

# Exchangeable pairs, switchings, and random regular graphs

Tobias Johnson\*

Department of Mathematics  
University of Southern California  
Los Angeles, CA, U.S.A.

tobias.johnson@usc.edu

Submitted: Sep 3, 2014; Accepted: Jan 3, 2015; Published: Feb 16, 2015

Mathematics Subject Classifications: 05C80, 60B10, 60B20

## Abstract

We consider the distribution of cycle counts in a random regular graph, which is closely linked to the graph's spectral properties. We broaden the asymptotic regime in which the cycle counts are known to be approximately Poisson, and we give an explicit bound in total variation distance for the approximation. Using this result, we calculate limiting distributions of linear eigenvalue statistics for random regular graphs.

Previous results on the distribution of cycle counts by McKay, Wormald, and Wysocka (2004) used the method of switchings, a combinatorial technique for asymptotic enumeration. Our proof uses Stein's method of exchangeable pairs and demonstrates an interesting connection between the two techniques.

**Keywords:** switchings, Stein's method, exchangeable pairs, random regular graphs, linear eigenvalue statistics

## 1 Introduction

Suppose that  $\lambda_1, \dots, \lambda_n$  are the eigenvalues of an  $n \times n$  random matrix. The random variable  $\sum_{i=1}^n f(\lambda_i)$  for a given function  $f$  is known as a *linear eigenvalue statistic*, and it is a common object of study in random matrix theory, typically as  $n$  tends to infinity.

Let  $G$  be chosen uniformly at random from the space of all simple  $d$ -regular graphs on  $n$  vertices, and consider its adjacency matrix. Brendan McKay determined the first-order behavior of its linear eigenvalue statistics, showing that  $n^{-1} \sum_{i=1}^n f(\lambda_i)$  converged in probability to a deterministic limit as  $n \rightarrow \infty$  [23]. In [13], the second-order behavior of

---

\*The author acknowledges support from the NSF by grants DMS-0847661 and DMS-1401479.

linear eigenvalue statistics was computed for a slightly different model of random regular graph, with improved results given in [27, Chapter 3]. The motivating goal of this paper is to prove similar results for uniformly chosen random regular graphs, which we carry out in Theorems 19 and 21.

We will defer further discussion of this problem and its background until Section 4. Until then, we discuss several combinatorial and probabilistic results interesting in their own right that we will achieve along the way. Let  $C_k$  denote the number of cycles of length  $k$  in the random regular graph  $G$ . The distribution of these random variables has been studied since [6, 36], where it was proven that  $(C_3, \dots, C_r)$  converges in law to a vector of independent Poisson random variables as  $n$  tends to infinity, with  $r$  held fixed. As early as [23], the cycle counts of a graph have been used to investigate properties of the graph's eigenvalues. We take this approach as well, converting our original problem into one of accurately estimating the distribution of this random vector.

The strongest results on the cycle counts of a random regular graph came in [26], 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  $G$  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 [26] to speculate that the Poisson approximation failed beyond this threshold. Surprisingly, this is not the case. In Theorem 11, we give a Poisson approximation for the cycle counts that holds so long as  $\sqrt{r}(d - 1)^{\frac{3}{2}r-1} = o(n)$ . We give a quantitative bound on the accuracy of the approximation, which is the necessary ingredient for our results on linear eigenvalue statistics. As a bonus, we give in Theorem 7 a distributional approximation not just of the cycle counts, but of a more general process defined by the cycles.

The Poisson approximation in [26, Theorem 1] uses a combinatorial technique for asymptotic enumeration known as the method of switchings. We adapt this technique to use Stein's method of exchangeable pairs for Poisson approximation. We discuss both methods further in the following section. As noted in [37], they have some obvious similarity, but we believe that this is the first time they have been connected in a rigorous way. This connection gives a novel construction of an exchangeable pair for use with Stein's method, and it allows the machinery of Stein's method to be used in some new combinatorial settings.

In Section 2, we give some basic definitions and preliminary estimates on random regular graphs. Section 3 presents our Poisson approximation. The core argument and the most general result is Theorem 7, and our main result on cycle counts is Theorem 11. In Section 4, we give the context and proofs of our results on linear eigenvalue statistics of random regular graphs.

## 1.1 Switchings and Stein's method

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 [16] and matrices with prescribed row and column sums [24, 25, 19]. It has seen its biggest use in analyzing regular graphs; see [22], [26], [21], and [5] for

some examples. A good summary of switchings in random regular graphs can be found in Section 2.4 of [35].

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 [30]. Louis Chen adapted the method for Poisson approximation [11]. Since then, Stein, Chen, and a score of others have adapted Stein's method to a wide variety of circumstances. The survey paper [28] gives a broad introduction to Stein's method, and [3] and [10] focus specifically on using it for Poisson approximation.

We will use the technique of exchangeable pairs, following the treatment in [10]. 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 [32, 16], and the analysis of random contingency tables in [12] is similar to combinatorial work like [18]. 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 switching 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 [9]). On the other hand, Stein's method cannot prove results as sharp as [26, 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.

## 2 Preliminaries

A *d-regular graph* is one for which all vertices have degree exactly  $d$ . We call a graph *simple* if it has no loops (edges between a vertex and itself) or parallel edges. By *random d-regular graph on n vertices*, we mean a random graph chosen uniformly from the space

of all simple  $d$ -regular graphs on  $n$  vertices (unless we specifically refer to another model). When  $d$  is odd, we always assume that  $n$  is even, since there are no  $d$ -regular graphs on  $n$  vertices with  $d$  and  $n$  odd. By *cycle*, we mean what is sometimes called a simple cycle: a walk on a graph starting and ending at the same vertex, and with no repeated edges or vertices along the way. 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 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 1.** *Let  $G$  be a random  $d$ -regular graph on  $n$  vertices, with  $d \leq n^{1/3}$ .*

(a) *Suppose  $H$  is a subgraph of the complete graph  $K_n$  in which every vertex has degree 2 or higher. Let  $e$  be the number of edges and  $v$  the number of vertices in  $H$ . Suppose  $e \leq 2n^{1/10}$ . Then*

$$\mathbf{P}[H \subseteq G] \leq \frac{c_1(d-1)^e}{n^e}.$$

(b) *Let  $\alpha$  be a cycle of length  $k \leq 2n^{1/10}$  in the complete graph  $K_n$ . Then*

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

(c) *Let  $\beta$  be another cycle in  $K_n$  of length  $j \leq n^{1/10}$ , and suppose that  $\alpha$  and  $\beta$  share  $f$  edges. Then*

$$\mathbf{P}[\alpha \cup \beta \subseteq G] \leq \frac{c_1(d-1)^{j+k-f}}{n^{j+k-f}}.$$

*Proof.* Statements (b) and (c) are specializations of (a), which follows directly from Theorem 3a in [26].  $\square$

### 3 Poisson approximation of cycle counts by Stein's method

#### 3.1 Stein's method background

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

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

for some constant  $c$ , then  $X$  is approximately Poisson with mean  $\lambda$ . (When  $X$  and  $X'$  are two steps in a stationary immigration-death chain whose invariant distribution is Poisson with mean  $\lambda$ , these equations hold exactly.) The following proposition gives a precise, multivariate version of this heuristic. Recall that the total variation distance between the laws of two random variables  $X$  and  $Y$  taking values in  $\mathbb{N} = \{0, 1, 2, \dots\}$  is given by

$$d_{TV}(X, Y) := \sup_{A \subseteq \mathbb{N}} |\mathbf{P}[X \in A] - \mathbf{P}[Y \in A]|.$$

**Proposition 2** ([10, 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  $\sigma$ -algebra  $\mathcal{F}$  with respect to which  $W$  is measurable and any choice of constants  $c_k$ ,*

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

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

$$\begin{aligned} \Delta_k^+ &= \{W'_k = W_k + 1, W_j = W'_j \text{ for } k < j \leq r\}, \\ \Delta_k^- &= \{W'_k = W_k - 1, W_j = W'_j \text{ for } k < j \leq r\}. \end{aligned}$$

*Remark 3.* We have changed the statement of the proposition from [10] 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 4.* 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 [31, Section 2], one can directly investigate ratios of probabilities of different values of  $X$  using the equation

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

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.

