \in \mathcal{W}_l/D_{2l}$ must have the form

$$w = \underbrace{u_1 \cdots u_1}_{a_1 \text{ times}} \underbrace{u_2 \cdots u_2}_{a_2 \text{ times}} \cdots\cdots \underbrace{u_K \cdots u_K}_{a_K \text{ times}},$$

with $a_i \ge 1$ and $a_1 + \cdots + a_K = l$. The number of choices for $a_1, \ldots, a_K$ is $\binom{l-1}{K-1}$, the number of compositions of l into K parts, and each of these corresponds to a choice of w and I. There are fewer than $a(d,K)$ choices for u, giving us a bound of $a(d,K)\binom{l-1}{K-1}$ choices of pairs w and I for any fixed $l > L$.

Next, we will show that for any pair w and I with $|w| = l$,

$$\mathbf{P}[F(w,I)] \le (1 - e^{-T})^{l-K}. \tag{4.8}$$

Condition on $\overleftarrow{G}_s(0)$ having n vertices. Consider any of the $[n]_l$ possible sequences of l vertices. Choose some representative $w' \in \mathcal{W}_l$ of w. For each of these sequences, the probability that it forms a cycle with word w' is at most $1/[n]_l$ (recall the original definition of our random graphs in terms of random permutations). Given that the sequence forms a cycle, the probability that the vertices of the cycle at positions I are all deleted within time T is $(1-e^{-T})^{l-K}$. Hence

$$\begin{aligned}\mathbf{P}\left[F(w,I)\mid \overleftarrow{G}_s(0)\text{ has }n\text{ vertices}\right] &\leq [n]_l\frac{1}{[n]_l}(1-e^{-T})^{l-K},\\ &\leq (1-e^{-T})^{l-K}.\end{aligned}$$

This holds for any n, establishing (4.8).

Applying all of this to (4.7),

$$\mathbf{P}[E_s(L)] \leq \sum_{l=L+1}^{\infty} a(d,K)\binom{l-1}{K-1}(1-e^{-T})^{l-k}.$$

This sum converges, which means that for any $\epsilon>0$, we have $\mathbf{P}[E_s(L)]<\epsilon$ for large enough L, independent of s.

Step 3. *Approximation of $\overleftarrow{N}_w(t)$ by $\overleftarrow{M}_w(t)$.*

Recall that we defined the processes $\{(\overleftarrow{M}_w(t),\, w\in\mathcal{W}'_K),\, 0\leq t\leq T\}$ and $\{(\overleftarrow{N}_w(t),\, w\in\mathcal{W}'_K),\, 0\leq t\leq T\}$ on the same probability space. We will show that for sufficiently large L, the two processes are identical with probability arbitrarily close to one.

By their definitions, these two processes are identical unless one of the processes $X_{u,s}(\cdot)$ started at each atom of χ grows from a word of size K or less to a word of size $L+1$ before time T; we call this event $E(L)$. Let

$$Y=\left|\left\{(u,s)\in\chi\colon |u|\leq K,\, s\leq T\right\}\right|,$$

the number of processes starting from a word of size K or less before time T.

Suppose that $X(\cdot)$ has law $\widetilde{P}_w$ for some word $w\in\mathcal{W}_k/D_{2k}$. We can choose L large enough that $\mathbf{P}\bigl[|X(T)|>L\bigr]<\epsilon$ for all $k\leq K$. Then $\mathbf{P}[E(L)\mid Y]<\epsilon Y$ by a union bound, and so $\mathbf{P}[E(L)]<\epsilon\mathbf{E}Y$. Since $\mathbf{E}Y<\infty$, we can make $\mathbf{P}[E(L)]$ arbitrarily small by choosing sufficiently large L.

Step 4. *Weak convergence of $\{(\overleftarrow{C}^{(s)}_w(t),\, w\in\mathcal{W}'_K),\, 0\leq t\leq T\}$ to $\{(\overleftarrow{N}_w(t),\, w\in\mathcal{W}'_K),\, 0\leq t\leq T\}$.*

If two processes are identical with probability $1-\epsilon$, then the total variation distance between their laws is at most ϵ. Thus, by steps 2 and 3, we can choose L large enough that the laws of the processes $\{(\overleftarrow{C}^{(s)}_w(t),\, w \in \mathcal{W}'_K),\, 0 \le t \le T\}$ and $\{(\phi_w(\overleftarrow{\Gamma}_s(t),\, w \in \mathcal{W}'_K),\, 0 \le t \le T)\}$ are arbitrarily close in total variation distance, uniformly in s, and so that the laws of $\{(\overleftarrow{M}_w(t),\, w \in \mathcal{W}'_K),\, 0 \le t \le T\}$ and $\{(\overleftarrow{N}_w(t),\, w \in \mathcal{W}'_K),\, 0 \le t \le T\}\}$ are arbitrarily close in total variation distance. Since total variation distance dominates the Prokhorov metric (or any other metric for the topology of weak convergence), we can choose L such that these two pairs are each within $\epsilon/3$ in the Prokhorov metric. Since $\{(\phi_w(\overleftarrow{\Gamma}_s(t)),\, w \in \mathcal{W}'_K),\, 0 \le t \le T\}$ converges in law to $\{(\overleftarrow{M}_w(t),\, w \in \mathcal{W}'_K),\, 0 \le t \le T\}$ as $s \to \infty$, there is an s_0 such that for all $s \ge s_0$, the laws of these processes are within $\epsilon/3$ in the Prokhorov metric. We have thus shown that for every $\epsilon > 0$, the laws of $\{(\overleftarrow{C}^{(s)}_w(t),\, w \in \mathcal{W}'_K),\, 0 \le t \le T\}$ and $\{(\overleftarrow{N}_w(t),\, w \in \mathcal{W}'_K),\, 0 \le t \le T\}$ are within ϵ for sufficiently large s, which proves that the first process converges in law to the second in the space $D_{\mathbb{R}^{|\mathcal{W}'_K|}}[0,T]$ as $s \to \infty$.

Step 5. *Weak convergence of* $\{(C^{(s)}_w(t),\, w \in \mathcal{W}'),\, t \ge 0\}$ *to* $\{(N_w(t),\, w \in \mathcal{W}'),\, t \ge 0\}$.

It follows immediately from the previous step that the (not time-reversed) process $\{(C^{(s)}_w(t),\, w \in \mathcal{W}'_K),\, 0 \le t \le T\}$ converges in law to $\{(N_w(t),\, w \in \mathcal{W}'_K),\, 0 \le t \le T\}$ for any $T > 0$. By Theorem 16.17 in [Bil99], $\{(C^{(s)}_w(t),\, w \in \mathcal{W}'_K),\, t \ge 0\}$ converges in law to $\{(N_w(t),\, w \in \mathcal{W}'_K),\, t \ge 0\}$. By [EK86, Section 3.11, Exercise 23], this also also proves that $\{(C^{(s)}_w(t),\, w \in \mathcal{W}'),\, t \ge 0\}$ converges in law to $\{(N_w(t),\, w \in \mathcal{W}'),\, t \ge 0\}$. □

Proof of Theorem 4.1. We will express the graph cycle counts as functionals of $\big(C^{(s)}_w(t),\, w \in \mathcal{W}'\big)$. The number of k-cycles in $G(s+t)$ is given by $C^{(s)}_k(t) = \sum_{w \in \mathcal{W}_k/D_{2k}} C^{(s)}_w(t)$. Let

$$N_k(t) = \sum_{w \in \mathcal{W}_k/D_{2k}} N_w(t).$$

By Lemma 4.15 and the continuous mapping theorem, $\{(C^{(s)}_k(t),\, k \in \mathbb{N}),\, t \ge 0\}$ converges in law to $\{(N_k(t),\, k \in \mathbb{N}),\, t \ge 0\}$ as $s \to \infty$.

It is not hard to see that this limit is Markov and admits the following representation: Cycles of size k appear spontaneously with rate $\sum_{w \in \mathcal{W}_k/D_{2k}} \mu(w)$. The size of each cycle then grows as a pure birth process with generator $Lf(i) = i\,(f(i+1) - f(i))$. The only thing we need to verify is that

$$\sum_{w \in \mathcal{W}_k/D_{2k}} \mu(w) = \sum_{w \in \mathcal{W}_k/D_{2k}} \frac{k - c(w)}{h(w)} = \frac{a(d,k) - a(d,k-1)}{2}. \tag{4.9}$$

This follows from Lemma 4.11 in the following way. From that lemma we get

$$\sum_{w\in\mathcal{W}_k/D_{2k}} \frac{c(w)}{h(w)} = (k-1) \sum_{w\in\mathcal{W}_{k-1}/D_{2(k-1)}} \frac{1}{h(w)}.$$

Thus

$$\sum_{w\in\mathcal{W}_k/D_{2k}} \mu(w) = \sum_{w\in\mathcal{W}_k/D_{2k}} \frac{k}{h(w)} - \sum_{w\in\mathcal{W}_{k-1}/D_{2(k-1)}} \frac{k-1}{h(w)}.$$

The two terms on the right side of the above equation are simply half the total number of cyclically reduced words possible, of size k and $k-1$ respectively. The total number of cyclically reduced words of size k on an alphabet of size d is by definition $a(d,k)$, showing (4.9) and completing the proof. □

So far, we have considered d as a constant. We now view it as a parameter of the graph and allow it to vary. Recall that $(\pi_d^{(n)},\, n\geq 1)$ are towers of random permutations independent for each d, and that $G(n,2d)$ is defined from $\pi_1^{(n)},\ldots,\pi_d^{(n)}$. For each d, we follow the construction used to define $G(t)$ and construct $G(t,2d)$, a continuous-time version of $(G(n,2d),\, n\in\mathbb{N})$. Let $\mathcal{W}'(d)$ be the set of equivalence classes of cyclically reduced words as before, with the parameter d made explicit. Define $C_{d,k}^{(s)}(t)$ as the number of k-cycles in $G(s+t,2d)$ and consider the convergence of the two-dimensional field $\{(C_{d,k}^{(s)}(t),\, d,k\in\mathbb{N}),\, t\geq 0\}$ as $s\to\infty$.

Again, we will consider this process as a functional of another one. Define $\mathcal{W}'(\infty) = \bigcup_{d=1}^{\infty}\mathcal{W}'(d)$, noting that $\mathcal{W}'(1)\subseteq\mathcal{W}'(2)\subseteq\cdots$. For any $w\in\mathcal{W}'(d)$, the number of cycles in $G(s+t,2d')$ with word w is the same for all $d'\geq d$. We define $C_w^{(s)}(t)$ by this, so that

$$C_{d,k}^{(s)}(t) = \sum_{\substack{w\in\mathcal{W}'(d)\\ |w|=k}} C_w^{(s)}(t).$$

Then we will prove convergence of $\{(C_w^{(s)}(t),\, w\in\mathcal{W}'(\infty)),\, t\geq 0\}$ as $s\to\infty$.

To define a limit for this process, we extend μ to a measure on all of $\mathcal{W}'(\infty)$ and define the Poisson point process χ on $\mathcal{W}'(\infty)\times[0,\infty)$. The rest of the construction is identical to the one in Section 4.2.2, giving us random variables $\big(N_w(t),\, w\in\mathcal{W}'(\infty)\big)$.

