

# GLOBAL EIGENVALUE FLUCTUATIONS OF RANDOM BIREGULAR BIPARTITE GRAPHS

IOANA DUMITRIU AND YIZHE ZHU

**ABSTRACT.** We compute the eigenvalue fluctuations of uniformly distributed random biregular bipartite graphs with fixed and growing degrees for a large class of analytic functions. As a key step in the proof, we obtain a total variation distance bound for the Poisson approximation of the number of cycles and cyclically non-backtracking walks in random biregular bipartite graphs, which might be of independent interest. We also prove a semicircle law for random  $(d_1, d_2)$ -biregular bipartite graphs when  $\frac{d_1}{d_2} \rightarrow \infty$ . As an application, we translate the results to adjacency matrices of uniformly distributed random regular hypergraphs.

## 1. INTRODUCTION

**1.1. Eigenvalue fluctuations of random matrices.** The study of fluctuations from the limiting empirical spectral distributions (ESDs) for random matrices is a well-established topic of interest in random matrix theory, originated in [41, 43, 58], see also [2] and all references therein. More recently, it has been extended to sparse random matrices and random graph-related matrices in various regimes of sparsity and independence ([55, 56, 10, 30, 9]), and the natural next target is hypergraphs ([34]).

The ultimate goal in these studies is to see the equivalent of the one-dimensional Central Limit Theorem (CLT) emerge, when examining linear statistics of the spectra of random matrices and random graphs. More precisely, denote by  $\lambda_1, \dots, \lambda_n$  the eigenvalues of the random matrix, suitably scaled to put them with high probability on a compact set, and let  $f$  be a suitably smooth function. When the matrices in question are not extremely sparse, one can almost invariably prove that the linear statistic

$$\mathcal{L}(f) = \sum_{i=1}^n f(\lambda_i)$$

has the property that, when centered, it converges to a normal distribution whose variance depends on  $f$ :

$$\mathcal{L}(f) - \mathbb{E}(\mathcal{L}(f)) \rightarrow N(0, \sigma_f^2).$$

**1.1.1. Dense and not-too-sparse Wigner cases.** There is an interesting phenomenon taking place with respect to sparsity; the variance  $\sigma_f^2$  is the same in the case of Gaussian Orthogonal Ensembles (GOEs) as in the case of the random regular graph under the permutation model with growing degrees [30]:

$$(1.1) \quad \sigma_f^2 = 2 \sum_{k=1}^{\infty} k a_k^2,$$

---

*Date:* January 19, 2023.

2000 *Mathematics Subject Classification.* Primary 60C05, 60B20; Secondary 05C50.

*Key words and phrases.* random biregular bipartite graph, random regular hypergraph, switching, non-backtracking walk, eigenvalue fluctuations.

This work was partially supported by NSF DMS-1949617.

where  $a_k$  is the  $k$ -th coefficient of  $f$  in the Chebyshev polynomial basis expansion. Small variations of this expression also occur in dense Wigner variants and the uniform regular graph model, as follows. For real Wigner and generalized Wigner matrices in the dense case [39, 6, 3, 22, 60],  $\sigma_f$  also depends on the fourth moments of the off-diagonal entries and the variance of the diagonal entries, which yields corrections to the constants in front of  $a_1^2, a_2^2$  (see Theorem 1.1 in [6] for an explicit expression). Similarly, in the uniform regular graph model [40], a correction must be introduced as there are no one- or two-cycles (as the graph is simple), and so the terms corresponding to  $k = 1$  and  $k = 2$  in the sum (1.1) are not present.

However, in the case of sparse Wigner matrices (corresponding to Erdős-Rényi graphs  $G(n, p)$  with  $p \rightarrow 0, np \rightarrow \infty$ , [56]), the fluctuations are impacted by the fact that the number of nonzero entries in each row (i.e., the degree of each vertex) fluctuates, and the 4th moment of the scaled adjacency matrix entries grows. The variance  $\sigma_f^2$  blows up, necessitating another multiplicative scaling of the linear statistic  $\mathcal{L}(f) - \mathbb{E}(\mathcal{L}(f))$  by  $\sqrt{p}$ , and extracting only part of the expression (1.1) (see Theorem 1 in [56]).