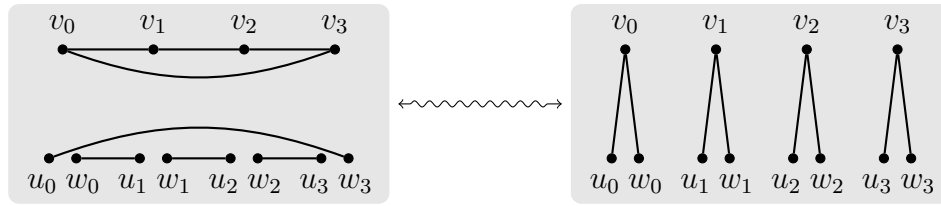


Figure 1: The change from left to right is a *forward switching*, and from right to left is a *backward switching*.

### 3.2 Counting switchings

We start by defining our switchings. Besides some small notational differences, the definitions will be the same as those in [26]. 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  $\alpha$  deleted (see Figure 1). We call this action induced given by the sequences  $(v_i)$ ,  $(u_i)$ , and  $(w_i)$  a *forward  $\alpha$ -switching*. We will consider forward  $\alpha$ -switchings only up to cyclic rotation of indices; that is, we identify the  $2k$  different  $\alpha$ -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  $\alpha$ -switching*. Again, we consider switchings only up to cyclic rotation of all indices.

We call an  $\alpha$ -switching *valid* if  $\alpha$  is the only short cycle created or destroyed by the switching. For each valid forward  $\alpha$ -switching taking  $G$  to  $G'$ , there is a corresponding valid backwards  $\alpha$ -switching taking  $G'$  to  $G$ . Let  $F_\alpha$  and  $B_\alpha$  be the number of valid forward and backwards  $\alpha$ -switchings, respectively, on some graph  $G$ . Using arguments drawn from [26, Lemma 3], we give some estimates on them.

**Lemma 5.** *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  $k$ ,*

$$F_\alpha \leq [n]_k d^k. \tag{1}$$

*If  $\alpha$  does not share an edge with another short cycle,*

$$F_\alpha \geq [n]_k d^k \left( 1 - \frac{2k \sum_{j=3}^r j C_j + c_2 k (d-1)^r}{nd} \right). \tag{2}$$

*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  $e'_0, \dots, e'_{k-1}$  that give a valid switching? There are at most  $[n]_k d^k$  choices of oriented edges  $e'_0, \dots, e'_{k-1}$ , which proves the upper bound (1). For the lower bound, we demonstrate a procedure to choose these edges that is guaranteed to give us a valid forward  $\alpha$ -switching. Suppose that  $e'_0, \dots, e'_{k-1}$  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 [26], 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  $\alpha$ .

Since  $\alpha$  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  $e'_0, \dots, e'_{k-1}$ . The switching does not create any short cycles either: Suppose otherwise, and let  $\beta$  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  $\alpha$  and has length less than  $r/2$ , then combining this path with a path in  $\alpha$  gives a short cycle in  $G$  that intersects  $\alpha$ ;
- if it starts in  $\alpha$  and finishes in  $W = \{u_0, w_0, \dots, u_{k-1}, w_{k-1}\}$  and has length less than  $r/2$ , then combining this path with a path in  $\alpha$  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  $\beta$  contains exactly one path in  $G \cap G'$ . The remainder of  $\beta$  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  $F_\alpha$ . We will do this by bounding from above the number of switchings out of the  $[n]_k d^k$  counted in (1) that fail each condition (a)–(d).

- There are a total of  $\sum_{j=3}^r j C_j$  edges in short cycles in  $G$ . Choosing one of the edges  $e'_0, \dots, e'_{k-1}$  from these and the rest arbitrarily, there are at most

$$[n-1]_{k-1} d^{k-1} k \sum_{j=3}^r 2j C_j$$

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

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

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

switchings out of the original  $[n]_k d^k$  fail conditions by (a)–(d), establishing (2).  $\square$

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

**Lemma 6.** *Let  $G$  be a random  $d$ -regular graph on  $n$  vertices, and let  $\alpha$  be a cycle of length  $k \leq r$  in the complete graph  $K_n$ . Then*

$$B_\alpha \leq (d(d-1))^k \tag{3}$$

and

$$\mathbf{E}B_\alpha \geq (d(d-1))^k \left( 1 - \frac{c_3 k (d-1)^{r-1}}{n} \right). \tag{4}$$

*Proof.* The question this time is given  $\alpha$ , how many choices of oriented paths yield a valid switching? For any fixed  $\alpha$ , there are at most  $(d(d-1))^k$  choices of oriented paths, proving (3). For the lower bound, let  $B = \sum_\beta B_\beta$ , where  $\beta$  runs over all cycles of length  $k$  in the complete graph. We will first show that

$$B \geq \frac{[n]_k (d(d-1))^k}{2k} \left( 1 - \frac{4k \sum_{j=3}^r j C_j + O(k(d-1)^r)}{nd} \right). \tag{5}$$

As in Lemma 5, we give conditions that ensure a valid switching. Let  $\beta = v_0 \cdots v_{k-1}$ , and suppose that the paths  $u_i v_i w_i$  in  $G$  for  $0 \leq i \leq k-1$  satisfy

- (a) the edges  $v_i u_i$  and  $v_i w_i$  are not contained in any short cycles;
- (b) for all  $1 \leq j \leq r/2$ , 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  $\beta$  are created or destroyed by the switching. By (a), none are destroyed. Suppose a short cycle  $\beta'$  other than  $\beta$  is created in  $G'$ . It consists of paths in  $G \cap G'$ , portions of  $\beta$ , 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  $1 \leq j \leq r/2$ , then (b) implies this;
- if it starts and ends at one of  $u_i, v_i$ , and  $w_i$ , then (a) implies this.

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

There are  $[n]_k d^k / 2k$  choices for  $\beta$ , and at most  $(d(d-1))^k$  choices for  $u_i, w_i$ ,  $0 \leq i < k$ . 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 5, we find that at most

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

of the switchings violate condition (a), and at most  $[n]_{k-1} (d(d-1))^{k-1} O((d-1)^{r+1})$  violate condition (b), which proves (5).

By Proposition 1b (or by [26, eq. 2.2]),

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

Applying this to (5) gives

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

By the exchangeability of the vertex labels of  $G$ , the law of  $B_\beta$  is the same for all  $k$ -cycles  $\beta$ . It follows that  $\mathbf{E}B = ([n]_k / 2k) \mathbf{E}B_\alpha$ , proving (4).  $\square$

### 3.3 Applying Stein's method

Rather than prove a theorem about the vector of cycle counts, we will give a result on a more general process. Let  $\mathcal{J}$  be an index set of possible cycles in  $K_n$  that the random graph  $G$  might contain, and for  $\alpha \in \mathcal{J}$ , let  $I_\alpha$  be an indicator on  $G$  containing  $\alpha$ . We will show that the entire process  $(I_\alpha, \alpha \in \mathcal{J})$  is well approximated by a vector of independent Poissons, with the accuracy of the approximation depending on the size of the set  $\mathcal{J}$ . We will also prove a slight variant in Proposition 10 which achieves a better error bound, at

the expense of considering a less general process. Our result on cycle counts, Theorem 11, will follow easily from this.

Though we have no need for these process approximations in our paper, similar results for the permutation model of random graph have proven useful in [20]. In any event, the machinery of Stein's method gives them to us with no extra effort.