Theorem 4.17. *The process $\big(C_w^{(s)}(\cdot),\, w\in\mathcal{W}'(\infty)\big)$ converges in law as $s\to\infty$ to $\big(N_w(\cdot),\, w\in\mathcal{W}'(\infty)\big)$.*

Proof. For every d, we have shown in Theorem 4.16 that $\left(C_w^{(s)}(\cdot),\, w \in \mathcal{W}'(d)\right)$ converges in law as $s \to \infty$ to $(N_w(\cdot),\, w \in \mathcal{W}'(d))$. The rest of the proof then just amounts to the statement that weak convergence in $D_{\mathbb{R}^k}[0,\infty)$ for each k amounts to convergence in $D_{\mathbb{R}^\infty}[0,\infty)$, just as in the very end of the proof of Theorem 4.16. $\square$

Theorem 4.18. *There is a joint process convergence of* $(C_{i,k}^{(s)}(t),\ k \in \mathbb{N},\ i \in [d],\ t \geq 0)$ *to a limiting process* $(N_{i,k}(t),\ k \in \mathbb{N},\ i \in [d],\ t \geq 0)$*. This limit is a Markov process whose marginal law for every fixed d is described in Theorem 4.1. Moreover, for any* $d \in \mathbb{N}$*, the process* $(N_{d+1,k}(\cdot) - N_{d,k}(\cdot),\ k \in \mathbb{N})$ *is independent of the process* $(N_{i,k}(\cdot),\ k \in \mathbb{N},\ i \in [d])$ *and evolves as a Markov process. Its generator (defined on functions dependent on finitely many coordinates) is given by*

$$Lf(x) = \sum_{k=1}^{\infty} k x_k \left[f\left(x + e_{k+1} - e_k\right) - f(x) \right] + \sum_{k=1}^{\infty} \nu(d,k) \left[f(x + e_k) - f(x) \right],$$

where x *is a nonnegative sequence,* $(e_k, k \in \mathbb{N})$ *are the canonical orthonormal basis of* ℓ^2*, and*

$$\nu(d,k) = \frac{1}{2}\left[a(d+1,k) - a(d+1,k-1) - a(d,k) + a(d,k-1) \right].$$

Proof. Let

$$N_{d,k}(t) = \sum_{\substack{w \in \mathcal{W}'(d) \\ |w|=k}} N_w(t).$$

By Lemma 4.15, the continuous mapping theorem, and Theorem 4.17, $(N_{d,k}(\cdot),\, d,k \in \mathbb{N})$ is the limit of $(C_{d,k}^{(s)}(\cdot),\, d,k \in \mathbb{N})$ as $s \to \infty$.

Let us now describe what the limiting process is. It is obvious that $(N_{d,k}(\cdot),\ k \in \mathbb{N},\ d \in \mathbb{N})$ is jointly Markov. For every fixed d, the law of the corresponding marginal is given by Theorem 4.1. To understand the relationship across d, notice that cycles of size k in $G(t, 2(d+1))$ consist of cycles of size k in $G(t, 2d)$ and the extra cycles that contain an edge labeled by π_{d+1} or π_{d+1}^{-1}. Thus

$$N_{d+1,k}(t) - N_{d,k}(t) = \sum_{\substack{w \in \mathcal{W}'(d+1)\setminus\mathcal{W}'(d) \\ |w|=k}} N_w(t) \tag{4.10}$$

This process is independent of $(N_{i,\cdot},\ i \in [d])$, since the set of words involved are disjoint. Moreover, the rates for this process are clearly the following: cycles of size k grow at rate k and new cycles of size k appear at rate $[a(d+1,k) - a(d+1,k-1) - a(d,k) + a(d,k-1)]/2$. This completes the proof of the result. $\square$

4.4 Process limit for linear eigenvalue statistics

4.4.1 The limiting cycle structure

As in Section 3.2, we must transfer our results from cycles to cyclically non-backtracking walks. Call a cyclically non-backtracking walk *bad* if it is anything other than a repeated walk around a cycle.

Proposition 4.19. *Fix an integer K. There is a random time T, almost surely finite, such that there are no bad cyclically non-backtracking walks of length K or less in $G(t)$ for all $t \geq T$.*

Proof. We will work with the discrete-time version of our process $(G_n,\ n \in \mathbb{N})$. We first define some machinery introduced in [LP10]. Consider some cyclically non-backtracking walk of length k on the edge-labeled complete graph K_n of the form

$$s_0 \xrightarrow{w_1} s_1 \xrightarrow{w_2} s_2 \xrightarrow{w_3} \cdots \xrightarrow{w_k} s_k = s_0.$$

Here, $s_i \in [n]$ and $w = w_1 \cdots w_k$ is the word of the walk (that is, each w_i is π_j or π_j^{-1} for some j, indicating which permutation provided the edge for the walk). We say that G_n contains the walk if the random permutations $\pi_1, \ldots, \pi_d$ satisfy $w_i(s_{i-1}) = s_i$. In other words, G_n contains a walk if considering both as edge-labeled directed graphs, the walk is a subgraph of G_n.

If $(s'_i,\ 0 \leq i \leq k)$ is another walk with the same word, we say that the two walks are of the same *category* if $s_i = s_j \iff s'_i = s'_j$. In other words, two walks are of the same category if they are identical up to relabeling vertices. The probability that G_n contains a walk depends only on its category. If a walk contains e distinct edges, then G_n contains the walk with probability at most $1/[n]_e$.

Let $X_k^{(n)}$ be the number of bad walks of length k in G_n that start at vertex n. We will first prove that with probability one, $X_k^{(n)} > 0$ for only finitely many n. Call a category bad if the walks in the category are bad. Let $\mathcal{T}_{k,d}$ be the number of bad categories of walks of length k. For any particular bad category whose walks contain v distinct vertices, there are $[n-1]_{v-1}$ walks of that category whose first vertex is n. Any bad walk contains more edges than vertices, so

$$\mathbf{E}X_k^{(n)} \leq \frac{\mathcal{T}_{k,d}[n-1]_{v-1}}{[n]_{v+1}} \leq \frac{\mathcal{T}_{k,d}}{n(n-k)}.$$

Since $X_k^{(n)}$ takes values in the nonnegative integers, $\mathbf{P}[X_k^{(n)} > 0] \leq \mathbf{E}X_k^{(n)}$. By the Borel-Cantelli lemma, $X_k^{(n)} > 0$ for only finitely many values of n.

Thus, for any fixed $r+1$, there exists a random time N such that there are no bad walks on G_n of length $r+1$ or less starting with vertex n, for $n \geq N$. We claim that for $n \geq N$, there are no bad walks at all on G_n with length r or less. Suppose that G_m contains some bad walk of length $k \leq r$, for some $m \geq N$. As the graph evolves, it is easy to compute that with probability one, a new vertex is eventually inserted into an edge of this walk. But at the time $n > m \geq N$ when this occurs, G_n will contain a bad walk of length $r+1$ or less starting with vertex n, a contradiction. Thus we have proven that G_n eventually contains no bad walks of length r or less. The equivalent statement for the continuous-time version of the graph process follows easily from this. $\square$

Define

$$\begin{aligned}
\widehat{\Gamma}_0(x) &= 1, \\
\widehat{\Gamma}_{2k}(x) &= 2T_{2k}(x) + \frac{2d-2}{(2d-1)^k} && \text{for } k \geq 1, \\
\widehat{\Gamma}_{2k+1}(x) &= 2T_{2k+1}(x) && \text{for } k \geq 0.
\end{aligned}$$

Note that $\widehat{\Gamma}_i(x)$ is the same as $\Gamma_i(x)$ from Section 3.1, except that x and d are replaced by $2x$ and $2d$.

Definition 4.20. Let G be a $2d$-regular graph on n vertices. Let $f(x)$ be a polynomial expressed in the basis $\{\widehat{\Gamma}_i(x),\, i \geq 0\}$ as

$$f(x) = \sum_{j=0}^{k} a_j \widehat{\Gamma}_j(x).$$

We define $\operatorname{tr} f(G)$ as

$$\sum_{i=1}^{n} f(\lambda_i) - n a_0,$$

where $\lambda_1 \geq \cdots \geq \lambda_n$ are the eigenvalues of the adjacency matrix of G divided by $2\sqrt{2d-1}$.

Remark 4.21. The polynomial $f(x) - a_0$ is orthogonal to 1 with respect to the Kesten–McKay law (1.1), since $\widehat{\Gamma}_1(x), \widehat{\Gamma}_2(x), \ldots$ are orthogonal to 1 with respect to this measure. (To prove this, observe that each of these polynomials can be written in terms of the orthogonal polynomials of [Sod07, Example 5.3]. This is done in the proof of [DJPP13, Proposition 32].) This orthogonalization keeps $\operatorname{tr} f(G_n)$ of constant order when $n \to \infty$. One can calulate a_0 by integrating f against the Kesten-McKay law:

$$a_0 = \int_{-2}^{2} f(x) \frac{2d(2d-1)\sqrt{4-x^2}}{2\pi\bigl(4d^2 - (2d-1)x^2\bigr)}\, dx.$$

The most important set of functions for us will be the Chebyshev polynomials. For $T_k(x)$ with $k \geq 1$,

$$a_0 = \begin{cases} 0 & \text{if } k \text{ is odd,} \\ -\frac{d-1}{(2d-1)^{k/2}} & \text{if } k \text{ is even.} \end{cases}$$

Proof of Theorem 4.3. Let $\mathrm{CNBW}_k^{(s)}(t)$ denote the number of cyclically non-backtracking walks of length k in $G(s+t)$. We decompose these into those that are repeated walks around cycles of length j for some j dividing k, and the remaining bad walks, which we denote $B_k^{(s)}(t)$, giving us

$$\mathrm{CNBW}_k^{(s)}(t) = \sum_{j \mid k} 2j C_j^{(s)}(t) + B_k^{(s)}(t).$$

Proposition 4.19 implies that

$$\lim_{s \to \infty} \mathbf{P}\bigl[B_k^{(s)}(t) = 0 \text{ for all } k \leq K,\, t \geq 0\bigr] = 1.$$

This together with Lemma 4.15 and Theorem 4.1 shows that as s tends to infinity,

$$\big(\mathrm{CNBW}_k^{(s)}(\cdot),\, 1 \leq k \leq K\big) \xrightarrow{\mathcal{L}} \bigg(\sum_{j|k} 2jN_j(\cdot),\, 1 \leq k \leq K\bigg). \tag{4.11}$$

Now, we modify the polynomials $\widehat{\Gamma}_k$ to form a new basis $\{f_k,\, k \in \mathbb{N}\}$ with the right properties, which amounts to expressing each $N_k(t)$ as a linear combination of terms $\sum_{j|l} 2jN_j(t)$. We do this with the Möbius inversion formula. Define the polynomial

$$f_k(x) = \frac{1}{2k} \sum_{j|k} \mu\left(\frac{k}{j}\right) (2d-1)^{j/2} \widehat{\Gamma}_j(x), \tag{4.12}$$

where μ is the Möbius function, given by

$$\mu(n) = \begin{cases} (-1)^a & \text{if } n \text{ is the product of } a \text{ distinct primes,} \\ 0 & \text{otherwise.} \end{cases}$$

From Proposition 3.3, (4.11), and the continuous mapping theorem,

$$\big(\operatorname{tr} f_k(G(s+\cdot)),\, k \in [K]\big) \xrightarrow{\mathcal{L}} \big(N_k(\cdot),\, k \in [K]\big)$$

as desired.

For an arbitrary polynomial f, let $\hat{f}$ denote $f - a_0$, the orthogonalized version of f from Definition 4.20. The polynomial $\hat{f}$ is a linear combination of $f_1, f_2, \ldots$, and so the process $\operatorname{tr} f(G(s+\cdot))$ converges to a linear combination of the coordinate processes of $(N_k(\cdot),\, k \in \mathbb{N})$. □

4.4.2 Some properties of the limiting object

To prove the process convergence in Theorem 4.4 and Proposition 4.5, we need to know more about the limiting cycle process $(N_k(\cdot),\, k \in \mathbb{N})$. Though the limiting object is not defined in terms of graphs, we will nonetheless refer to $N_k(t)$ as the number of k-cycles at time t in the limiting object. Similarly, if one of the Yule processes counted to define the limiting object increases from j to k, we will refer to this as a cycle growing from size j to k.

We start our study of the limiting object by decomposing $N_k(t)$ into independent summands in terms of the process at time s. We first give a definition related to this decomposition.