**1.1.2. Dense Wishart cases.** A similar phenomenon occurs in the Wishart case, i.e., for sample covariance matrices (corresponding to bipartite graphs); in the case of dense matrices with converging aspect ratio, the variance is given in different forms in [4, 5]. Although these expressions are not explicit in terms of a Chebyshev polynomial expansion, in [19, 44], it is shown that the covariance between two linear statistics is diagonalized by shifted Chebyshev polynomials. When the aspect ratio goes to  $\infty$ , [24] computes the variance which is consistent with the Wigner case in [6]. So far, we are not aware of any CLT results for sparse bipartite Erdős-Rényi graphs, but a similar argument as in [56] should apply.

For dependent entries (biregular bipartite graphs), we obtain here the variance of the eigenvalue fluctuation in Theorem 4.7, and it matches the one in [24], except for the first coefficient.

**1.1.3. Constant (expected or deterministic) degree.** When  $p = \frac{c}{n}$ , the explicit limiting spectral distribution for Erdős-Rényi graphs  $G(n, p)$  is not known, although it is known that the measure  $\mu_c$  exists for every  $c$  (given, e.g., by a Stieltjes transform equation as in [14]), and if  $c > 1$  it consists of a continuous part and an atomic part [15]. Convergence of  $\mu_c$  to the semicircular distribution is studied in [33, 42], where asymptotic expressions for the moments of  $\mu_c$  with an  $o(1/c)$  term are computed (as  $c \rightarrow \infty$ ,  $\mu_c$  converges to the semicircle law).

However, a CLT for Erdős-Rényi graphs  $G(n, \frac{c}{n})$  still holds [55, 10] with a more complicated variance that does not follow the same expression as in (1.1), see Theorem 2.2 in [10]. By contrast, in the random  $d$ -regular graph case with  $d$  finite, the fluctuations are no longer Gaussian. Instead, they are modeled by an infinitely divisible distribution, expressed as a sum of Poisson variables (see [30] for the permutation model, and [40, 50] for the uniform model). Notably, in the case when the matrix is not symmetric and corresponds to the (directed) cycle structure of a random permutation, [9] showed that the global fluctuations could be computed, and whether or not the limiting distribution is Gaussian depends on how smooth the test function is. For random regular graphs with fixed degree  $d$ , we will not see the effect of degree fluctuation. But for Erdős-Rényi graphs  $G(n, \frac{d}{n})$  with expected degree  $d$ , the degree fluctuation contributes to the fluctuation of linear statistics, and there is an extra  $n^{-1/2}$  normalization in the CLT (see [55, Theorem 4]). Such a difference shows that eigenvalues of random  $d$ -regular graphs for fixed  $d$  are more rigid than the corresponding Erdős-Rényi graph  $G(n, \frac{d}{n})$ .

For the bipartite Erdős-Rényi case, once again, the limiting distribution is not known, but results that are similar to [33] can be found in [53]. We are not aware of any CLT-like results for the fluctuations in this case. We compute the fluctuations for the uniformly random biregular bipartite with fixed degrees. Just like in the regular case [30], we see that the fluctuations are modeled by a sum of Poisson variables (Theorem 4.4).

Another important class of random graphs is the configuration model. The ESD of the configuration model with a large mean degree is not generally given by the semicircle law [28, 51], but no linear statistics result is known. When the degrees grow with the size of the graph, similar to Erdős-Rényi graphs and random regular graphs, we expect a central limit theorem for linear statistics holds. There are linear statistics results for inhomogeneous matrix models beyond Wigner matrices [22, 1], and it might be possible to apply their techniques, together with a coupling argument introduced in [28], to study the linear statistics for the configuration model.

**1.2. Random biregular bipartite graphs.** Biregular bipartite graphs have found applications in error correcting codes, matrix completion, and community detection, see for example [38, 61, 59, 36, 16, 18, 11, 26, 27]. An  $(n, m, d_1, d_2)$ -biregular bipartite graph is a bipartite graph  $G = (V_1, V_2, E)$  where  $|V_1| = n, |V_2| = m$  and every vertex in  $V_1$  has degree  $d_1$  and every vertex in  $V_2$  has degree  $d_2$ . Here we must have  $nd_1 = md_2 = |E|$ . When the number of vertices is clear, we call it a  $(d_1, d_2)$ -biregular bipartite graph for simplicity.