**Theorem 7.** *Let  $G$  be a random  $d$ -regular graph on  $n$  vertices. For some collection  $\mathcal{J}$  of cycles in the complete graph  $K_n$  of maximum length  $r$ , we define  $\mathbf{I} = (I_\alpha, \alpha \in \mathcal{J})$ , with  $I_\alpha = \mathbf{1}\{G \text{ contains } \alpha\}$ . Let  $\mathbf{Z} = (Z_\alpha, \alpha \in \mathcal{J})$  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  $\alpha$ .*

*For some absolute constant  $c_4$ , for all  $n$  and  $d, r \geq 3$  satisfying  $r \leq n^{1/10}$  and  $d \leq n^{1/3}$ ,*

$$d_{TV}(\mathbf{I}, \mathbf{Z}) \leq \sum_{\alpha \in \mathcal{J}} \frac{c_4 |\alpha| (d-1)^{|\alpha|+r-1}}{n^{|\alpha|+1}}.$$

Before we give the proof, we show the result of applying this theorem when  $\mathcal{J}$  is all cycles of length  $r$  or less:

**Corollary 8.** *Let  $G$  be a random  $d$ -regular graph on  $n$  vertices, and let  $\mathcal{J}$  be the collection of all cycles of length  $r$  or less in the complete graph  $K_n$ . Define  $\mathbf{I}$  and  $\mathbf{Z}$  as in the previous theorem. For some absolute constant  $c_5$ , for all  $n$  and  $d, r \geq 3$ ,*

$$d_{TV}(\mathbf{I}, \mathbf{Z}) \leq \frac{c_5 (d-1)^{2r-1}}{n}.$$

*Proof of the corollary.* If  $r > n^{1/10}$  or  $d > n^{1/3}$ , then  $c_5 (d-1)^{2r-1}/n > 1$  for a sufficiently large choice of  $c_5$ , and the total variation bound is trivial. Thus we can assume that this is not the case and apply the previous theorem:

$$\begin{aligned} d_{TV}(\mathbf{I}, \mathbf{Z}) &\leq \sum_{\alpha \in \mathcal{J}} \frac{c_4 |\alpha| (d-1)^{|\alpha|+r-1}}{n^{|\alpha|+1}} \\ &= \sum_{k=3}^r \frac{[n]_k}{2k} \left( \frac{c_4 k (d-1)^{k+r-1}}{n^{k+1}} \right) \\ &= O\left(\frac{(d-1)^{2r-1}}{n}\right). \quad \square \end{aligned}$$

The strength of Theorem 7 is that one can consider a smaller set  $\mathcal{J}$  of possible cycles and get a tighter total variation bound. For instance, if  $\mathcal{J}$  is the set of all cycles in  $K_n$  of length  $r$  or less containing vertex 1, then  $\mathbf{I}$  and  $\mathbf{Z}$  are within  $O(r(d-1)^{2r-1}/n^2)$  in total variation norm.

*Remark 9.* Since the cycle counts  $(C_3, \dots, C_r)$  are a functional of  $\mathbf{I}$ , this corollary implies that

$$d_{TV}((C_3, \dots, C_r), (Z_3, \dots, Z_r)) \leq \frac{c_5 (d-1)^{2r-1}}{n},$$

where  $(Z_3, \dots, Z_r)$  is a vector of independent Poisson random variables with  $\mathbf{E}Z_k = (d-1)^k/2k$ . In fact, we will give a slightly better result in Theorem 11.

*Proof of Theorem 7.* 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  $\alpha$ -switching with  $\alpha \in \mathcal{J}$  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  $\alpha$ -switchings from  $G_0$  to  $G_1$  are in bijective correspondence with valid backwards  $\alpha$ -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  $I'_\alpha$  be an indicator on  $G'$  containing the cycle  $\alpha$ , and define  $\mathbf{I}' = (I'_\alpha, \alpha \in \mathcal{J})$ . It follows from the exchangeability of  $G$  and  $G'$  that  $\mathbf{I}$  and  $\mathbf{I}'$  are exchangeable, and we can apply Proposition 2 on this pair. Define the events  $\Delta_\alpha^+$  and  $\Delta_\alpha^-$  as in that proposition. By our construction,

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

Thus by Proposition 2 with all constants set to  $d_0$ ,

$$d_{TV}(\mathbf{I}, \mathbf{Z}) \leq \sum_{\alpha \in \mathcal{J}} \mathbf{E} \left| \frac{(d-1)^{|\alpha|}}{[n]_{|\alpha|}} - \frac{B_\alpha}{[n]_{|\alpha|}d^{|\alpha|}} \right| + \sum_{\alpha \in \mathcal{J}} \mathbf{E} \left| I_\alpha - \frac{F_\alpha}{[n]_{|\alpha|}d^{|\alpha|}} \right|. \quad (6)$$

We will bound these two sums. Fix some  $\alpha \in \mathcal{J}$ , and let  $|\alpha| = k$ . Applying first the upper bound and then the lower bound from Lemma 6,

$$\begin{aligned} \mathbf{E} \left| \frac{(d-1)^k}{[n]_k} - \frac{B_\alpha}{[n]_k d^k} \right| &= \mathbf{E} \left[ \frac{(d-1)^k}{[n]_k} - \frac{B_\alpha}{[n]_k d^k} \right] \\ &\leq \frac{c_3 k (d-1)^{k+r-1}}{n [n]_k}. \end{aligned} \quad (7)$$

To bound the other sum, partition the 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 5 apply, giving us

$$\left| I_\alpha - \frac{F_\alpha}{[n]_k d^k} \right| \leq \frac{2k \sum_{j=3}^r j C_j + c_2 k (d-1)^r}{nd}.$$

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

$$\begin{aligned} \mathbf{E} \left| I_\alpha - \frac{F_\alpha}{[n]_k d^k} \right| &\leq \mathbf{E} \left[ \mathbf{1}_{A_2} \frac{2k \sum_{j=3}^r j C_j + c_2 k (d-1)^r}{nd} + \mathbf{1}_{A_3} \right] \\ &= \frac{2k}{nd} \mathbf{E} \left[ \mathbf{1}_{A_2} \sum_{j=3}^r j C_j \right] + \frac{c_2 k (d-1)^r}{nd} \mathbf{P}[A_2] + \mathbf{P}[A_3]. \end{aligned}$$

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

$$\begin{aligned} \mathbf{E} \left| I_\alpha - \frac{F_\alpha}{[n]_k d^k} \right| &\leq \frac{2k^2}{nd} \mathbf{E} \mathbf{1}_{A_2} + \frac{2k}{nd} \sum_{\beta \in \mathcal{J}} |\beta| \mathbf{E} \mathbf{1}_{A_2} I_\beta + \frac{c_2 k (d-1)^r}{nd} \mathbf{P}[A_2] + \mathbf{P}[A_3] \\ &\leq \frac{2k^2}{nd} \mathbf{E} I_\alpha + \frac{2k}{nd} \sum_{\beta \in \mathcal{J}} |\beta| \mathbf{E} I_\alpha I_\beta + \frac{c_2 k (d-1)^r}{nd} \mathbf{E} I_\alpha + \mathbf{P}[A_3]. \end{aligned} \quad (8)$$

By Proposition 1b,

$$\frac{2k^2}{nd} \mathbf{E} I_\alpha = O\left(\frac{k^2 (d-1)^{k-1}}{n^{k+1}}\right) \quad (9)$$

and

$$\frac{c_2 k (d-1)^r}{nd} \mathbf{E} I_\alpha = O\left(\frac{k (d-1)^{k+r-1}}{n^{k+1}}\right). \quad (10)$$