Definition 4.22. Let the random variable $\alpha_{s,t}(j,k)$ be the portion of j-cycles at time s that grow to be k-cycles at time t in the limiting object. When s and t are clear from context, we will just write this as $\alpha(j,k)$.

Lemma 4.23. *For $j \le k$ and $s \le t$,*

$$\mathbf{E}\alpha_{s,t}(j,k) = \binom{k-1}{k-j} e^{j(s-t)}\bigl(1 - e^{s-t}\bigr)^{k-j}. \tag{4.13}$$

Proof. The quantity $\mathbf{E}\alpha_{s,t}(j,k)$ is the probability that a Yule process started from j is at k at time $t-s$. It is known that this is given by (4.13) (see [Lig10, Exercise 2.11], for example), but we will give a proof of it anyhow.

We start with the case that $j = 1$, and we assume $s = 0$. Let X_t be a Yule process from 1. We would like to show that

$$\mathbf{P}[X_t = k] = e^{-t}\bigl(1 - e^{-t}\bigr)^{k-1}, \tag{4.14}$$

or equivalently that $X_t - 1 \sim \mathrm{Geo}\bigl(e^{-t}\bigr)$. Let $S_1, S_2, \ldots$ be the holding times of the Yule process. By definition, they are independent, with $S_i \sim \mathrm{Exp}(i)$. Then

$$\mathbf{P}[X_t > k] = \mathbf{P}[S_1 + \cdots + S_k \le t].$$

Now, let $\tau_1, \ldots, \tau_k$ be i.i.d. with distribution Exp(1), and consider a counting process with these k points as its jump times. Then the first holding time is $\mathrm{Exp}(k)$, the next $\mathrm{Exp}(k-1)$, and so on. Thus

$$\mathbf{P}[S_1 + \cdots + S_k \le t] = \mathbf{P}[\tau_1, \ldots, \tau_k \le t] = \bigl(1 - e^{-t}\bigr)^k,$$

which shows that $X_t - 1 \sim \mathrm{Geo}\bigl(e^{-t}\bigr)$, confirming (4.14).

To extend this to $j > 1$, let Y_t be the sum of j independent Yule processes starting from 1. This makes Y_t a Yule process starting from j. The random variable $Y_t - j$ is a sum of independent $\mathrm{Geo}(e^{-t})$ random variables and thus is negative binomial, the distribution of the number of failures before j successes occur in independent Bernoulli trials with a success rate of e^{-t}. Consulting [Fel68, eq. VI.8.1] for a formula for this distribution,

$$\mathbf{P}[Y_t - j = k - j] = \binom{k-1}{j-1} e^{-jt}(1 - e^{-t})^{k-j},$$

which matches (4.13) after the substitution of $t - s$ for t. □

We now give our decomposition of $N_k(t)$:

Lemma 4.24. *Let $j \leq k$ and $s \leq t$. The random variable $N_k(t)$ can be decomposed into independent, Poisson-distributed summands as*

$$N_k(t) = \sum_{j=1}^{k} \alpha_{s,t}(j,k) N_j(s) + Z. \tag{4.15}$$

Proof. All k-cycles at time t are either j-cycles at time s that grow to size k, or they are spontaneously formed. The random variable $\alpha_{s,t}(j,k)N_j(s)$ is the number of j-cycles that grow to size k, and we define Z to be the number of cycles that form spontaneously at times in $(s,t]$ and have size k at time t. We then have (4.15), and we just need to to confirm that the summands are independent and Poisson. Cycles at time s grow independently of each other and of the spontaneously formed cycles, which confirms the independence. By Raikov's theorem on decompositions of the Poisson distribution into independent sums [Loè60, 19.2A], each summand is Poisson, completing the proof.

This last step is needlessly slick: The random variable $\alpha_{s,t}(j,k)N_j(s)$ is a thinned version of a Poisson random variable and hence Poisson itself. A similar argument applies to Z. $\square$

Next, we compute the covariance structure of our limiting object.

Proposition 4.25. *For any $s \leq t$ and $j, k \in \mathbb{N}$,*

$$\mathbf{Cov}(N_k(t), N_j(s)) = \begin{cases} \frac{a(d,j)}{2j}\binom{k-1}{k-j} e^{j(s-t)}\left(1 - e^{s-t}\right)^{k-j} & \text{if } j \leq k, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Suppose $j > k$. As $\{N_i(s),\, i \geq 1\}$ are independent, the decomposition (4.15) shows that $N_k(t)$ is independent of $N_j(s)$. Intuitively, cycles of size greater than k at time s do not affect the cycles of size k at time t.

When $j \leq k$, the result follows immediately from Lemmas 4.23 and 4.24 by decomposing $N_k(t)$ and taking expectations. $\square$

4.4.3 The process convergence

Lemma 4.26. *Let ξ be a Poisson random measure on $[0, \infty)$ with arbitrary σ-finite intensity measure. Let $T_1, T_2, \ldots$ denote the atoms of ξ. Let $(\tau_i,\, i \geq 1)$ be arbitrary nonnegative i.i.d.*

random variables. Form a new point process ζ with atoms $T_i + \tau_i$. Then ζ is also a Poisson random measure on $[0, \infty)$.

Proof. Let μ be the intensity measure of ξ, and let P be the distribution of τ_i. We have made ξ into a marked point process, giving each atom T_i an independent mark τ_i. This is equivalent to defining $\{(T_i, \tau_i),\, i \geq 1\}$ to be the atoms of a Poisson random measure on $[0, \infty)^2$ with intensity measure $\mu \otimes P$ [Çın11, Corollary VI.3.5]. The point process ζ is a deterministic transformation of this one by the map $(x, y) \mapsto x + y$, and is hence also a Poisson random measure [Çın11, Remark VI.2.4b]. □

The following technical lemma will be used in both Theorem 4.4 and Proposition 4.5.

Lemma 4.27. *Fix k and T, and consider $\big\{(2d-1)^{-k/2}\big(2kN_k(\cdot) - a(d,k)\big),\, d \geq 1\big\}$, a collection of processes in $D[0,T]$ indexed by d. This collection is tight.*

Proof. Fix d, and define $Y_k(t)$ as the process that starts at 0 and increases at each point of increase of $N_k(t)$; define $Z_k(t)$ as the process that starts at 0 and increases at each point of decrease of $N_k(t)$. Thus, we have $N_k(t) - N_k(0) = Y_k(t) - Z_k(t)$. As $N_k(t)$ almost surely jumps only by 1 and -1, both $Y_k(t)$ and $Z_k(t)$ are counting processes. Observe that $Y_k(t)$ counts k-cycles formed spontaneously or by growth in the time interval $(0, t]$, and $Z_k(t)$ counts k-cycles that jump to size $k+1$ in the time interval $(0, t]$.

Claim 4.28. The processes $Y_k(t)$ and $Z_k(t)$ are (non-independent) Poisson processes with rate $a(d,k)/2$.

Proof. We argue by induction on k. As our base case, the process $Y_1(t)$ jumps when 1-cycles form spontaneously, which happen according to a Poisson process of rate $a(d,1)/2$. Now, assume that $Y_k(t)$ is a Poisson process of rate $a(d,k)/2$. First, we argue that $Z_k(t)$ is as well. Let ξ be the Poisson point process whose atoms are the points of increase of $Y_k(t)$, with an extra $N_k(0)$ atoms at 0. Each atom T_i of ξ is the time that a k-cycle forms (or 0 if it was present from the start). Let τ_i be the amount of time after T_i that the corresponding k-cycle jumps to $k+1$. Then $(\tau_i,\, i \geq 1)$ are i.i.d., and $T_i + \tau_i$ are the jump times of $Z_k(t)$. By Lemma 4.26, $Z_k(t)$ is a (possibly inhomogeneous) Poisson process. By the stationarity of the limiting object, $\mathbf{E}N_k(t) = \mathbf{E}N_k(0)$, and hence $\mathbf{E}Z_k(t) = \mathbf{E}Y_k(t) = a(d,k)/2$, showing that $Z_k(t)$ is a homogeneous Poisson process with rate $a(d,k)/2$.

To complete the induction, we must show that $Y_{k+1}(t)$ is a Poisson process of rate $a(d,k+1)/2$. To see this, observe that $Z_k(t)$ counts all $(k+1)$-cycles that form by growth in the time interval $(0,t]$. As $Y_{k+1}(t)$ counts all $(k+1)$-cycles that form by growth or spontaneously in that interval, it is the sum of $Z_k(t)$ and an independent Poisson process of rate $\big(a(d,k+1) - a(d,k)\big)/2$. Thus it is a Poisson process of rate $a(d,k+1)/2$. $\square$

Now, fix k and let $X_d(t) = (2d-1)^{-k/2}\big(2kN_k(t) - a(d,k)\big)$. We need to show that $\{X_d,\, d \geq 1\}$ is tight. As $N_k(t) - N_k(0) = Y_k(t) - Z_k(t)$, we have

$$\frac{1}{2k}X_d(t) = A_d + B_d(t) - C_d(t)$$

where

$$\begin{aligned} A_d &= (2d-1)^{-k/2}\left(N_k(0) - \frac{a(d,k)}{2k}\right), \\ B_d(t) &= (2d-1)^{-k/2}\left(Y_k(t) - \frac{a(d,k)t}{2}\right), \\ C_d(t) &= (2d-1)^{-k/2}\left(Z_k(t) - \frac{a(d,k)t}{2}\right). \end{aligned}$$

with $Y_k(t)$ and $Z_k(t)$ implicitly depending on d.

As $d \to \infty$, the random variable A_d converges in law to Gaussian, and $B_d(t)$ and $C_d(t)$ converge in law to Brownian motion. Viewing A_d, $B_d(t)$, and $C_d(t)$ as elements of $D[0,T]$, each thus converges weakly to a limit in $C[0,T]$. As tightness in a product space is equivalent to tightness of the marginals, the sequence (A_d, B_d, C_d) is tight in $D^3[0,t]$, with all weak limit points lying in $C^3[0,t]$.

Given a subsequence of $\{X_d(\cdot)\}$, choose a further subsequence $\{X_{d_i}(\cdot)\}$ such that $(A_{d_i}, B_{d_i}, C_{d_i})$ converges. The map

$$(x(t), y(t), z(t)) \mapsto x(t) + y(t) - z(t)$$

is not in general continuous from $D^3[0,T] \to D[0,T]$, but it is continuous at $C^3[0,T]$. (This holds because Skorokhod convergence to a continuous function implies uniform convergence.)

By the continuous mapping theorem, $A_{d_i} + B_{d_i} - C_{d_i}$ has a weak limit. Thus we have shown that every subsequence of $\{X_d(\cdot)\}$ has a subsequence with a weak limit. □

Proof of Theorem 4.4. By Proposition 3.3 and (4.11),

$$2 \operatorname{tr} T_k \left(G(\infty + t)\right) = (2d-1)^{-k/2} \sum_{j|k} 2jN_j(t). \tag{4.16}$$

Now, we will prove finite-dimensional convergence to the stated Ornstein-Uhlenbeck process. Fix $K \in \mathbb{N}$ and a sequence of times $t_1 < \cdots < t_n$. We first show that the random vector

$$\Big((2d-1)^{-k/2}\big(N_k(t_i) - \mathbf{E}N_k(t_i)\big),\ k \in [K],\ i \in [n]\Big) \tag{4.17}$$

converges to a multivariate Gaussian, using a slight extension of the decomposition from Lemma 4.24. Let $\mathcal{S}$ be the set of sequences $s_1, \ldots, s_n$ with $s_i \in \{\delta\} \cup \mathbb{N}$ that satisfy a certain set of conditions. Each sequence will represent the history of a growing cycle, with s_i the size of the cycle at time t_i. The symbol δ will mean "not yet born." Thus, a sequence is in $\mathcal{S}$ if it consists of zero or more δs followed by a nondecreasing sequence of positive integers. We do not include the sequence of all δs in $\mathcal{S}$.