Let  $X \in \{0, 1\}^{n \times m}$  be a matrix indexed by  $V_1 \times V_2$  such that  $X_{ij} = 1$  if and only if  $(i, j) \in E$ . The adjacency matrix of a  $(d_1, d_2)$ -biregular bipartite graph with  $V_1 = [n], V_2 = [m]$  can be written as

$$(1.2) \quad A = \begin{bmatrix} 0 & X \\ X^\top & 0 \end{bmatrix}.$$

All eigenvalues of  $A$  come in pairs as  $\{-\lambda, \lambda\}$ , where  $|\lambda|$  is a singular value of  $X$ , along with extra  $|n - m|$  zero eigenvalues. It's easy to see  $\lambda_1(A) = -\lambda_{n+m}(A) = \sqrt{d_1 d_2}$ .

The empirical spectral distribution for uniformly distributed random biregular bipartite graphs (RBBGs), which is the equivalent of the Kesten-McKay law, was first computed in [37, 52] for the case of fixed  $d_1, d_2$ . For growing degrees, when  $\frac{d_1}{d_2}$  converges to a positive constant, the analogue to the Marčenko-Pastur law was proved in [29, 63, 65].

In this paper, instead of examining the spectrum of  $A$ , we will be looking at the spectrum of the matrix  $XX^\top - d_1 I$ . This serves two purposes: one, it allows for an immediate parallel to the sample covariance matrix (Wishart) case, and two, it allows us to deal with all regimes in a unitary fashion. The eigenvalues of  $XX^\top - d_1 I$  are the shifted squares of the eigenvalues of  $A$ . Any result on global fluctuations for linear statistics of the spectrum of  $XX^\top - d_1 I$  is automatically converted into an equivalent result for the spectrum of  $A$ . However, because any result of fluctuations must necessarily put most of the eigenvalues (with the exception of the deterministic outliers) on a compact interval, scaling must be involved. This works perfectly fine when the ratio  $d_1/d_2$  is bounded, but it becomes tricky when it is not, and the matrix  $XX^\top - d_1 I$  allows us to do the scaling in a more natural way, similarly to the sample covariance (Wishart) matrix with unbounded aspect ratio in [24].

To prove a result on eigenvalue fluctuations, we need two special ingredients: *eigenvalue confinement on a compact interval* and *asymptotic behavior of cycle counts*. For the former, we make use of the spectral gap shown in [16] for the fixed degree case and [66] for the growing degree case. Previous results of this kind were obtained for random regular graphs [35, 13] for a fixed degree, and [17, 25, 62, 8] for growing degrees.

For the latter, we use Stein's method to approximate cycle counts as Poisson random variables by bounding the total variation distance (Theorem 2.10) and obtain a Poisson approximation of the number of cyclically non-backtracking walks (Corollary 2.15). Note that computing cycle counts is a fundamental problem in the study of random graphs, ever since the seminal papers of [47] and more general [48, 49].

To prove our results, we follow the recipe of [40] by using switching to construct exchangeable pairs of graphs that allow us to estimate cycle counts. The switching we use here differs from [40] and is suitable for biregular bipartite graphs. In the analysis of switchings, a new challenge is an imbalance between the parameters  $d_1, d_2$  when the aspect ratio is unbounded. Our results on cycle

counts hold for a large range of  $d_1, d_2$ , and are notably independent of the aspect ratio as long as the cycle length is small. It is also worth noting that the method of switching has been applied to other problems on random biregular bipartite graphs, for example, [21, 20, 54].

Finally, we also obtain an algebraic relation between linear eigenvalue statistics on modified Chebyshev polynomials and cyclically non-backtracking walks (Theorem 2.17). Then based on the spectral gap results in [16, 66] and approximation theory for Chebyshev polynomials [64], we extend the eigenvalue fluctuation results to a general class of analytic functions.

**1.3. Main results.** Our main contributions are represented by Theorems 4.4, 4.7, establishing the behavior of the global fluctuations for the linear statistics of eigenvalues of RBBGs in the fixed  $d_1, d_2$ , respectively, in the  $d_1 \cdot d_2 \rightarrow \infty$  cases. Note that Theorem 4.7 describes the behavior of the fluctuations even in the case when the limiting ESD does not exist since it merely requires  $d_1/d_2$  to be bounded, rather than to converge to a number in  $[1, \infty)$  (which would be the necessary condition for the ESD to converge). In addition, we show that the covariance between two linear statistics with different test functions is given by the coefficients in their Chebyshev expansions.