By Proposition 1c with  $f = 0$ , for any  $\beta \in \mathcal{J}$  we have  $\mathbf{E} I_\alpha I_\beta \leq c_1 (d-1)^{j+k} / n^{j+k}$ , where  $j = |\beta|$ . For each  $3 \leq j \leq r$ , there are at most  $[n]_j / 2j$  cycles in  $\mathcal{J}$  of length  $j$ . Therefore

$$\begin{aligned} \frac{2k}{nd} \sum_{\beta \in \mathcal{J}} |\beta| \mathbf{E} I_\alpha I_\beta &\leq \frac{2k}{nd} \sum_{j=3}^r \frac{[n]_j}{2j} \left( \frac{j c_1 (d-1)^{j+k}}{n^{j+k}} \right) \\ &\leq \frac{k}{nd} \sum_{j=3}^r \frac{c_1 (d-1)^{j+k}}{n^k} = O\left(\frac{k (d-1)^{k+r-1}}{n^{k+1}}\right). \end{aligned} \quad (11)$$

The last term of (8) is the most difficult to bound. Let  $\mathcal{K}$  be the set of short cycles in  $K_n$  that share an edge with  $\alpha$ , not including  $\alpha$  itself. By a union bound,

$$\mathbf{P}[A_3] \leq \sum_{\beta \in \mathcal{K}} \mathbf{E} I_\alpha I_\beta. \quad (12)$$

Now, we classify and count the cycles  $\beta \in \mathcal{K}$  according to the structure of  $\alpha \cup \beta$ . Suppose that  $\beta$  has length  $j$ , and consider the intersection of  $\alpha$  and  $\beta$  (the graph consisting of all vertices and edges contained in both  $\alpha$  and  $\beta$ ). Suppose this intersection graph has  $p$

components and  $f$  edges. As computed on [26, 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  $\beta$  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  $\alpha$ . In defining  $\beta$ , the remaining  $j-p-f$  vertices can be chosen to be anything, and the intersection of  $\alpha$  and  $\beta$  can be rotated around  $\alpha$  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  $\alpha$  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  $\beta$ , we have  $\mathbf{E}I_\alpha I_\beta \leq c_1(d-1)^{j+k-f}/n^{j+k-f}$  by Proposition 1c. Applying this to (12),

$$\begin{aligned} \mathbf{P}[A_3] &\leq \sum_{j=3}^r \sum_{p,f \geq 1} \frac{(2r^3)^{p-1}}{(p-1)!^2} 2kn^{j-p-f} \frac{c_1(d-1)^{j+k-f}}{n^{j+k-f}} \\ &= \sum_{j=3}^r \sum_{p,f \geq 1} \frac{(2r^3)^{p-1}}{(p-1)!^2} \frac{2kc_1(d-1)^{j+k-f}}{n^{k+p}} \\ &= \sum_{j=3}^r O\left(\frac{k(d-1)^{j+k-1}}{n^{k+1}}\right) = O\left(\frac{k(d-1)^{k+r-1}}{n^{k+1}}\right). \end{aligned} \tag{13}$$

Combining (9), (10), (11), and (13), we have

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

Applying this and (7) to (6) establishes the theorem.  $\square$

As mentioned in Remark 9, we can apply this theorem to give a total variation bound on the law of any functional of  $\mathbf{I}$ . This bound is often less than optimal, since this theorem fails to exploit the  $\lambda_k^{-1/2}$  factors in Proposition 2. We will take advantage of these factors in the following proposition, and then apply this to prove Theorem 11.

**Proposition 10.** *With the set-up of Theorem 7, divide up the collection of cycles  $\mathcal{J}$  into bins  $\mathcal{B}_1, \dots, \mathcal{B}_s$ . Let*

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

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

$$d_{TV}((I_1, \dots, I_s), (Z_1, \dots, Z_s)) \leq c_4 \sum_{k=1}^s \xi_k \sum_{\alpha \in \mathcal{B}_k} \frac{|\alpha| (d-1)^{|\alpha|+r-1}}{n^{|\alpha|+1}},$$

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

*Proof.* Define the exchangeable pair  $(G, G')$  as in Theorem 7, and define  $I'_1, \dots, I'_s$  as the analogous quantities in  $G'$ . Define  $\Delta_k^+$  and  $\Delta_k^-$  as in Proposition 2, noting that

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

By Proposition 2,

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

These summands were already bounded in expectation in Theorem 7, and applying these bounds proves the proposition.  $\square$

**Theorem 11.** *Let  $G$  be a random  $d$ -regular graph on  $n$  vertices 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  $r, d \geq 3$ ,*

$$d_{TV}((C_3, \dots, C_r), (Z_3, \dots, Z_r)) \leq \frac{c_6 \sqrt{r} (d-1)^{3r/2-1}}{n}$$

*Proof.* If  $d > n^{1/3}$  or  $r > n^{1/10}$ , then  $c_6 \sqrt{r} (d-1)^{3r/2-1}/n > 1$  for a sufficiently large choice of  $c_6$ , and the theorem holds trivially. Thus we can assume that  $d \leq n^{1/3}$  and  $r \leq n^{1/10}$ .

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

$$\begin{aligned} d_{TV}((C_3, \dots, C_r), (Z_3, \dots, Z_r)) &\leq c_4 \sum_{k=3}^r 1.4 \lambda_k^{-1/2} \sum_{\alpha \in \mathcal{J}_k} \frac{k(d-1)^{k+r-1}}{n^{k+1}} \\ &= \sum_{k=3}^r O\left(\frac{\sqrt{k}(d-1)^{k/2+r-1}}{n}\right) \\ &= O\left(\frac{\sqrt{r}(d-1)^{3r/2-1}}{n}\right). \end{aligned} \quad \square$$

## 4 Eigenvalue fluctuations of random regular graphs

Consider a random symmetric  $n \times n$  matrix  $X_n$  with eigenvalues  $\lambda_1 \geq \dots \geq \lambda_n$ . As we mentioned in the introduction, a *linear eigenvalue statistic* is a random variable of the

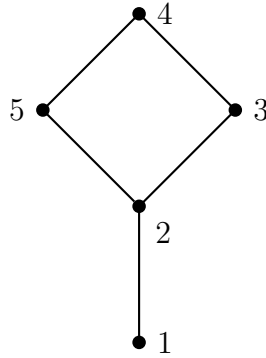


Figure 2: The walk  $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 2 \rightarrow 1$  is non-backtracking, but not cyclically non-backtracking.

form  $\sum_{i=1}^n f(\lambda_i)$  for some function  $f$ . A common problem in random matrix theory is to understand the asymptotic behavior of linear eigenvalue statistics. Typically, one shows convergence to a deterministic limit under one scaling (the first-order behavior), and to a distributional limit under another scaling (the second-order behavior). The prototypical example is when  $X_n$  is a Wigner matrix: the first-order behavior is given by Wigner’s semicircle law (see [1] for a modern account of Wigner’s result), and for sufficiently smooth  $f$ , the fluctuations from this are normal [29, 2].

Recently, the problem of finding the fluctuations of linear eigenvalue statistics was considered for random permutation matrices [4], where for sufficiently smooth  $f$ , the limiting distribution is non-Gaussian. This is striking because this behavior is *non-universal*. The naïve expectation would have been that the eigenvalues of these matrices should behave as in the *Gaussian orthogonal ensemble*, which consists of Wigner matrices with Gaussian entries. In [13], the same problem was considered for the adjacency matrices of random  $d$ -regular graphs drawn from the permutation model. As with random permutation matrices, for sufficiently smooth  $f$ , the limiting fluctuations are non-Gaussian if  $d$  is a fixed constant. On the other hand, if  $d$  grows to infinity with  $n$ , the limiting fluctuations are Gaussian. The first-order behavior of linear eigenvalue statistics shows the same dichotomy, with the non-universal Kesten-McKay limit when  $d$  is fixed replaced by the semicircle law when  $d$  grows with  $n$  [14, 33].

Our goal is to extend these fluctuation results to the uniform model of random regular graph. Following the approach of [13], we will use Theorem 11 to estimate the distribution of counts of *cyclically non-backtracking walks*. Using a connection between these counts and the graph’s eigenvalues, we compute the non-Gaussian limiting fluctuations in Theorem 19. We will then show in Theorem 21 that when  $d$  grows with  $n$ , the eigenvalue fluctuations converge to nearly the same limit as in the GOE.