Let $S = (s_1, \ldots, s_n) \in \mathcal{S}$ and suppose that s_i is the first non-δ in the sequence. When $i = 1$, define X_S as the number of cycles that have size s_j at time t_j for all $1 \le j \le n$. If $i > 1$, define X_S as the number of cycles that form spontaneously between times t_{i-1} and t_i and have size s_j at time t_j for $j \ge i$.

We claim that $\{X_S,\, S \in \mathcal{S}\}$ is a collection of independent Poisson random variables. The number of cycles of each size at time t_1 and the number of cycles of each size at time t_i that formed after time t_{i-1} for all $2 \le i \le n$ are independent Poissons. Each of these random variables is then thinned to form $\{X_S,\, S \in \mathcal{S}\}$, which thus consists of independent Poissons as well.

Now, we will write (4.17) in terms of this Poisson field. First, let $\varphi(S)$ denote the first non-δ character in S, and consider the normalized field

$$\Big\{(2d-1)^{-\varphi(S)/2}(X_S - \mathbf{E}X_S),\ S \in \mathcal{S}\Big\}. \tag{4.18}$$

Fix some $S = (s_1, \ldots, s_n) \in \mathcal{S}$ with $s_i = \varphi(S)$ the first non-δ character. The expected number of cycles that form spontaneously between times t_{i-1} and t_i with size $\varphi(S)$ at time t_i is $O\big((2d-1)^{\varphi(S)}\big)$ (here, we are interpreting all elements of the big-O expression as constants except for d). The portion of these that grow according to S is in expectation a fixed fraction of these, with no dependence on d. Thus $\mathbf{E}X_S = O\big((2d-1)^{\varphi(S)}\big)$. By the Gaussian approximation to Poisson, the field (4.18) converges as $d \to \infty$ to independent Gaussians.

For each $k \in [K]$ and $i \in [n]$, we have

$$(2d-1)^{-k/2}\big(N_k(t_i) - \mathbf{E}N_k(t_i)\big) = \sum_S (2d-1)^{-k/2}(X_S - \mathbf{E}X_S),$$

where the sum ranges over all $S = (s_1, \ldots, s_n) \in \mathcal{S}$ with $s_i = k$. Every term of the sum with $\varphi(S) < k$ vanishes in probability, and the terms with $\varphi(S) = k$ are elements of the field (4.18). By the Gaussian convergence of (4.18), the random vector (4.17) converges to Gaussian as $d \to \infty$.

Now, consider a finite-dimensional slice of the process

$$\left(\operatorname{tr} T_k\left(G(\infty+t)\right) - \mathbf{E}\operatorname{tr} T_k\left(G(\infty+t)\right),\ t \geq 0,\ k \in \mathbb{N}\right), \tag{4.19}$$

choosing finitely many choices of k and t and forming a random vector. Each component has the form given by (4.16) for some k and t. The scaling causes all the terms of the sum there with $j < k$ to vanish in probability. Subtracting off these terms, we have a random vector whose components are a subset of those of (4.17). Thus the finite-dimensional distributions of (4.19) converge to Gaussian as $d \to \infty$.

Next, we compute the covariances. For a fixed d, from (4.16) we have

$$\mathbf{Cov}\left(\operatorname{tr} T_i\left(G(\infty+t)\right), \operatorname{tr} T_j\left(G(\infty+s)\right)\right) = \frac{1}{4}\left(2d-1\right)^{-(i+j)/2} \sum_{k|i,\ l|j} 4lk\,\mathbf{Cov}\left(N_k(t), N_l(s)\right) \tag{4.20}$$

for $s \leq t$. We now fix any i, j, t, s and take d to infinity, using the following expression from Proposition 4.25:

$$\mathbf{Cov}(N_k(t), N_l(s)) = \begin{cases} \frac{a(d,l)}{2l}\binom{k-1}{k-l}p^l(1-p)^{k-l}, & p = e^{s-t}, \quad \text{if } k \geq l. \\ 0, \quad \text{otherwise.} \end{cases}$$

Any term $a(d,r)$ is asymptotically the same as $(2d-1)^r$. Thus the highest order term in d on the right side of (4.20) is $(2d-1)^{\min(i,j)}$. Unless $i=j$, this term is negligible compared to $(2d-1)^{(i+j)/2}$. This shows that the limiting covariance is zero unless $i=j$. On the other hand, when $i=j$, every term on the right side of (4.20) vanishes, except when $k=i=l=j$. Hence,

$$\lim_{d\to\infty} \mathbf{Cov}\left(\operatorname{tr} T_i\left(G(\infty+t)\right), \operatorname{tr} T_i\left(G(\infty+s)\right)\right) = \frac{1}{4}2ip^i = \frac{i}{2}e^{i(s-t)}.$$

Thus we have shown convergence of the finite-dimensional distributions to the limiting process.

To show the process convergence, we appeal to Lemma 4.27. This lemma shows that all but the highest term of the sum in (4.16) vanishes in probability, and the remainder is a tight sequence in d. This immediately gives the convergence of

$$\left(\operatorname{tr} T_k\left(G(\infty+t)\right) - \mathbf{E}\operatorname{tr} T_k\left(G(\infty+t)\right),\ t\geq 0,\ k\in\mathbb{N}\right)$$

to the limiting process not in $D_{\mathbb{R}^\infty}[0,\infty)$, but in $D^\infty[0,\infty)$. As the limit lies in $C^\infty[0,\infty)$, an argument as in the end of Lemma 4.27 shows that the convergence holds in $D_{\mathbb{R}^\infty}[0,\infty)$ as well. □

4.4.4 Diagonal convergence

We now consider eigenvalue statistics where d increases with the size of the graph. One approach would be to give a quantitative version of Theorem 4.16 that would hold even as d grew, possibly with some conditions on its growth. We have opted for something much simpler, choosing d to grow however slowly is necessary to make the convergence still hold. The point here is more to explain what Theorem 4.4 has to do with the GFF than to study the graph process with d growing.

Proof of Proposition 4.5. Fix $K\in\mathbb{N}$ and $T>0$, and let

$$\Theta_d^{(s)}(t) = \Big(\operatorname{tr} T_k(G(s+t,2d)) - \mathbf{E}\operatorname{tr} T_k(G(\infty+t,2d)),\, 1\leq k\leq K\Big).$$

Considering this as a random element of $D_{\mathbb{R}^K}[0,T]$, Theorem 4.1 shows that with d held fixed, $\Theta_d^{(s)}(\cdot)$ converges weakly to a limit $\Theta_d^{(\infty)}(\cdot)$ described by (4.16). Theorem 4.4 then shows that $\Theta_d^{(\infty)}(\cdot)$ converges weakly to a collection of independent Ornstein-Uhlenbeck processes as $d \to \infty$. To take a diagonal limit, we simply take d to grow slowly enough that we can almost consider it as fixed. The argument will be highly technical but with little more than formal content.

Let ρ be a metric for the topology of weak convergence for probability measures on $D_{\mathbb{R}^K}[0,T]$, and use $\rho(X,Y)$ as a shorthand for the distance in this metric between the laws of X and Y. Recall the processes $C_{d,k}^{(s)}(t)$ and $N_{d,k}(t)$ from Theorem 4.18. Also recall that $B_k^{(s)}(t)$ is the number of bad cyclically non-backtracking walks of length k in $G(s+t, 2d)$, and introduce the notation $B_{d,k}^{(s)}(t)$ to indicate the dependence on d. For each d, choose s_d large enough that for all $s \geq s_d$,

$$\rho\Big(\Theta_{d+1}^{(s)},\, \Theta_{d+1}^{(\infty)}\Big) < \frac{1}{d}, \tag{4.21}$$

$$\rho\Big(\big(C_{d,k}^{(s)},\, C_{d+1,k}^{(s)}\big)_{k=1}^K,\, \big(N_{d,k},\, N_{d+1,k}\big)_{k=1}^K\Big) < \frac{1}{d}, \tag{4.22}$$

$$\mathbf{P}\Big[B_{d+1,k}^{(s)}(t) > 0 \text{ for any } k \leq K,\, 0 \leq t \leq T\Big] < \frac{1}{d}, \tag{4.23}$$

and for all $1 \leq k \leq K$,

$$\Big|\mathbf{E}\operatorname{tr} T_k\big(G(\lfloor e^{s/2}\rfloor, 2d(s+t))\big) - \mathbf{E}\operatorname{tr} T_k\big(G(\infty+t, 2d(s+t))\big)\Big| < \frac{1}{d}. \tag{4.24}$$

It is possible to find s_d satisfying (4.21)–(4.23) by Theorems 4.1 and 4.18 and Proposition 4.19, respectively. For (4.24), we clarify that $G(\lfloor e^{s/2}\rfloor, 2d(s+t))$ refers to the discrete-time graph defined in Section 1.4.3. For any fixed d, one can check by a combinatorial calculation that $\mathbf{E}\operatorname{tr} T_k\big(G(n, 2d)\big)$ converges as $n \to \infty$ to $\mathbf{E}\operatorname{tr} T_k\big(G(\infty+t, 2d)\big)$, which establishes that one can choose s_d to satisfy (4.24). We can take s_d and $s_{d+1} - s_d$ to be increasing sequences in d by choosing larger values for s_d if necessary. Define $d(s)$ to be the right-continuous function with $d(s) = 1$ that jumps from $i-1$ to i at s_i.

Our first goal is to show that $\Theta_{d(s+t)}^{(s)}(t)$ converges to the limiting Ornstein-Uhlenbeck processes as $d \to \infty$. From (4.21) and Theorem 4.4, we know that $\Theta_{d(s)}^{(s)}(t)$ converges to this limit. Thus it suffices to show that the distance between $\big(\Theta_{d(s+t)}^{(s)}(t),\, 0 \leq t \leq T\big)$ and $\big(\Theta_{d(s)}^{(s)}(t),\, 0 \leq t \leq T\big)$ in $D_{\mathbb{R}^K}[0,T]$ vanishes in probability as $s \to \infty$.

Consider the kth component of

$$\big(\Theta^{(s)}_{d(s+t)}(t),\, 0 \le t \le T\big) - \big(\Theta^{(s)}_{d(s)}(t),\, 0 \le t \le T\big) \tag{4.25}$$

at time t, which by Proposition 3.3 is equal to

$$\begin{aligned} \frac{1}{2}\big(2d(s+t)-1\big)^{-k/2}\big(\mathrm{CNBW}^{(s)}_{d(s+t),k}(t) - \mathbf{E}\mathrm{CNBW}^{(s)}_{d(s+t),k}(t)\big)& \\ -\frac{1}{2}\big(2d(s)-1\big)^{-k/2}\big(\mathrm{CNBW}^{(s)}_{d(s),k}(t) - \mathbf{E}\mathrm{CNBW}^{(s)}_{d(s),k}(t)\big),& \end{aligned} \tag{4.26}$$

with $\mathrm{CNBW}^{(s)}_{d,k}(t)$ denoting the number of cyclically non-backtracking walks in $G(s+t, 2d)$. We will show that this vanishes in probability as $s \to \infty$. For sufficiently large s and $0 \le t \le T$, we have either $d(s+t) = d(s)$ or $d(s+t) = d(s)+1$. In the first case, (4.26) is 0, so it suffices to show that

$$\begin{aligned} \frac{1}{2}\big(2d(s)+1\big)^{-k/2}\big(\mathrm{CNBW}^{(s)}_{d(s)+1,k}(t) - \mathbf{E}\mathrm{CNBW}^{(s)}_{d(s)+1,k}(t)\big)& \\ -\frac{1}{2}\big(2d(s)-1\big)^{-k/2}\big(\mathrm{CNBW}^{(s)}_{d(s),k}(t) - \mathbf{E}\mathrm{CNBW}^{(s)}_{d(s),k}(t)\big)& \end{aligned}$$

vanishes in probability. By (4.23), the difference between this expression and

$$\begin{aligned} \frac{1}{2}\big(2d(s)+1\big)^{-k/2}\sum_{j\mid k}\Big(2jC^{(s)}_{d(s)+1,j}(t) - 2j\mathbf{E}C^{(s)}_{d(s)+1,j}(t)\Big)& \\ -\frac{1}{2}\big(2d(s)-1\big)^{-k/2}\sum_{j\mid k}\Big(2jC^{(s)}_{d(s),j}(t) - 2j\mathbf{E}C^{(s)}_{d(s),j}(t)\Big)& \end{aligned}$$

converges to 0 in probability as $s \to \infty$. The scaling makes all terms of the sums besides $j = k$ vanish in probability. Thus it sufficies to show that

$$\begin{aligned} k\big(2d(s)+1\big)^{-k/2}\Big(C^{(s)}_{d(s)+1,k}(t) - \mathbf{E}C^{(s)}_{d(s)+1,k}(t)\Big)& \\ -k\big(2d(s)-1\big)^{-k/2}\Big(C^{(s)}_{d(s),k}(t) - \mathbf{E}C^{(s)}_{d(s),k}(t)\Big)& \end{aligned}$$

vanishes in probability. By (4.22), it suffices to show this for

$$\big(2d(s)+1\big)^{-k/2}\Big(N_{d(s)+1,k}(t) - \mathbf{E}N_{d(s)+1,k}(t)\Big) - \big(2d(s)-1\big)^{-k/2}\Big(N_{d(s),k}(t) - \mathbf{E}N_{d(s),k}(t)\Big).$$