As part of the proofs for our main results, we also describe the asymptotic behavior of the cycle counts (Theorem 2.10). Based on the cycle counts estimates, we then use the locally tree-like structure of RBBGs to prove a global semicircle law in the case when the degree goes slowly ( $d_1 = n^{o(1)}$ ) and  $d_1/d_2$  is unbounded (Theorem 5.3).

Finally, as an important application, we obtain equivalent results for uniformly distributed random regular hypergraphs, including cycle counts, global laws, spectral gaps, and eigenvalue fluctuations.

**1.4. Organization of the paper.** In Section 2 we prove our results on cycle counts in random biregular bipartite graphs. Section 3 collects relevant results for the spectral gap and eigenvalue confinement on a compact interval from the literature. Section 4 proves our main results, Theorems 4.4 and 4.7. Section 5 proves a global semicircle law for RBBGs when  $d_1/d_2$  is unbounded. In Section 6, we use the connections established in [32] to prove several results on uniformly distributed regular hypergraphs.

## 2. CYCLE COUNTS

**2.1. Counting switchings.** In this section, we estimate the number of switchings that create or delete a cycle in a biregular bipartite graph. The precise definitions of switchings for our purposes are given in Definition 2.4 and Definition 2.5. These estimates will be used in Section 2.2 to show that cycle counts converge in distribution to Poisson random variables.

**Definition 2.1** (cycle). Throughout the paper, when we say a *cycle*, we mean a *simple cycle*, i.e., all vertices in a cycle are distinct.

Let  $K_{n,m}$  be the complete bipartite graph on  $n + m$  vertices with  $V_1 = [n], V_2 = [m]$ . Let  $H \subseteq K_{n,m}$  be a subgraph with  $v$  vertices. For any  $i \in K_{n,m}$ , let  $g_i, h_i$  denote the degree of  $i$  considered as a vertex in a biregular bipartite graph  $G = (V_1, V_2, E)$  and the subgraph  $H$ , respectively. Let  $h_{\max}$  be the largest value of  $h_i$  and  $|H|$  be the number of edges of  $H$ . Denote by

$$[x]_a = x(x-1) \cdots (x-a+1)$$

the falling factorial. The following estimate is given in [48].

**Proposition 2.2** (Theorem 3.5 in [48]). *Assume  $d_1 \geq d_2$  and  $nd_1 \geq 2d_1(d_1 + h_{\max} - 2) + |H| + 1$ . Then*

$$\mathbb{P}(H \subseteq G) \leq \frac{\prod_{i=1}^v [g_i]_{h_i}}{[nd_1 - 4d_1^2 - 1]_{|H|}}.$$

We first prove several estimates on random biregular bipartite graphs based on Proposition 2.2.

**Lemma 2.3.** *Let  $G$  be a random  $(d_1, d_2)$ -biregular bipartite graph with  $d_2 \leq d_1 \leq n^{1/3}$ .*

(1) *Suppose  $H$  is a subgraph of the complete graph  $K_{n,m}$  in which every vertex has degree at least 2. Let  $e$  be the number of edges in  $H$ . Suppose  $e = o(n^{1/3})$ . Then*

$$(2.1) \quad \mathbb{P}(H \subseteq G) \leq c_1 \left( \frac{(d_1 - 1)(d_2 - 1)}{nm} \right)^{e/2}.$$

(2) *Let  $\alpha$  be a cycle of length  $2k$  in the complete bipartite graph  $K_{n,m}$ . Suppose  $k \leq n^{1/10}$ , then*

$$(2.2) \quad \mathbb{P}(\alpha \subseteq G) \leq c_1 \left( \frac{(d_1 - 1)(d_2 - 1)}{nm} \right)^k.$$

(3) *Let  $\beta$  be another cycle of length  $2j \leq 2n^{1/10}$  in the complete bipartite graph  $K_{n,m}$ . Suppose  $\alpha, \beta$  share  $f$  edges. Then*

$$(2.3) \quad \mathbb{P}(\alpha \cup \beta \subseteq G) \leq c_1 \left( \frac{(d_1 - 1)(d_2 - 1)}{nm} \right)^{j+k-f/2}.$$