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 (that is, the walk does not look like the one given in Figure 2). 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 from the uniform model, and let  $C_k^{(n)}$  be the number of cycles of length  $k$  in  $G_n$ . We define the random variable  $\text{CNBW}_k^{(n)}$  to be the number of cyclically non-backtracking walks of length  $k$  in  $G_n$ . Define  $(C_k^{(\infty)}, k \geq 3)$  to be independent Poisson random variables, with  $C_k^{(\infty)}$  having mean  $\lambda_k = (d-1)^k/2k$ . It will be convenient to define  $C_1^{(\infty)}, C_2^{(\infty)}, C_1^{(n)}$ , and  $C_2^{(n)}$  as zero.

Define

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

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 the following proposition will show.

**Proposition 12.** *Suppose  $d \leq n^{1/3}$  and  $k \leq n^{1/10}$ . Let*

$$B_k^{(n)} := \text{CNBW}_k^{(n)} - \sum_{j|k} 2jC_j^{(n)},$$

*the number of cyclically non-backtracking walks in the random  $d$ -regular graph  $G_n$  that are not repeated walks around cycles. Then*

$$\mathbf{E}B_k^{(n)} \leq \frac{c_7 k^6 (d-1)^k}{n}.$$

*Proof.* Call a cyclically non-backtracking walk *bad* if it is not a repeated walk around a cycle. We just need to enumerate the possible bad walks and apply Proposition 1a to bound the probability of each one. First, we give some notation first used in [7]. Let  $v_0, \dots, v_k \in \{1, \dots, n\}$  satisfying  $v_0 = v_k$  be a sequence of vertices that forms a bad cyclically non-backtracking walk. Let  $1 \leq i \leq k$ . We say that the  $i$ th step of the walk is

- *free* if  $v_i$  did not previously occur in the walk;
- a *coincidence* if  $v_i$  previously occurred in the walk, but the edge  $v_{i-1}v_i$  did not;
- and *forced* if the edge  $v_{i-1}v_i$  previously occurred in the walk.

Let  $\chi + 1$  be the number of coincidences and  $f$  the number of forced steps in the walk. With  $v$  the number of vertices and  $e$  the number of edges in the graph formed by the walk, we then have

$$\begin{aligned} v &= k - \chi - f, \\ e &= k - f. \end{aligned}$$

It follows from the walk being bad that  $\chi \geq 1$ .



**Claim 13.** Consider walks on  $K_n$  such that when the walk is viewed as a subgraph, all vertices have degree at most  $d$ . The number of such walks with given values  $\chi \geq 1$  and  $f \geq 0$  is at most  $k^{3\chi+2}(d-1)^f n^{k-\chi-f}$ .

*Proof of the claim.* Imagine laying out the coincidences, then the forced steps, and then the free steps. Given that there are  $\chi + 1$  coincidences, there are  $\binom{k}{\chi+1} \leq k^{\chi+1}$  possible subsets of indices  $\{1, \dots, k\}$  where the coincidences can occur. The vertex at a coincidence has already occurred in the walk, so there are fewer than  $k$  choices for each of them, giving us a total of  $k^{2\chi+2}$  choices so far.

Forced steps can occur only after a coincidence or another forced step. After each coincidence, imagine assigning some number of the steps to be forced. The number of ways to do this is at most the number of weak compositions of  $f$  elements into  $\chi + 1$  parts,  $\binom{f+\chi}{\chi}$ , which we can bound by  $k^\chi$ . At each forced step, the walk can only move along an edge that has already been traversed, so there are at most  $d - 1$  possible choices of vertices at each forced step. In all, this gives us at most  $k^\chi(d-1)^f$  choices for the forced steps.

At each of the  $k - \chi - 1 - f$  free steps, we have at most  $n$  choices of where to move, and we have an additional  $n$  choices for  $v_0$ , giving us another  $n^{k-\chi-f}$  choices in all. Multiplying together these three bounds proves the claim.  $\square$

The probability of a given bad walk being found in  $G_n$  is at most  $c_1((d-1)/n)^{k-f}$  by Proposition 1a. Applying Claim 13 and summing over all possible bad walks,

$$\begin{aligned} \mathbf{E}B_k^{(n)} &\leq \sum_{\chi \geq 1} \sum_{f=0}^{k-1} k^{3\chi+2}(d-1)^f n^{k-\chi-f} \frac{c_1(d-1)^{k-f}}{n^{k-f}} \\ &= \sum_{\chi \geq 1} \sum_{f=0}^{k-1} O\left(\frac{k^{3\chi+2}(d-1)^k}{n^\chi}\right) \\ &= k^3(d-1)^k \sum_{\chi \geq 1} O\left(\frac{k^{3\chi}}{n^\chi}\right) = O\left(\frac{k^6(d-1)^k}{n}\right), \end{aligned}$$

completing the proof of Proposition 12.  $\square$

**Corollary 14.**

$$d_{TV}\left(\left(\text{CNBW}_k^{(n)}, 3 \leq k \leq r\right), \left(\text{CNBW}_k^{(\infty)}, 3 \leq k \leq r\right)\right) \leq \frac{c_8 \sqrt{r}(d-1)^{3r/2-1}}{n}.$$

*Proof.* For any measurable function  $f$  and random variables  $X$  and  $Y$ , it holds that  $d_{TV}(f(X), f(Y)) \leq d_{TV}(X, Y)$ . It follows by Theorem 11 that

$$d_{TV}\left(\left(\sum_{j|k} 2jC_j^{(n)}, 3 \leq k \leq r\right), \left(\text{CNBW}_k^{(\infty)}, 3 \leq k \leq r\right)\right) \leq \frac{c_6 \sqrt{r}(d-1)^{3r/2-1}}{n}. \quad (14)$$

By the previous proposition,  $\mathbf{P}[B_k^{(n)} \geq 1] \leq c_7 k^6 (d-1)^k / n$ . Summing these probabilities from  $k = 3, \dots, r$ ,

$$\left( \sum_{j|k} 2j C_j^{(n)}, 1 \leq k \leq r \right) = (\text{CNBW}_k^{(n)}, 1 \leq k \leq r) \quad (15)$$

with probability  $1 - O(r^6 (d-1)^r / n)$ . If two random variables are equal with probability  $1 - \epsilon$ , then the total variation distance between their laws is at most  $\epsilon$ . Thus the two random vectors in (15) have total variation distance  $O(r^6 (d-1)^r / n)$ . This fact and (14) prove the corollary.  $\square$

To relate Corollary 14 to the eigenvalues of the adjacency matrix of  $G_n$ , we define a set of polynomials

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

Here  $\{T_n(x)\}_{n \in \mathbb{N}}$  are the Chebyshev polynomials of the first kind on the interval  $[-1, 1]$ , defined inductively by

$$\begin{aligned} T_0(x) &= 1, \\ T_1(x) &= x, \\ T_{n+1}(x) &= 2xT_n(x) - T_{n-1}(x), \quad n \geq 2. \end{aligned}$$

**Proposition 15** ([13, Proposition 32]). *Let  $A$  be the adjacency matrix of a (deterministic)  $d$ -regular graph  $G$ , and let  $\lambda_1 \geq \dots \geq \lambda_n$  be the eigenvalues of  $(d-1)^{-1/2}A$ . Let  $\text{CNBW}_k$  be the number of cyclically non-backtracking walks of length  $k$  in  $G$ . Then*

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

By Corollary 14, 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 note the following bounds on the eigenvalues of uniform random regular graphs.