By observing that the second moment of $\big((2d(s)+1)^{-k/2} - (2d(s)-1)^{-k/2}\big)\big(N_{d(s)+1,k}(t) - \mathbf{E}N_{d(s)+1,k}(t)\big)$ vanishes, it sufficies to show this for

$$\big(2d(s)-1\big)^{-k/2}\Big(N_{d(s)+1,k}(t) - N_{d(s),k}(t) - \mathbf{E}\big[N_{d(s)+1,k}(t) - N_{d(s),k}(t)\big]\Big). \tag{4.27}$$

By (4.10), the random variable $N_{d(s)+1,k}(t) - N_{d(s),k}(t)$ is distributed as $\mathrm{Poi}\big((a(d+1) - a(d))/2k\big)$, and the second moment of (4.27) vanishes. Thus we have shown that for any k and t, the expression (4.26) converges to 0 in probability. From (4.27), we also see that each component of (4.25) is tight. It follows from this that supremum norm of each component of (4.25) on $[0,T]$ converges to 0 in probability. This then shows that $\Theta^{(s)}_{d(s+t)}(t)$ converges to the same weak limit as $\Theta^{(s)}_{d(s)}(t)$.

The next step is showing that

$$\Big(\operatorname{tr} T_k\big(G(s+t, 2d(s+t))\big) - \mathbf{E}\big[\operatorname{tr} T_k\big(G(s+t, 2d(s+t))\big) \,\big|\, N(t)\big],\ 1 \le k \le K\Big)$$

converges to the same weak limit in $D_{\mathbb{R}^K}[0,T]$ as $\Theta^{(s)}_{d(s+t)}(t)$. The difference between the kth component of these two processes is

$$\mathbf{E}\big[\operatorname{tr} T_k\big(G(s+t, 2d(s+t))\big) \,\big|\, N(t)\big] - \mathbf{E} \operatorname{tr} T_k(G(\infty + t, 2d)),$$

and we would like to show that this vanishes in probability in the supremum norm as $s \to \infty$. By (4.24), it suffices to show that as $t \to \infty$,

$$\mathbf{P}[N(t) < e^{t/2}] \to 0. \tag{4.28}$$

By definition of our continuous-time process, $N(t) + 1$ is a Yule process starting from 2. It is well known that $(N(t)+1)e^{-t} \to Z$ a.s., where $Z \sim \mathrm{Exp}(1)$, which establishes (4.28). (To prove this, show that $(N(t)+1)e^{-t} \xrightarrow{\mathcal{L}} Z$ by a direct calculation, and then observe that if Y_t is a Yule process, then $Y_t e^{-t}$ is a positive martingale and hence converges a.s.)

The weak convergence of the process

$$\Big(\operatorname{tr} T_k\big(G(s+t, 2d(s+t))\big) - \mathbf{E}\big[\operatorname{tr} T_k\big(G(s+t, 2d(s+t))\big) \,\big|\, N(t)\big],\ k \in \mathbb{N},\, t \ge 0\Big)$$

in $D_{\mathbb{R}^K}[0,T]$ for arbitrary K and T gives the desired convergence in $D_{\mathbb{R}^\infty}[0,\infty)$ by the same argument as at the end of the proof of Theorem 4.1. □

4.5 Convergence to the Gaussian free field

The Gaussian free field is a generalization of Brownian motion where the indexing set has dimension greater than one. Physicists have long been interested in the GFF because of its importance in quantum field theory. Mathematicians have come to the GFF more recently, as it it became clear that it was the limit of a variety of discrete random surfaces and height functions [NS97, GOS01, Ken01, RV07, Ken08, BF14, JLS14, Bor10a, Kua11, Dui13, Pet12] and was closely related to Schramm-Loewner evolution [Dub09, SS09, SS13, MS12a, MS12b, MS12c, MS13].

At its most basic level, the GFF on the upper half-plane with zero Dirichlet boundary conditions can be thought of as a centered Gaussian field $(h(z),\, z \in \mathbb{H})$ with covariances given by

$$\mathbf{E}\big[h(z)h(w)\big] = -\frac{1}{2\pi}\log\left|\frac{z-w}{z-\overline{w}}\right|.$$

The problem with this definition is that no such random function h exists. If it did exist, then the collection of random variables $\int_{\mathbb{H}} f(z)h(z)\,dz$ indexed by smooth compactly supported functions f would also be a Gaussian field. This field does truly exist, and we will use it to define the GFF.

We start by giving a bare-bones treatment of the GFF that gives only the very few properties we need. After this, we give a more languorous account based on [She07], [HMP10], and [Dub09].

4.5.1 Bare-bones background on the Gaussian free field

Let h denote the GFF on $\mathbb{H}$ (with zero Dirichlet boundary conditions, the only kind we will consider). The only property we use in this thesis is that if $f(z)$ is a smooth function defined on a smooth path γ satisfying (4.29), one can define a collection of random variables denoted $\int_\gamma f(z)h(z)\,dz$ that form a centered Gaussian field. (Again, h is not really a function, and we are not really integrating against it. The notation is from [Bor10a], [BG13], and other papers. In Section 4.5.2, we explain the real definitions.) The covariances are given by the following proposition:

Proposition 4.29 ([BG13, Lemma 4.6]). *Let f_1, f_2 be smooth functions defined on the image of a smooth curve γ such that*

$$\int_\gamma\int_\gamma f_i(z)\left(-\frac{1}{2\pi}\log\left|\frac{z-w}{z-\overline{w}}\right|\right)f_i(w)\,dz\,dw < \infty \tag{4.29}$$

for $i = 1, 2$. *Then*

$$\mathbf{E}\Big[\Big(\int_\gamma f_1(z)h(z)\,dz\Big)\Big(\int_\gamma f_2(z)h(z)\,dz\Big)\Big] = \int_\gamma\int_\gamma f_1(z)\Big(-\frac{1}{2\pi}\log\Big|\frac{z-w}{z-\overline{w}}\Big|\Big)f_2(w)\,dz\,dw.$$

4.5.2 More background on the Gaussian free field

We will build up the GFF from scratch, mostly following [She07] with a sprinkling of [HMP10] and [Dub09]. Our goal will be to present it in as simply as possible and explain how it meshes with the more concrete information from the previous section. To make this account friendlier without bogging it down too much, we present background material on partial differential equations and Sobolev spaces in italics. For a proper introduction, see [Eva10], [Hun], and [Bre11].

Definition and construction of the Gaussian free field

Let $D \subseteq \mathbb{R}^d$ be a domain (that is, a connected open set). We define $H_s(D)$ as the space of all smooth, compactly supported, real-valued functions on D, and we endow this space with the *Dirichlet inner product*, given by $\langle f, g\rangle_\nabla = \int_D \nabla f(x)\cdot\nabla g(x)\,dx$. When $d = 2$, this inner product is conformally invariant, meaning that $\langle f\circ\varphi, g\circ\varphi\rangle_\nabla = \langle f, g\rangle_\nabla$ for any conformal map φ. We denote the Hilbert space closure of $H_s(D)$ by $H(D)$. When D is bounded, $H(D)$ is the subspace $H_0^1(D)$ of the Sobolev space $H^1(D) = W^{1,2}(D)$.

The Sobolev space $H^1(D)$ is a Hilbert space consisting of all functions in $L^2(D)$ whose (weak or distributional) first-order derivatives are also in L^2. When D is bounded, the Dirichlet inner product on $H_s(D)$ gives a norm equivalent to the standard one in $H^1(D)$ by the Poincaré inequality [Eva10, Section 5.6.1, Theorem 3]. The Hilbert space completion of $H_s(D)$ is then the closure of $C_c^\infty(D)$ in $H^1(D)$, with an inner product equivalent to the usual Sobolev one. This closure is denoted as $H_0^1(D)$, and it consists of the elements of $H^1(D)$ that are zero on the boundary in the sense of traces [Eva10, Section 5.5].

When D is unbounded, the situation is slightly messier, but we need to address it so that we can talk about the GFF on regions like the upper half-plane. To take advantage of the conformal invariance of the Dirichlet inner product, we will assume that D is an unbounded domain in $\mathbb{R}^2$ that admits a conformal map φ onto a bounded domain D'. The space $H^1_{loc}(D)$ consists of all functions on D whose restrictions belong to $H^1(U)$ for all open sets U with compact closure *in D. A sequence converges in $H^1_{loc}(D)$ if its restrictions converge in $H^1(U)$ for all such U, which makes this a Fréchet space. We will show that $H(D) \subseteq H^1_{loc}(D)$.*

Suppose that f_n forms a Cauchy sequence in $H_s(D)$. Then $f_n \circ \varphi^{-1}$ is a Cauchy sequence in $H_s(D')$, and it converges to a limit $g \in H^1_0(D')$. Let $f = g \circ \varphi$. By the local invariance of Sobolev spaces under smooth coordinate changes, $f \in H^1_{loc}(D)$ and $f_n \to f$ in that space [Fol95, Theorem 6.24, Corollary 6.25]. By conformal invariance, $f_n \to f$ in the Dirichlet inner product. Thus $H(D) \subseteq H^1_{loc}(D)$. In particular, elements of $H(D)$ are locally L^2-integrable.

Note that by integration by parts, the Dirichlet inner product on $H_s(D)$ can be expressed in terms of the usual inner product in L^2 by

$$\langle f, g\rangle_\nabla = \langle f, -\Delta g\rangle. \tag{4.30}$$

Suppose we have a probability space $(\Omega, \mathcal{F}, P)$. A closed subspace of $L^2(\Omega, \mathcal{F}, P)$ consisting of centered Gaussian random variables is called a *Gaussian Hilbert space.* We will assume throughout that $\mathcal{F}$ is the σ-algebra generated by these random variables. A trivial example of a Gaussian Hilbert space is the one-dimensional space $\{t\xi,\, t \in \mathbb{R}\}$, where ξ is a centered Gaussian. A non-trivial one is the closed linear span of the collection of random variables $\{B_t,\, t \geq 0\}$, where B_t is a standard Brownian motion. The definition and both examples can be found in much more detail in [Jan97].

We are now ready to define the GFF, though it will take some work afterwards to make sense of it. In the following definition, h has no meaning on its own. For each $f \in H(D)$, the notation $\langle h, f\rangle_\nabla$ indicates a random variable, with no assumptions at all on the map $f \mapsto \langle h, f\rangle$.