*Proof.* It suffices to prove (2.1). Then (2.2) and (2.3) follow as special cases. Since  $H$  has  $e$  edges, and  $H$  is bipartite, it satisfies

$$\sum_{i \in V_1} h_i = \sum_{i \in V_2} h_i = e.$$

Since  $h_i \geq 2$  for all  $i \in V(H)$ , we know  $[g_i]_{h_i} \leq (g_i(g_i - 1))^{h_i/2}$ . Therefore from Proposition 2.2,

$$\begin{aligned} \mathbb{P}(H \subseteq G) &\leq \frac{(d_1(d_1 - 1))^{e/2}(d_2(d_2 - 1))^{e/2}}{[nd_1 - 4d_1^2 - 1]_e} \\ &= \left( \frac{(d_1 - 1)(d_2 - 1)}{nm} \right)^{e/2} \frac{(nd_1)^e}{[nd_1 - 4d_1^2 - 1]_e}. \end{aligned}$$

Recall  $d_1 \leq n^{1/3}$ ,  $e = o(n^{1/3})$ , and  $(1 + x)^r = 1 + O(rx)$  if  $rx \rightarrow 0$ . We have for some absolute constant  $c_1 > 0$ ,

$$(2.4) \quad \frac{(nd_1)^e}{[nd_1 - 4d_1^2 - 1]_e} \leq \left( \frac{nd_1}{nd_1 - 4d_1^2 - e} \right)^e = \left( 1 + \frac{4d_1^2 + e}{nd_1 - 4d_1^2 - e} \right)^e \leq c_1.$$

This proves (2.1).  $\square$

Let  $G$  be a  $(d_1, d_2)$ -biregular bipartite graph. Let  $C_j$  be the number of cycles of length  $2j$  in  $G$ . We will always represent a cycle by a vertex sequence starting from a vertex in  $V_1$ . Suppose  $\alpha = (x_1, y_1, \dots, x_k, y_k)$  is a cycle of length  $2k$  in  $G$  with  $x_i \in V_1, y_i \in V_2, 1 \leq i \leq k$ , where  $y_k$  is connected to  $x_1$  in the cycle  $\alpha$ .

Let  $e_i = u_i v_i, e'_i = u'_i v'_i$  be the edges with  $u_i, u'_i \in V_1, v_i, v'_i \in V_2, 1 \leq i \leq k$  such that neither  $u_i, u'_i$  is adjacent to  $y_i$  for  $1 \leq i \leq k$  and neither  $v_i, v'_i$  is adjacent to  $x_i$ . See the left part of Figure 1 for an example.

We now introduce our definitions of switching for biregular bipartite graphs.

**Definition 2.4** (forward  $\alpha$ -switching). Consider the action of deleting all  $4k$  edges  $\tilde{e}_i, 1 \leq i \leq 2k$  and  $e_i, e'_i, 1 \leq i \leq k$ , and replacing them by the edges  $x_i v_i, x_i v'_i, y_i u_i, y_i u'_i$  for  $1 \leq i \leq k$ . We obtain a new biregular bipartite graph  $G'$  with the cycle  $\alpha$  deleted. We call this action induced by the 6 sequences  $(x_i), (y_i), (u_i), (u'_i), (v_i), (v'_i), 1 \leq i \leq k$  a *forward  $\alpha$ -switching*. See Figure 1 for an example. We will consider forward  $\alpha$ -switchings only up to cyclic rotation and inversion of indices in  $[k]$ ; that is, we identify the  $2k$  different forward  $\alpha$ -switchings obtained by applying the same cyclic rotation or inversion on  $[k]$  to the 6 sequences  $(x_i), (y_i), (u_i), (u'_i), (v_i), (v'_i), 1 \leq i \leq k$ .

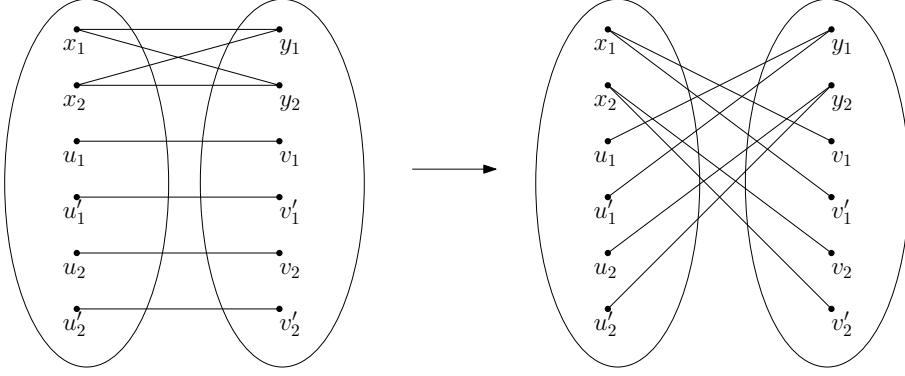


FIGURE 1. A forward  $\alpha$  switching from the left to the right, where  $\alpha = (x_1, y_1, x_2, y_2)$ .