**Proposition 16.** *Let  $G_n$  be a random  $d$ -regular graph on  $n$  vertices with eigenvalues  $\lambda_1 \geq \dots \geq \lambda_n$ . Let  $\lambda = \max_{i=2, \dots, n} |\lambda_i|$ , the maximum nontrivial eigenvalue in absolute value.*

(a) *Suppose that  $d \geq 3$  is fixed. For any  $\epsilon > 0$ ,*

$$\mathbf{P}[\lambda > 2\sqrt{d-1} + \epsilon] \rightarrow 0$$

*as  $n \rightarrow \infty$ .*

(b) Suppose that  $d = d(n)$  satisfies  $d = o(n^{1/2})$ . Then for some constant  $K$ ,

$$\mathbf{P}[\lambda > K\sqrt{d}] \leq \frac{c_9}{n^2}$$

for all  $n$ .

*Proof.* It is well known that (a) follows from the results in [15] 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 [15, 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 [17, Corollary 1.1] transfers the result to the uniform model. When  $d$  is odd (and  $n$  even, as it has to be), we apply [15, 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 [35, Corollary 4.17] transfers the result to the uniform model.

Fact (b) is proven in a more general context in [8, Lemma 18]. □

Following some facts from approximation theory, we will state the main result on the limiting distribution of linear eigenvalue statistics.

**Definition 17.** For  $\rho > 1$ , let  $E_\rho$  denote the image under the map  $z \mapsto \frac{z+z^{-1}}{2}$  of the open disc of radius  $\rho$  in the complex plane, centered at the origin. We call this the *Bernstein ellipse* of radius  $\rho$ . The ellipse has foci at  $\pm 1$ , and the sum of the major semiaxis and the minor semiaxis is exactly  $\rho$ .

**Proposition 18** ([34, Theorem 8.1]). *Suppose that  $f: [-1, 1] \rightarrow \mathbb{R}$  can be analytically extended to  $E_\rho$  and is bounded by  $M$  there. Then  $f$  has a unique expansion on  $[-1, 1]$  as*

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

and the coefficients of this expansion satisfy

$$|a_0| \leq M, \quad |a_k| \leq \frac{2M}{\rho^k}.$$

By applying the bound  $|T_k(x)| \leq 1$  and summing, we see that the approximations  $f_k(x) = \sum_{i=0}^k a_i T_i(x)$  satisfy

$$|f(x) - f_k(x)| \leq \frac{2M}{\rho^k(\rho - 1)} \tag{16}$$

for  $x \in [-1, 1]$ .

**Theorem 19.** Fix  $d \geq 3$ , and let  $G_n$  be a random  $d$ -regular graph on  $n$  vertices with adjacency matrix  $A_n$ . Let  $\lambda_1 \geq \dots \geq \lambda_n$  be the eigenvalues of  $(d-1)^{-1/2}A_n$ .

Suppose that  $f$  is a function such that  $f(2z)$  is analytic on  $E_\rho$ , where  $\rho = (d-1)^\alpha$  for some  $\alpha > 3/2$ . Then  $f(x)$  can be expanded on  $[-2, 2]$  as

$$f(x) = \sum_{k=0}^{\infty} a_k \Gamma_k(x), \tag{17}$$

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

$$Y_f := \sum_{k=1}^{\infty} \frac{a_k}{(d-1)^{k/2}} \text{CNBW}_k^{(\infty)}.$$

*Proof.* Let  $f_k(x) = \sum_{i=0}^k a_i \Gamma_i(x)$ . First, we show that  $f_k(x)$  is a good approximation to  $f(x)$ . Applying Proposition 18 to  $f(2x)$  gives the expansion (17) and shows that

$$|a_k| \leq c_{10}(d-1)^{-\alpha k} \tag{18}$$

for all  $k \geq 1$ . On any interval  $[-A, A]$  with  $A > 1$ , the maximum of  $|T_k(x)|$  occurs at the endpoints. Using a well-known expression for  $T_k(x)$ , we have

$$\max_{|x| \leq A} |T_k(x)| = \frac{(A - \sqrt{A^2 - 1})^k + (A + \sqrt{A^2 - 1})^k}{2}. \tag{19}$$

Applying this, one can see that for any  $\delta > 0$ , it is possible to choose  $\epsilon > 0$  such that for  $|x| \leq 2 + \epsilon$ ,

$$|\Gamma_k(x)| \leq (1 + \delta)^k$$

for all sufficiently large  $k$ . Choosing  $\delta$  small enough, this shows in combination with (18) that

$$\sup_{|x| \leq 2 + \epsilon} |f(x) - f_k(x)| \leq c_{11}(d-1)^{-\alpha' k} \tag{20}$$

for some  $\frac{3}{2} < \alpha' < \alpha$ . We also note that applying (18) and (19) in the same way with  $A = d/2\sqrt{d-1}$  shows that  $f_k \rightarrow f$  uniformly on  $[-d/\sqrt{d-1}, d/\sqrt{d-1}]$ , which deterministically contains all the eigenvalues of  $(d-1)^{-1/2}A_n$ .

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

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

and this is a sum of independent random variables, bounded in  $L^2$  by (18). Choose  $\beta$  satisfying  $\frac{1}{\alpha'} < \beta < \frac{2}{3}$  and define

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

$$X_f^{(n)} = \sum_{k=1}^{r_n} \frac{a_k}{(d-1)^{k/2}} \text{CNBW}_k^{(n)}.$$

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 Proposition 15. By Corollary 14 and the fact that  $\beta < \frac{2}{3}$ , the total variation distance between  $X_f^{(n)}$  and  $\sum_{k=1}^{r_n} (d-1)^{-k/2} a_k \text{CNBW}_k^{(\infty)}$  vanishes as  $n$  tends to infinity. 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

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

We have

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

As noted before,  $f_k(x) \rightarrow f(x)$  for any  $|x| \leq d/\sqrt{d-1}$ . In particular, for the deterministic top eigenvalue  $\lambda_1 = d/\sqrt{d-1}$ , we have  $f_k(\lambda_1) \rightarrow f(\lambda_1)$ . 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 (20),

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

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 16a. □

In our next theorem, we extend this theorem to the case when the degree grows with  $n$ . We will need a technical lemma on a normal approximation for the Poisson distribution:

**Lemma 20** (Lemma 19 in [27]). *Suppose  $X \sim \text{Poi}(\lambda)$  and  $W = (X - \lambda)/\sqrt{\lambda}$ . Then  $X$  can be coupled with  $Z \sim N(0, 1)$  so that  $\mathbf{E}|W - Z| \leq 1/\sqrt{\lambda}$ .*

An entire function  $f$  is said to be of order less than  $m$  if  $|f(z)| \leq |e^{z^m}|$  for all sufficiently large  $z$ . Given a test function of order less than  $m$ , our theorem gives conditions on the growth of the degree of the random regular graphs such that the eigenvalue fluctuations converge to Gaussian. The theorem also gives the limiting variance:

**Theorem 21.** *Let  $G_n$  be a random  $d_n$ -regular graph on  $n$  vertices, with  $d_n \rightarrow \infty$  as  $n \rightarrow \infty$ . Let  $A_n$  be the adjacency matrix of  $G_n$ , and let  $\lambda_1 \geq \dots \geq \lambda_n$  be the eigenvalues of  $(d_n - 1)^{-1/2} A_n$ .*

*Suppose that  $f$  is entire with order less than  $m$ , which implies that it can be expressed as  $f(x) = \sum_{k=0}^{\infty} a_k T_k(x/2)$ . If  $d_n \leq (\log n)^{\frac{2}{3m} - \epsilon}$  for some  $\epsilon > 0$ , then*

$$\sum_{i=1}^n f(\lambda_i) - \mathbf{E} \sum_{i=1}^n f(\lambda_i) \tag{21}$$

*converges in law to normal with mean zero and variance  $\frac{1}{2} \sum_{k=3}^{\infty} k a_k^2$ .*

*Proof.* Note that the sum defining the limiting variance is finite by Proposition 18. Choose  $\beta$  satisfying  $\frac{2}{3} - m\epsilon < \beta < \frac{2}{3}$ , and let

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

First, we show that the expression

$$\sum_{k=3}^{r_n} \frac{k a_k}{(d_n - 1)^{k/2}} \left( C_k^{(\infty)} - \mathbf{E} C_k^{(\infty)} \right). \tag{22}$$

converges to the desired limit. Then, we will gradually change this expression while maintaining the same limit until we arrive at (21).