Definition 4.30. The Gaussian free field on a domain D (with zero Dirichlet boundary conditions) is the Gaussian Hilbert space of random variables $\big\{\langle h, f\rangle_\nabla,\, f \in H(D)\big\}$ with covariances given by

$$\mathbf{E}\big[\langle h, f\rangle_\nabla \langle h, g\rangle_\nabla\big] = \langle f, g\rangle_\nabla. \tag{4.31}$$

The notation $\langle h, f\rangle_\nabla$ suggests that the map $f \mapsto \langle h, f\rangle_\nabla$ should be linear, and this definition implies that it is: By applying (4.31), we can show that the variance of $\langle h, af + bg\rangle_\nabla - \big(a\langle h, f\rangle_\nabla + b\langle h, g\rangle_\nabla\big)$ is zero.

By the monotone class lemma, the law of $\{\langle h,f\rangle_\nabla,\ f\in H(D)\}$ is determined by the finite-dimensional distributions; see [Jan97, Example A.3]. This is where we use the assumption that the σ-algebra associated with a Gaussian Hilbert space is the smallest one that makes $\langle h,f\rangle_\nabla$ measurable for all $f\in H(D)$. Thus the definition determines at most one family $\{\langle h,f\rangle_\nabla,\ f\in H(D)\}$ in law. It is not clear, however, that there even exists such a Gaussian Hilbert space at all. We resolve this by constructing one:

Proposition 4.31. *There exists a Gaussian Hilbert space satisfying Definition* 4.30.

Proof. Let $\{f_i,\ i\in\mathbb{N}\}$ be an ordered orthonormal basis for $H(D)$ (this space is separable and hence has a countable orthonormal basis). Let $\{\alpha_i,\ i\in\mathbb{N}\}$ be independent standard Gaussians. For any $f\in H(D)$ with expansion $f=\sum\beta_i f_i$, we define

$$\langle h,f\rangle_\nabla=\lim_{k\to\infty}\sum_{i=1}^{k}\beta_i\alpha_i. \tag{4.32}$$

The sum is a martingale bounded in L^2 by Parseval's equality and hence converges a.s. and in L^2. Note that it was necessary to fix an order for the sum, as the sequence need not be absolutely summable. Thus we have constructed a Gaussian field $\{\langle h,f\rangle_\nabla,\ f\in H(D)\}$. If $f=\sum\beta_i f_i$ and $g=\sum\gamma_i f_i$, then it follows from the L^2 convergence of (4.32) that

$$\lim_{n\to\infty}\mathbf{E}\left[\left(\sum_{i=1}^{n}\beta_i\alpha_i\right)\left(\sum_{i=1}^{n}\gamma_i\alpha_i\right)\right]=\mathbf{E}\big[\langle h,f\rangle_\nabla\langle h,g\rangle_\nabla\big]$$

Thus

$$\mathbf{E}\big[\langle h,f\rangle_\nabla\langle h,g\rangle_\nabla\big]=\sum_{i=1}^{\infty}\beta_i\gamma_i=\langle f,g\rangle_\nabla$$

as desired. □

An example

We have defined and constructed the GFF without developing much of an intuition for it. We show now that the GFF on $D=(0,\infty)$ is Brownian motion. More precisely, let B_t be a standard Brownian motion and define $\langle h,f\rangle=\int_0^\infty f(t)B_t\,dt$ for $f\in H_s(D)$. Then define

$\langle h, f\rangle_\nabla = -\langle h, f''\rangle$ in analogy with (4.30). We confirm that this (or rather, its extension to all $f \in H(D)$) is the GFF according to Definition 4.30. For $f, g \in H_s(D)$,

$$\begin{aligned}
\mathbf{E}\big[\langle h, f\rangle_\nabla \langle h, g\rangle_\nabla\big] &= \mathbf{E}\int_0^\infty \int_0^\infty B_t f''(t) B_u g''(u)\, du\, dt \\
&= \int_0^\infty \int_0^\infty f''(t) g''(u) \min(u,t)\, du\, dt \\
&= \int_0^\infty \biggl(f''(t) \int_0^t u g''(u)\, du + t f''(t) \int_t^\infty g''(u)\, du \biggr) dt \\
&= \int_0^\infty \Bigl(f''(t)\bigl(t g'(t) - g(t)\bigr) - f''(t) t g'(t) \Bigr) dt \\
&= -\int_0^\infty f''(t) g(t) = \langle f, g\rangle_\nabla .
\end{aligned}$$

Green's functions and an alternate form of the GFF

The GFF can be written in an alternate form inspired by (4.30). Let $H(D)^*$ denote the dual space of $H(D)$, considered as a space of distributions, and denote the action of $f \in H(D)^*$ on $g \in H(D)$ by $\langle f, g\rangle$.

When D is bounded and hence $H(D) = H_0^1(D)$, the space $H(D)^*$ has a well-known characterization. Though Hilbert spaces are self-dual, we can instead view the dual space of $H_0^1(D)$ as a space of distributions. Viewed in this way, the dual space is denoted $H^{-1}(D)$. It consists of all sums of L^2-functions (viewed as distributions) and first-order distributional derivatives of L^2-functions [Bre11, Proposition 9.20]. When $f \in H^{-1}(D) \cap L^2(D)$, the distributional action of f coincides with the L^2 inner product; that is, for $\phi \in H_0^1(D)$, we have $\langle f, \phi\rangle = \int_D f\phi$.

Definition 4.32 (The GFF indexed by $H(D)^*$)**.** Let $f \in H(D)^*$. By the self-duality of Hilbert spaces, there exists $u \in H(D)$ such that $\langle f, \phi\rangle = \langle u, \phi\rangle_\nabla$ for all $\phi \in H(D)$. We define $\langle h, f\rangle = \langle h, u\rangle_\nabla$.

The significance of this definition is as follows. Suppose $f \in C_c^\infty(D)$, and we view it as an element of $H(D)^*$. Then the function $u \in H(D)$ associated with it solves the partial differential equation $-\Delta u = f$, and we have

$$\langle h, -\Delta u\rangle = \langle h, u\rangle_\nabla,$$

as in (4.30).

This version of the GFF also lends some insight on why the GFF in dimensions two and higher cannot be represented as a random function. Dirac δ-measures are elements of $H^{-1}(D)$ when $d = 1$ but not when $d \geq 2$. Thus it makes sense to evaluate h at a single point x by $\langle h, \delta_x \rangle$ only in the one-dimensional case.

Remark 4.33. The GFF can also be constructed as a random element of $H^{-\epsilon}(D)$ for any $\epsilon > 0$; see [HMP10, p. 7] and [She07, Proposition 2.7, Remark 2.8] for more details. The basic idea is to take $\{f_i\}$ and $\{\alpha_i\}$ as in Proposition 4.31 and define

$$h = \sum_{i=1}^{\infty} \alpha_i f_i,$$

which converges a.s. in $H^{-\epsilon}(D)$. This defines $\langle h, f \rangle$ for $f \in C_c^\infty(D)$ and coincides with our definition of $\langle h, f \rangle$.

The covariances of the Gaussian field $\{\langle h, f \rangle,\ f \in H(D)^*\}$ have a nice expression in terms of the *Green's function* for the Laplacian operator on D.

The Green's function $G(x, y)$ for the operator $-\Delta$ on a region D with Dirichlet boundary conditions is a solution to $-\Delta G(x, \cdot) = \delta_x$ (in the distributional sense) that satisfies $G(x, y) = 0$ if $x \in \partial D$ or $y \in \partial D$. The Green's function in general exists and is unique when D is bounded with C^1 boundary. The Green's function for the upper half-plane also exists and can be given explicitly:

$$G(x, y) = -\frac{1}{2\pi} \log \left| \frac{x - y}{x - \overline{y}} \right|,$$

thinking of x and y as complex. If $f \in H_s(D)$, then $u(x) = \int_D G(x, y) f(y)\, dy$ is in $H(D)$ and satisfies $-\Delta u = f$. The equivalent statement holds for $u(x) = \int_D G(x, y) \mu(dy)$ if $\mu \in H(D)^*$ is a locally finite measure with compact support in D. See [Fol95, Chapter 2] for a reference on Green's functions and related ideas.

Let G be the Green's function for $-\Delta$ on D with Dirichlet boundary conditions, and let $\Delta^{-1} f(x) \triangleq -\int_d G(x, y) f(y)\, dy$. For $f, g \in H_s(D)$,

$$\begin{aligned} \mathbf{E}\big[\langle h, f \rangle \langle h, g \rangle\big] &= \mathbf{E}\big[\langle h, -\Delta^{-1} f \rangle_\nabla \langle h, -\Delta^{-1} g \rangle_\nabla\big] \\ &= \langle -\Delta^{-1} f, -\Delta^{-1} g \rangle_\nabla \\ &= \langle f, -\Delta^{-1} g \rangle = \int_D \int_D f(x) G(x, y) g(y)\, dy\, dx. \end{aligned} \tag{4.33}$$

Similarly, if $\mu, \nu \in H(D)^*$ are locally finite, compactly supported measures, then

$$\mathbf{E}\big[\langle h, \mu\rangle\langle h, \nu\rangle\big] = \int_D \int_D G(x,y)\mu(dx)\nu(dy). \tag{4.34}$$

Traces

In this section, we explain how to define $\langle h, \mu\rangle$ when μ is a measure supported on a curve γ in $\overline{D}$, which along with (4.34) explains Proposition 4.29. Suppose that γ is a simple closed curve in $\overline{D}$, and suppose it forms the boundary of an open set E and is locally a graph of a Lipschitz function. Suppose that μ is supported on γ and bounded with respect to the natural measure there. Precisely, let $\mathcal{H}$ denote 1-dimensional Hausdorff measure and suppose that $\mu = \rho\, d\mathcal{H}$ for a bounded function ρ. Our goal is to define $\langle h, \mu\rangle$ by showing that $\mu \in H(D)^*$.

Lemma 4.34. *If $D \subseteq \mathbb{R}^2$ is bounded, or it is unbounded and its complement contains an open set, then the functional $f \mapsto \int f\, d\mu$ for $f \in H_s(D)$ extends to an element of $H(D)^*$.*

Proof. First, suppose that D is bounded. It suffices to show that $f \mapsto \int f\, d\mu$ is a bounded linear functional with respect to the Sobolev norm, since this is equivalent to the one given by the Dirichlet inner product. The restriction map $H_0^1(D) \to H^1(E)$ is obviously linear and bounded. By the Sobolev trace theorem [EG92, Theorem 4.3.1], there is a bounded trace operator $T\colon H^1(E) \to L^2(d\mathcal{H})$ such that $Tf = f|_\gamma$ when f is continuous. Thus for $f \in H_s(D)$, we have

$$\left|\int f\, d\mu\right| \le \int |Tf|\|\rho\|_\infty\, d\mathcal{H} \le \|\rho\|_\infty \|Tf\|_{L^2(d\mathcal{H})}\mathcal{H}(\gamma)^{1/2} \le C\|f\|_{H_0^1(D)}.$$

Thus $f \mapsto \int f\, d\mu$ is bounded and admits a unique extension to all $f \in H(D)$.

Now, suppose that $D \subseteq \mathbb{R}^2$ is unbounded. We will identify $\mathbb{R}^2$ with $\mathbb{C}$. Suppose that there is a neighborhood of $z_0 \in \mathbb{C}$ disjoint from D. Consider the conformal map $\varphi(z) = 1/(z - z_0)$, and let $D' = \varphi(D)$, a bounded set. The pushforward measure $\mu' = \mu \circ \varphi$ is supported on $\varphi(\gamma)$, and it has a bounded density with respect to $\mathcal{H}$. By the previous paragraph, for some C and any $f \in H_s(D)$ we have

$$\left|\int_D f\, d\mu\right| = \left|\int_{D'} f \circ \varphi\, d\mu'\right| \le C\big\|f \circ \varphi\big\|_\nabla = C\|f\|_\nabla$$

by the conformal invariance of $\|\cdot\|_\nabla$. Thus $f \mapsto \int f\, d\mu$ extends to a bounded linear functional on $H(D)$. $\square$

Identifying μ with its associated element of $H(D)^*$, we have justified the existence of $\langle h, \mu\rangle$. This is the random variable denoted by $\int_\gamma \rho(z)h(z)\,dz$ in Proposition 4.29. Together with (4.34), this explains Proposition 4.29.

4.5.3 Convergence of fluctuation process to the Gaussian free field

Recall that $F_t(x)$ counts the eigenvalues of $G(t, 2d(t))$ that are less than or equal to $2\sqrt{2d(t)-1}x$ and that

$$\overline{F}_t(x) = F_t(x) - \mathbf{E}[F_t(x) \mid N(t)].$$

Our goal is to show that $\overline{F}_{s+t}(x)$, considered as a function is x and t, converges in some sense to the Gaussian free field. First, we show that integrals against $\overline{F}_t(x)$ can be expressed in terms of traces. As usual, $T_k(x)$ and $U_k(x)$ denote the Chebyshev polynomials of order k on $[-1, 1]$ of the first and second kind, respectively.

Lemma 4.35.

$$\int_{-\infty}^{\infty} U_{k-1}(x)\overline{F}_t(x)\,dx = -\frac{1}{k}\Big(\operatorname{tr} T_k\big(G(t,2d(t))\big) - \mathbf{E}\big[\operatorname{tr} T_k\big(G(t,2d(t))\big) \bigm| N(t)\big]\Big).$$

Proof. As $x \to \pm\infty$, we have $\overline{F}_t(x) \to 0$ almost surely. Integrating by parts and using the relation $T_k'(x) = kU_{k-1}(x)$,

$$\begin{aligned}\int_{-\infty}^{\infty} U_{k-1}(x)\overline{F}_t(x)\,dx &= -\frac{1}{k}\int_{-\infty}^{\infty} T_k(x)\,d\overline{F}_t(x)\\ &= -\frac{1}{k}\sum_{i=1}^{N(t)} T_k(\lambda_i) + \frac{1}{k}\mathbf{E}\left[\sum_{i=1}^{N(t)} T_k(\lambda_i) \,\middle|\, N(t)\right],\end{aligned}$$

where $\lambda_1 \geq \cdots \geq \lambda_{N(t)}$ are the eigenvalues of $G(t)$ divided by $2\sqrt{2d(t)-1}$. This is equal to

$$\int_{-\infty}^{\infty} U_{k-1}(x)\overline{F}_t(x)\,dx = -\frac{1}{k}\Big(\operatorname{tr} T_k\big(G(t,2d(t))\big) - \mathbf{E}\big[\operatorname{tr} T_k\big(G(t,2d(t))\big) \bigm| N(t)\big]\Big).$$

Note that when k is even, the na_0 term introduced by the trace (see Definition 4.20) is cancelled by the same term in the expectation. $\square$

Combining this lemma with Proposition 4.5, integrals of the form $\int p(x)\overline{F}_{s+t}(x)\,dx$ converge jointly as $s \to \infty$ to a Gaussian field indexed by t and by polynomials $p(x)$. We now express this field in terms of the GFF.

Proof of Theorem 4.6. Proposition 4.5 and Lemma 4.35 prove that the integrals

$$\int_{-\infty}^{\infty} p_i(x) H_s(x, t_i)\,dx, \qquad i = 1, \dots, n$$

converge jointly to a centered multivariate normal distribution, which is also the distribution of the integrals against the GFF. We just need to check that the covariances match up. It suffices to confirm this on a polynomial basis. By Proposition 4.5,

$$\lim_{s\to\infty} \mathbf{E}\biggl[\biggl(\int_{-\infty}^{\infty} U_{j-1}(x) H_s(x,t_0)\,dx\biggr)\biggl(\int_{-\infty}^{\infty} U_{k-1}(x) H_s(x,t_1)\,dx\biggr)\biggr] = \delta_{jk}\frac{\pi}{4k} e^{k(t_0-t_1)} \quad (4.35)$$

for $t_0 \le t_1$. By Proposition 4.29, the covariance of

$$\int_{-1}^{1} U_{j-1}(x) h(\Omega(x,t_0))\,dx \quad \text{and} \quad \int_{-1}^{1} U_{k-1}(x) h(\Omega(x,t_1))\,dx$$

is

$$I \triangleq -\frac{1}{2\pi}\int_{-1}^{1}\int_{-1}^{1} U_{j-1}(x) \log\left|\frac{\Omega(x,t_0)-\Omega(y,t_1)}{\Omega(x,t_0)-\overline{\Omega}(y,t_1)}\right| U_{k-1}(y)\,dx\,dy.$$

Substituting $x = \cos u$ and $y = \cos v$, we have

$$I = -\frac{1}{2\pi}\int_0^{\pi}\int_0^{\pi} U_{j-1}(\cos u)\sin u \log\left|\frac{e^{t_0+iu}-e^{t_1+iv}}{e^{t_0+iu}-e^{t_1-iv}}\right| U_{k-1}(\cos v)\sin v\,du\,dv. \quad (4.36)$$

Assume that $t_0 < t_1$. For any constant $w \in \mathbb{C}$ with $|w| = t_1$, we can define functions $\log(z-w)$ and $\log(z-\overline{w})$ that are analytic on $|z| < t_1$. For each v, we choose two such logarithm functions with $w = e^{t_1+iv}$ to get

$$\log\left|\frac{e^{t_0+iu}-e^{t_1+iv}}{e^{t_0+iu}-e^{t_1-iv}}\right| = \frac{1}{2}\Bigl(\log\bigl(e^{t_0+iu}-e^{t_1+iv}\bigr) + \log\bigl(e^{t_0-iu}-e^{t_1-iv}\bigr) - \log\bigl(e^{t_0+iu}-e^{t_1-iv}\bigr) - \log\bigl(e^{t_0-iu}-e^{t_1+iv}\bigr)\Bigr).$$

Using the relation $U_{n-1}(\cos x) = \sin(nx)/\sin x$, we then have

$$I = -\frac{1}{4\pi}\int_0^{\pi}\int_0^{2\pi} \sin(ju)\sin(kv)\Bigl(\log\bigl(e^{t_0+iu}-e^{t_1+iv}\bigr) - \log\bigl(e^{t_0+iu}-e^{t_1-iv}\bigr)\Bigr)du\,dv,$$

and by integrating by parts in u,

$$\begin{aligned} I &= -\frac{1}{4j\pi}\int_0^{\pi}\int_0^{2\pi}\cos(ju)\sin(kv)\left(\frac{ie^{t_0+iu}}{e^{t_0+iu}-e^{t_1+iv}} - \frac{ie^{t_0+iu}}{e^{t_0+iu}-e^{t_1-iv}}\right)du\,dv \\ &= -\frac{1}{4j\pi}\int_0^{2\pi}\int_0^{2\pi}\cos(ju)\sin(kv)\frac{ie^{t_0+iu}}{e^{t_0+iu}-e^{t_1+iv}}\,du\,dv. \end{aligned}$$

We then integrate by parts in v to get

$$I = \frac{1}{4jk\pi}\int_0^{2\pi}\int_0^{2\pi}\cos(ju)\cos(kv)\frac{e^{t_0+t_1+i(u+v)}}{\left(e^{t_0+iu}-e^{t_1+iv}\right)^2}\,du\,dv.$$

Let γ denote a counterclockwise path around the unit disc.

$$I = \frac{1}{4jk\pi}\int_0^{2\pi}\cos(kv)\int_\gamma \frac{z^j+z^{-j}}{2iz}\frac{e^{t_0+t_1+iv}z}{\left(e^{t_0}z-e^{t_1+iv}\right)^2}dz\,dv.$$

The integrand of the path integral has a single pole in the unit disc at 0, and the residue there is $je^{j(t_0-t_1-iv)}/2i$. This gives

$$\begin{aligned} I &= \frac{1}{4k}\int_0^{2\pi}\cos(kv)e^{j(t_0-t_1-iv)}\,dv \\ &= \frac{1}{4k}\int_\gamma \frac{w^k+w^{-k)}}{2iw}e^{j(t_0-t_1)}w^{-j}\,dw \\ &= \frac{e^{j(t_0-t_1)}}{8ik}\int_\gamma \left(w^{k-j-1}+w^{-k-j-3}\right)dw. \end{aligned}$$

By computing residues, this is $\pi e^{j(t_0-t_1)}/4k$ if $j=k$ and 0 otherwise, agreeing with (4.35) for all $t_0 < t_1$. To extend this to $t_0 = t_1$ by a limiting argument, we apply the dominated convergence theorem to the integral in (4.36). One can show that

$$\left|\log\left|\frac{e^{t_0+iu}-e^{t_1+iv}}{e^{t_0+iu}-e^{t_1-iv}}\right|\right| \le \log\left|\frac{e^{iu}-e^{-iv}}{e^{iu}-e^{iv}}\right|$$

for all $t_0 \le t_1$. The right-hand side of this equation is integrable over $0 \le u, v \le \pi$. The other factors of the integrand in (4.36) are bounded there. Thus by the dominated convergence theorem we can compute I when $t_0 = t_1$ by letting $t_0 \to t_1$ from below. $\square$

BIBLIOGRAPHY

[AGG89] Richard Arratia, Larry Goldstein, and Louis Gordon. Two moments suffice for Poisson approximations: the Chen-Stein method. *Ann. Probab.*, 17(1):9–25, 1989.

[AGK13] Richard Arratia, Larry Goldstein, and Fred Kochman. Size bias for one and all. Preprint. Available at arXiv:1308.2729, 2013.

[ANvM12] Mark Adler, Eric Nordenstam, and Pierre van Moerbeke. The Dyson Brownian minor process. Preprint. Available at arXiv:1006.2956, 2012.

[AT92] Richard Arratia and Simon Tavaré. The cycle structure of random permutations. *Ann. Probab.*, 20(3):1567–1591, 1992.

[Bar01] Yu. Baryshnikov. GUEs and queues. *Probab. Theory Related Fields*, 119(2):256–274, 2001.

[BF14] Alexei Borodin and Patrik L. Ferrari. Anisotropic growth of random surfaces in $2+1$ dimensions. *Comm. Math. Phys.*, 325(2):603–684, 2014.

[BFSU99] Andrei Z. Broder, Alan M. Frieze, Stephen Suen, and Eli Upfal. Optimal construction of edge-disjoint paths in random graphs. *SIAM J. Comput.*, 28(2):541–573 (electronic), 1999.

[BG13] Alexei Borodin and Vadim Gorin. General beta Jacobi corners process and the Gaussian Free Field. Preprint. Available at arXiv:1305.3627, 2013.

[BHJ92] A. D. Barbour, Lars Holst, and Svante Janson. *Poisson approximation*, volume 2 of *Oxford Studies in Probability*. The Clarendon Press Oxford University Press, New York, 1992. Oxford Science Publications.

[Bil99] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[Bol80] Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, 1(4):311–316, 1980.

[Bor10a] Alexei Borodin. CLT for spectra of submatrices of Wigner random matrices. Preprint. Available at arXiv:1010.0898, 2010.

[Bor10b] Alexei Borodin. CLT for spectra of submatrices of Wigner random matrices II. Stochastic evolution. Preprint. Available at arXiv:1011.3544, 2010.

[Bre11] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.

[BS87] Andrei Broder and Eli Shamir. On the second eigenvalue of random regular graphs. In *28th Annual Symposium on Foundations of Computer Science (Los Angeles, 1987)*, pages 286–294. IEEE Comput. Soc. Press, Washington, D.C., 1987.

[BSK09] Sonny Ben-Shimon and Michael Krivelevich. Random regular graphs of non-constant degree: concentration of the chromatic number. *Discrete Math.*, 309(12):4149–4161, 2009.

[CDM05] Sourav Chatterjee, Persi Diaconis, and Elizabeth Meckes. Exchangeable pairs and Poisson approximation. *Probab. Surv.*, 2:64–106 (electronic), 2005.

[Cha07] Sourav Chatterjee. Stein's method for concentration inequalities. *Probab. Theory Related Fields*, 138(1-2):305–321, 2007.