Let  $\{Z_k\}_{k \in \mathbb{N}}$  be i.i.d. standard Gaussians. By Lemma 20, this collection can be coupled with  $\{C_k^{(\infty)}\}_{k \in \mathbb{N}}$  so that

$$\mathbf{E} \left| \frac{\sqrt{2k}}{(d_n - 1)^{k/2}} \left( C_k^{(\infty)} - \mathbf{E} C_k^{(\infty)} \right) - Z_k \right| \leq \frac{\sqrt{2k}}{(d_n - 1)^{k/2}}.$$

Thus the  $L^1$  distance between (22) and  $\sum_{k=3}^{r_n} \sqrt{k} a_k Z_k / \sqrt{2}$  is at most

$$\mathbf{E} \sum_{k=3}^{r_n} \frac{\sqrt{k} a_k}{\sqrt{2}} \left| \frac{\sqrt{2k}}{(d_n - 1)^{k/2}} \left( C_k^{(\infty)} - \mathbf{E} C_k^{(\infty)} \right) - Z_k \right| \leq \sum_{k=3}^{\infty} \frac{k a_k}{(d_n - 1)^{k/2}},$$

which vanishes as  $n \rightarrow \infty$ . This implies that (22) converges in law to a centered Gaussian with variance  $\sum_{k=3}^{\infty} \frac{k}{2} a_k^2$ .

Now, we present some expressions and show that they converge to the same limit.

**Expression 1:** 
$$\sum_{k=3}^{r_n} \frac{a_k}{2(d_n - 1)^{k/2}} \left( \text{CNBW}_k^{(\infty)} - \mathbf{E} \text{CNBW}_k^{(\infty)} \right)$$

The difference between Expression 1 and (22) is

$$\begin{aligned} \sum_{k=3}^{r_n} \frac{a_k}{(d_n - 1)^{k/2}} \sum_{\substack{j|k \\ j < k}} j \left( C_j^{(\infty)} - \mathbf{E}C_j^{(\infty)} \right) &\leq \sum_{j=3}^{\infty} \sum_{i=2}^{\infty} \frac{j a_{ij}}{(d_n - 1)^{ij/2}} \left( C_j^{(\infty)} - \mathbf{E}C_j^{(\infty)} \right) \\ &= \sum_{j=3}^{\infty} O\left( \frac{j a_{2j}}{(d_n - 1)^j} \right) \left( C_j^{(\infty)} - \mathbf{E}C_j^{(\infty)} \right), \end{aligned}$$

and the variance of this vanishes as  $n \rightarrow \infty$ . Thus the difference between Expression 1 and (22) converges to 0 in probability.

**Expression 2:** 
$$\sum_{k=3}^{r_n} \frac{a_k}{2(d_n - 1)^{k/2}} \left( \text{CNBW}_k^{(n)} - \mathbf{E}\text{CNBW}_k^{(\infty)} \right)$$

By Corollary 14 and our choice of  $r_n$ , the total variation distance between Expressions 1 and 2 vanishes as  $n \rightarrow \infty$ .

**Expression 3:** 
$$\sum_{k=3}^{r_n} \frac{a_k}{2(d_n - 1)^{k/2}} \left( \text{CNBW}_k^{(n)} - \mathbf{E}\text{CNBW}_k^{(n)} \right)$$

The difference between Expressions 2 and 3 is the deterministic quantity

$$\sum_{k=3}^{r_n} \frac{a_k}{2(d_n - 1)^{k/2}} \left( \mathbf{E}\text{CNBW}_k^{(\infty)} - \mathbf{E}\text{CNBW}_k^{(n)} \right). \quad (23)$$

Using the decomposition

$$\text{CNBW}_k^{(n)} = \sum_{j|k} 2j C_j^{(n)} + B_k^{(n)}$$

from Proposition 12, we have

$$\mathbf{E}\text{CNBW}_k^{(\infty)} - \mathbf{E}\text{CNBW}_k^{(n)} = \sum_{j|k} \left( (d-1)^j - 2j \mathbf{E}C_j^{(n)} \right) - \mathbf{E}B_k^{(n)}.$$

By [26, eq. (2.2)] and Proposition 12, this is  $O(k^6(k+d)(d-1)^k/n)$ . By our choice of  $r_n$  and the fact that  $a_k \rightarrow 0$  as  $k \rightarrow \infty$ , equation (23) vanishes as  $n \rightarrow \infty$ .

**Expression 4:** 
$$\sum_{i=1}^n f(\lambda_i) - \mathbf{E} \sum_{i=1}^n f(\lambda_i)$$

Let  $f_k := \sum_{i=0}^k a_k T_k(x/2)$ . By Proposition 15 and the fact that  $\lambda_1$  is deterministic, Expression 3 is equal to

$$\sum_{i=2}^n f_{r_n}(\lambda_i) - \mathbf{E} \sum_{i=2}^n f_{r_n}(\lambda_i).$$

Thus it suffices to show that  $\sum_{i=2}^n (f(\lambda_i) - f_{r_n}(\lambda_i))$  vanishes in  $L^1$ .

Let  $E$  be the event that  $\sup_{i=2, \dots, n} |\lambda_i| \leq K$ , where  $K$  is the constant from Proposition 16b that makes  $\mathbf{P}[E^C] \leq c_9 n^{-2}$ . We have

$$\mathbf{E} \left| \sum_{i=2}^n (f(\lambda_i) - f_{r_n}(\lambda_i)) \right| \leq \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3,$$

where

$$\mathcal{E}_1 := \mathbf{E} \left[ \mathbf{1}_E \sum_{i=2}^n |f(\lambda_i) - f_{r_n}(\lambda_i)| \right],$$

$$\mathcal{E}_2 := \mathbf{E} \left[ \mathbf{1}_{E^C} \sum_{i=2}^n |f(\lambda_i)| \right],$$

$$\mathcal{E}_3 := \mathbf{E} \left[ \mathbf{1}_{E^C} \sum_{i=2}^n |f_{r_n}(\lambda_i)| \right],$$

and we need to show that these quantities vanish as  $n \rightarrow \infty$ .

By Proposition 18,

$$|a_k| \leq 2 \left[ \sup_{z \in E_\rho} f(2z) \right] \rho^{-k} \leq 2 \exp \left[ \left( \rho + \frac{1}{\rho} \right)^m \right] \rho^{-k} \quad (24)$$

for sufficiently large  $\rho$ . By (19),  $\sup_{x \in [-K, K]} |T_k(x/2)| = O(K^k)$ . This gives us

$$\sup_{x \in [-K, K]} |f(x) - f_k(x)| \leq \sum_{i=k+1}^{\infty} |a_i T_i(x/2)| = O \left( \exp \left[ \left( \rho + \frac{1}{\rho} \right)^m \right] \left( \frac{K}{\rho} \right)^{k+1} \right).$$

Set  $\rho = k^{1/m}$  to approximately optimize this, and substitute  $k = r_n$  to get

$$\sup_{x \in [-K, K]} |f(x) - f_{r_n}(x)| \leq O(1) \exp \left( \frac{\beta \log n}{\log(d_n - 1)} \left( O(1) - \frac{\log \log n - \log \log(d_n - 1)}{m} \right) \right).$$

By the condition  $d_n \leq (\log n)^{\frac{2}{3m} - \epsilon}$ ,

$$\begin{aligned} \sup_{x \in [-K, K]} |f(x) - f_{r_n}(x)| &\leq O(1) \exp \left( \frac{\beta \log n}{\log(d_n - 1)} \right. \\ &\quad \cdot \left. \left( O(1) - \frac{\frac{1}{\frac{2}{3m} - \epsilon} \log(d_n - 1) - \log \log(d_n - 1)}{m} \right) \right) \\ &= O(1) \exp \left( O \left( \frac{\beta \log n \log \log(d_n - 1)}{m \log(d_n - 1)} \right) - \frac{\beta}{\frac{2}{3} - m\epsilon} \log n \right) \\ &= O(1) \exp \left( \log n \left( o(1) - \frac{\beta}{\frac{2}{3} - m\epsilon} \right) \right). \end{aligned}$$



Since  $\beta > \frac{2}{3} - m\epsilon$ , this expression is  $o(n^{-1})$ , and

$$\mathcal{E}_1 \leq (n-1) \sup_{x \in [-K, K]} |f(x) - f_{r_n}(x)| \rightarrow 0$$

as  $n \rightarrow \infty$ .