[Che75] Louis H. Y. Chen. Poisson approximation for dependent trials. *Ann. Probability*, 3(3):534–545, 1975.

[Chu97] Fan R. K. Chung. *Spectral graph theory*, volume 92 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1997.

[Çın11] Erhan Çınlar. *Probability and stochastics*, volume 261 of *Graduate Texts in Mathematics*. Springer, New York, 2011.

[DJPP13] Ioana Dumitriu, Tobias Johnson, Soumik Pal, and Elliot Paquette. Functional limit theorems for random regular graphs. *Probab. Theory Related Fields*, 156(3–4):921–975, 2013.

[DS98] Persi Diaconis and Bernd Sturmfels. Algebraic algorithms for sampling from conditional distributions. *Ann. Statist.*, 26(1):363–397, 1998.

[Dub09] Julien Dubédat. SLE and the free field: partition functions and couplings. *J. Amer. Math. Soc.*, 22(4):995–1054, 2009.

[Dui13] Maurice Duits. Gaussian free field in an interlacing particle system with two jump rates. *Comm. Pure Appl. Math.*, 66(4):600–643, 2013.

[EG92] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions.* Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

[EPR+10] László Erdős, Sandrine Péché, José A. Ramírez, Benjamin Schlein, and Horng-Tzer Yau. Bulk universality for Wigner matrices. *Comm. Pure Appl. Math.*, 63(7):895–925, 2010.

[ERS+10] László Erdős, José Ramírez, Benjamin Schlein, Terence Tao, Van Vu, and Horng-Tzer Yau. Bulk universality for Wigner Hermitian matrices with subexponential decay. *Math. Res. Lett.*, 17(4):667–674, 2010.

[ERSY10] László Erdős, José A. Ramírez, Benjamin Schlein, and Horng-Tzer Yau. Universality of sine-kernel for Wigner matrices with a small Gaussian perturbation. *Electron. J. Probab.*, 15:no. 18, 526–603, 2010.

[ESY09a] László Erdős, Benjamin Schlein, and Horng-Tzer Yau. Local semicircle law and complete delocalization for Wigner random matrices. *Comm. Math. Phys.*, 287(2):641–655, 2009.

[ESY09b] László Erdős, Benjamin Schlein, and Horng-Tzer Yau. Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. *Ann. Probab.*, 37(3):815–852, 2009.

[Eva10] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, second edition, 2010.

[Fel68] William Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons, Inc., New York-London-Sydney, 1968.

[Fer14] Patrik L. Ferrari. Why random matrices share universal processes with interacting particle systems? Lecture notes. Available at arXiv:1312.1126, 2014.

[Fol95] Gerald B. Folland. *Introduction to partial differential equations.* Princeton University Press, Princeton, NJ, second edition, 1995.

[Fri91] Joel Friedman. On the second eigenvalue and random walks in random d-regular graphs. *Combinatorica*, 11:331–362, 1991.

[Fri08] Joel Friedman. A proof of Alon's second eigenvalue conjecture and related problems. *Mem. Amer. Math. Soc.*, 195(910):viii+100, 2008.

[GJKW02] Catherine Greenhill, Svante Janson, Jeong Han Kim, and Nicholas C. Wormald. Permutation pseudographs and contiguity. *Combin. Probab. Comput.*, 11(3):273–298, 2002.

[GM90] C. D. Godsil and B. D. McKay. Asymptotic enumeration of Latin rectangles. *J. Combin. Theory Ser. B*, 48(1):19–44, 1990.

[GM08] Catherine Greenhill and Brendan D. McKay. Asymptotic enumeration of sparse nonnegative integer matrices with specified row and column sums. *Adv. in Appl. Math.*, 41(4):459–481, 2008.

[GMW06] Catherine Greenhill, Brendan D. McKay, and Xiaoji Wang. Asymptotic enumeration of sparse 0-1 matrices with irregular row and column sums. *J. Combin. Theory Ser. A*, 113(2):291–324, 2006.

[GOS01] Giambattista Giacomin, Stefano Olla, and Herbert Spohn. Equilibrium fluctuations for $\nabla\phi$ interface model. *Ann. Probab.*, 29(3):1138–1172, 2001.

[GS14] Vadim Gorin and Mykhaylo Shkolnikov. Multilevel Dyson Brownian motions via Jack polynomials. Preprint. Available at arXiv:1401.5595, 2014.

[HMP10] Xiaoyu Hu, Jason Miller, and Yuval Peres. Thick points of the Gaussian free field. *Ann. Probab.*, 38(2):896–926, 2010.

[Hun] John K. Hunter. Notes on Partial Differential Equations. Available at `https://www.math.ucdavis.edu/~hunter/pdes/pdes.html`.

[Jan97] Svante Janson. *Gaussian Hilbert spaces*, volume 129 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1997.

[JLS14] David Jerison, Lionel Levine, and Scott Sheffield. Internal DLA and the Gaussian free field. *Duke Math. J.*, 163(2):267–308, 2014.

[JN06] Kurt Johansson and Eric Nordenstam. Eigenvalues of GUE minors. *Electron. J. Probab.*, 11:no. 50, 1342–1371, 2006.

[Joh12] Tobias Johnson. Exchangeable pairs, switchings, and random regular graphs. Preprint. Available at arXiv:1112.0704, 2012.

[JP12] Tobias Johnson and Soumik Pal. Cycles and eigenvalues of sequentially growing random regular graphs. To appear in *Annals of Probability*. Available at arXiv:1203.1113, 2012.

[JP13] Tobias Johnson and Elliot Paquette. Quantitative small subgraph conditioning. Preprint. Available at arXiv:1307.4858, 2013.

[Ken01] Richard Kenyon. Dominos and the Gaussian free field. *Ann. Probab.*, 29(3):1128–1137, 2001.

[Ken08] Richard Kenyon. Height fluctuations in the honeycomb dimer model. *Comm. Math. Phys.*, 281(3):675–709, 2008.

[KSV07] Jeong Han Kim, Benny Sudakov, and Van Vu. Small subgraphs of random regular graphs. *Discrete Math.*, 307(15):1961–1967, 2007.

[KSVW01] Michael Krivelevich, Benny Sudakov, Van H. Vu, and Nicholas C. Wormald. Random regular graphs of high degree. *Random Structures Algorithms*, 18(4):346–363, 2001.

[Kua11] Jeffrey Kuan. The Gaussian free field in interlacing particle systems. To appear in *Electronic Journal of Probability*. Available at arXiv:1109.4444, 2011.

[Lig10] Thomas M. Liggett. *Continuous time Markov processes*, volume 113 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. An introduction.

[Loè60] Michel Loève. *Probability theory*. 2nd ed. The University Series in Higher Mathematics. D. Van Nostrand Co., Inc., Princeton, N. J.-Toronto-New York-London, 1960.

[LP10] N. Linial and D. Puder. Word maps and spectra of random graph lifts. *Random Structures Algorithms*, 37(1):100–135, 2010.

[McK81] Brendan D. McKay. The expected eigenvalue distribution of a large regular graph. *Linear Algebra Appl.*, 40:203–216, 1981.

[McK84] Brendan D. McKay. Asymptotics for 0-1 matrices with prescribed line sums. In *Enumeration and design (Waterloo, Ont., 1982)*, pages 225–238. Academic Press, Toronto, ON, 1984.

[MS12a] Jason Miller and Scott Sheffield. Imaginary geometry I: interacting SLEs. Preprint. Available at arXiv:1201.1496, 2012.

[MS12b] Jason Miller and Scott Sheffield. Imaginary geometry II: reversibility of $\mathrm{SLE}_\kappa(\rho_1;\rho_2)$ for $\kappa \in (0,4)$. Preprint. Available at arXiv:1201.1497, 2012.

[MS12c] Jason Miller and Scott Sheffield. Imaginary geometry III: reversibility of SLE_κ for $\kappa \in (4, 8)$. Preprint. Available at arXiv:1201.1498, 2012.

[MS13] Jason Miller and Scott Sheffield. Imaginary geometry IV: interior rays, whole-plane reversibility, and space-filling trees. Preprint. Available at arXiv:1302.4738, 2013.

[MW03] Brendan D. McKay and Xiaoji Wang. Asymptotic enumeration of 0-1 matrices with equal row sums and equal column sums. *Linear Algebra Appl.*, 373:273–287, 2003. Special issue on the Combinatorial Matrix Theory Conference (Pohang, 2002).

[MWW04] Brendan D. McKay, Nicholas C. Wormald, and Beata Wysocka. Short cycles in random regular graphs. *Electron. J. Combin.*, 11(1):Research Paper 66, 12 pp. (electronic), 2004.

[NS97] Ali Naddaf and Thomas Spencer. On homogenization and scaling limit of some gradient perturbations of a massless free field. *Comm. Math. Phys.*, 183(1):55–84, 1997.

[Pet12] Leonid Petrov. Asymptotics of uniformly random lozenge tilings of polygons. Gaussian free field. To appear in *Annals of Probability*. Available at arXiv:1206.5123, 2012.

[Pit06] Jim Pitman. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard.

[Ros11] Nathan Ross. Fundamentals of Stein's method. *Probab. Surv.*, 8:210–293, 2011.

[RV07] Brian Rider and Bálint Virág. The noise in the circular law and the Gaussian free field. *Int. Math. Res. Not. IMRN*, (2):Art. ID rnm006, 33, 2007.

[She07] Scott Sheffield. Gaussian free fields for mathematicians. *Probab. Theory Related Fields*, 139(3-4):521–541, 2007.

[Sod07] Sasha Sodin. Random matrices, nonbacktracking walks, and orthogonal polynomials. *J. Math. Phys.*, 48(12):123503, 21, 2007.

[Spi12] Daniel Spielman. Spectral graph theory. In *Combinatorial scientific computing*, Chapman & Hall/CRC Comput. Sci. Ser., pages 495–524. CRC Press, Boca Raton, FL, 2012.

[SS09] Oded Schramm and Scott Sheffield. Contour lines of the two-dimensional discrete Gaussian free field. *Acta Math.*, 202(1):21–137, 2009.

[SS13] Oded Schramm and Scott Sheffield. A contour line of the continuum Gaussian free field. *Probab. Theory Related Fields*, 157(1-2):47–80, 2013.

[Ste72] Charles Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*, pages 583–602, Berkeley, Calif., 1972. Univ. California Press.

[Ste78] Charles M. Stein. Asymptotic evaluation of the number of Latin rectangles. *J. Combin. Theory Ser. A*, 25(1):38–49, 1978.

[Ste92] Charles Stein. A way of using auxiliary randomization. In *Probability theory (Singapore, 1989)*, pages 159–180. de Gruyter, Berlin, 1992.

[TV10] Terence Tao and Van Vu. Random matrices: universality of local eigenvalue statistics up to the edge. *Comm. Math. Phys.*, 298(2):549–572, 2010.

[TV11] Terence Tao and Van Vu. Random matrices: universality of local eigenvalue statistics. *Acta Math.*, 206(1):127–204, 2011.

[War07] Jon Warren. Dyson's Brownian motions, intertwining and interlacing. *Electron. J. Probab.*, 12:no. 19, 573–590, 2007.

[Wor81] Nicholas C. Wormald. The asymptotic distribution of short cycles in random regular graphs. *J. Combin. Theory Ser. B*, 31(2):168–182, 1981.

[Wor99a] N. C. Wormald. Models of random regular graphs. In *Surveys in combinatorics, 1999 (Canterbury)*, volume 267 of *London Math. Soc. Lecture Note Ser.*, pages 239–298. Cambridge Univ. Press, Cambridge, 1999.

[Wor99b] Nicholas C. Wormald. Models of random regular graphs. In *Surveys in combinatorics, 1999 (Canterbury)*, volume 267 of *London Math. Soc. Lecture Note Ser.*, pages 239–298. Cambridge Univ. Press, Cambridge, 1999.