Next, we show that  $\mathcal{E}_2 \rightarrow 0$ . For large enough  $d_n$ , it holds that if  $|x| \leq d_n(d_n - 1)^{-1/2}$ , then

$$|f(x)| \leq \exp\left(\frac{d_n^m}{(d_n - 1)^{m/2}}\right) = \exp(O(d_n^{m/2})) = \exp(O((\log n)^{1/3})).$$

As  $E^C$  occurs with probability at most  $c_9 n^{-2}$ ,

$$\mathcal{E}_2 \leq O(n^{-2})(n-1) \exp(O((\log n)^{1/3})) \rightarrow 0$$

as  $n \rightarrow \infty$ .

Last, we consider  $\mathcal{E}_3$ . We apply (24) with  $\rho = k^{1/m}$  to show that for some  $C$  depending on  $m$  but not  $k$ ,

$$|a_k| \leq C^k k^{-k/m}$$

for all  $k \geq 1$ . By (19), for all  $|x| \leq d_n(d_n - 1)^{-1/2}$ ,

$$|T_k(x/2)| \leq d_n^{k/2}.$$

Thus

$$f_{r_n}(x) \leq a_0 + \sum_{k=1}^{r_n} \left(\frac{C d_n^{1/2}}{k^{1/m}}\right)^k \leq a_0 + \sum_{k=1}^{\infty} \left(\frac{C(\log n)^{\frac{1}{3m}}}{k^{1/m}}\right)^k.$$

Let  $N_n = \lfloor (2C)^m (\log n)^{1/3} \rfloor$ , and break the sum into two pieces, one from 1 to  $N_n$  and the other from  $N_n + 1$  to  $\infty$ . Each term in the first piece is at most  $C^{N_n} (\log n)^{N_n/3m}$ , and a bit of analysis shows that

$$N_n C^{N_n} (\log n)^{N_n/3m} = o(n).$$

The second piece is  $o(1)$ , as can be seen by comparing it to a geometric series. Thus

$$\mathcal{E}_3 \leq O(n^{-2})(n-1)o(n) \rightarrow 0$$

as  $n \rightarrow \infty$ . □

*Remark 22.* The only difference between the limiting distributions of Theorems 19 and 21 and those of the permutation model of random graph in [13] derives from the slightly different expectations of  $\text{CNBW}_k^{(\infty)}$  in the two models, and from the fact that  $\text{CNBW}_1^{(\infty)} = \text{CNBW}_2^{(\infty)} = 0$  in the uniform model. The limiting variance in Theorem 21 is the same as for eigenvalue fluctuations of the GOE, except that the coefficients  $a_1$  and  $a_2$  are ignored. (As the variance term for the GOE fluctuations can be expressed in many different ways, this is not entirely obvious. See Section 1.5 and in particular Proposition 3 from [27].)

## Acknowledgments

The author gratefully acknowledges Ioana Dumitriu for pointing out similarities between [13] and [26], and Soumik Pal and Elliot Paquette for their general assistance.

## References

- [1] Zhidong Bai and Jack W. Silverstein. *Spectral analysis of large dimensional random matrices*. Springer Series in Statistics. Springer, New York, second edition, 2010.
- [2] Zhidong Bai and Jianfeng Yao. On the convergence of the spectral empirical process of Wigner matrices. *Bernoulli*, 11(6):1059–1092, 2005.
- [3] 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.
- [4] Gérard Ben Arous and Kim Dang. On fluctuations of eigenvalues of random permutation matrices. Preprint. Available at arXiv:1106.2108, 2011.
- [5] 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.
- [6] Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, 1(4):311–316, 1980.
- [7] 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.
- [8] 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.
- [9] Sourav Chatterjee. Stein’s method for concentration inequalities. *Probab. Theory Related Fields*, 138(1-2):305–321, 2007.
- [10] Sourav Chatterjee, Persi Diaconis, and Elizabeth Meckes. Exchangeable pairs and Poisson approximation. *Probab. Surv.*, 2:64–106 (electronic), 2005.
- [11] Louis H. Y. Chen. Poisson approximation for dependent trials. *Ann. Probability*, 3(3):534–545, 1975.
- [12] Persi Diaconis and Bernd Sturmfels. Algebraic algorithms for sampling from conditional distributions. *Ann. Statist.*, 26(1):363–397, 1998.
- [13] 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.
- [14] Ioana Dumitriu and Soumik Pal. Sparse regular random graphs: Spectral density and eigenvectors. *Ann. Probab.*, 40(5):2197–2235, 2012.

- [15] Joel Friedman. A proof of Alon’s second eigenvalue conjecture and related problems. *Mem. Amer. Math. Soc.*, 195(910):viii+100, 2008.
- [16] Chris D. Godsil and Brendan D. McKay. Asymptotic enumeration of Latin rectangles. *J. Combin. Theory Ser. B*, 48(1):19–44, 1990.
- [17] Catherine Greenhill, Svante Janson, Jeong Han Kim, and Nicholas C. Wormald. Permutation pseudographs and contiguity. *Combin. Probab. Comput.*, 11(3):273–298, 2002.
- [18] 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.
- [19] 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.
- [20] Tobias Johnson and Soumik Pal. Cycles and eigenvalues of sequentially growing random regular graphs. *Ann. Probability*, 42(4):1396–1437, 2014.
- [21] Jeong Han Kim, Benny Sudakov, and Van Vu. Small subgraphs of random regular graphs. *Discrete Math.*, 307(15):1961–1967, 2007.
- [22] 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.
- [23] Brendan D. McKay. The expected eigenvalue distribution of a large regular graph. *Linear Algebra Appl.*, 40:203–216, 1981.
- [24] 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.
- [25] 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).
- [26] 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.
- [27] Elliot Paquette. *Eigenvalue Fluctuations of Random Matrices beyond the Gaussian Universality Class*. PhD thesis, University of Washington, 2013.
- [28] Nathan Ross. Fundamentals of Stein’s method. *Probab. Surv.*, 8:210–293, 2011.
- [29] Yakov Sinai and Alexander Soshnikov. Central limit theorem for traces of large random symmetric matrices with independent matrix elements. *Bol. Soc. Brasil. Mat. (N.S.)*, 29(1):1–24, 1998.
- [30] 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.
- [31] Charles Stein. A way of using auxiliary randomization. In *Probability theory (Singapore, 1989)*, pages 159–180. de Gruyter, Berlin, 1992.
- [32] Charles M. Stein. Asymptotic evaluation of the number of Latin rectangles. *J. Combin. Theory Ser. A*, 25(1):38–49, 1978.
- [33] Linh V. Tran, Van H. Vu, and Ke Wang. Sparse random graphs: Eigenvalues and eigenvectors. *Random Structures Algorithms*, 42(1):110–134, 2013.
- [34] Lloyd N. Trefethen. *Approximation theory and approximation practice*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2013.
- [35] 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.
- [36] Nicholas C. Wormald. The asymptotic distribution of short cycles in random regular graphs. *J. Combin. Theory Ser. B*, 31(2):168–182, 1981.
- [37] Nicholas C. Wormald. The perturbation method and triangle-free random graphs. Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995). *Random Structures Algorithms*, 9(1-2):253–270, 1996.