**Definition 2.5** (backward  $\alpha$ -switching). Suppose  $G$  contains paths  $v_i x_i v'_i$  and  $u_i y_i u'_i$  for  $1 \leq i \leq k$ , where  $x_i, u_i, u'_i \in V_1, y_i, v_i, v'_i \in V_2$ . Consider deleting all  $4k$  edges  $v_i x_i, v'_i x_i, u_i y_i, u'_i y_i$  for  $1 \leq i \leq k$ , and replacing them with  $u_i v_i, u'_i v'_i, x_i y_i, y_i x_{i+1}$  for  $1 \leq i \leq k$ . We obtain a new graph  $G'$  with a cycle  $\alpha = (x_1, y_1, \dots, x_k, y_k)$ . Such action is called a *backward  $\alpha$ -switching* induced by the sequences  $(x_i), (y_i), (u_i), (u'_i), (v_i), (v'_i), 1 \leq i \leq k$ . We also identify the  $2k$  different backward  $\alpha$ -switchings obtained by applying the same cyclic rotation or inversion on the index set  $[k]$ .

**Definition 2.6** (short cycles). Let  $r$  be an integer; we say that a cycle is *short* if its length is less than or equal to  $2r$ .

We call a  $\alpha$ -switching *valid* if  $\alpha$  is the only short cycle created or destroyed by the switching. For each forward  $\alpha$ -switching from  $G$  to  $G'$ , there is a corresponding backward  $\alpha$ -switching from  $G'$  to  $G$  by simply reversing the operation (i.e. from right to left in Figure 1).

Let  $F_\alpha$  be the number of all valid forward  $\alpha$ -switchings from  $G$  to some  $G'$  and let  $B_\alpha$  be the number of all valid backward  $\alpha$ -switchings from some  $G'$  to  $G$ . In the following two lemmas, we estimate  $F_\alpha$  and  $B_\alpha$  for biregular bipartite graphs.

**Lemma 2.7.** Let  $G$  be a deterministic  $(d_1, d_2)$ -biregular bipartite graph with  $d_1 \geq d_2$  and cycle counts  $C_k, 2 \leq k \leq r$ . For any short cycle  $\alpha \subseteq G$  of length  $2k$ , we have

$$(2.5) \quad F_\alpha \leq [n]_k [m]_k d_1^k d_2^k.$$

If  $\alpha$  does not share an edge with another short cycle, then for an absolute constant  $c_1 > 0$ , we have

$$(2.6) \quad F_\alpha \geq [n]_k [m]_k d_1^k d_2^k \left( 1 - \frac{4k \sum_{j=2}^r j C_j + c_1 k (d_1 - 1)^r (d_2 - 1)^r}{nd_1} \right).$$

*Proof.* Consider a cycle denoted by  $\alpha = (x_1, y_1, \dots, x_k, y_k)$ . Denote edges

$$(2.7) \quad \tilde{e}_i = x_i y_i, \quad \tilde{e}_{i+k} = y_i x_{i+1}, \quad 1 \leq i \leq k,$$

where  $x_{k+1} := x_1$ . There are at most  $[n]_k d_1^k [m]_k d_2^k$  many ways to choose edges  $e_i = u_i v_i$  and  $e'_i = u'_i v'_i$  for  $1 \leq i \leq k$ , which gives the upper bound (2.5). For the  $k$  edges  $e_i, 1 \leq i \leq k$ , we require distinct  $u_i \in V_1, 1 \leq i \leq k$ , and we have  $d_1$  choices for each  $v_i$ , given the degree constraint on  $u_i$ . This gives  $[n]_k d_1^k$  many choices altogether. For the remaining edges  $e'_i, 1 \leq i \leq k$  we require distinct  $v'_i \in V_2, 1 \leq i \leq k$  and each for each  $u'_i$  we have  $d_2$  choices, giving us a factor of  $[m]_k d_2^k$ . Therefore (2.5) holds.

For the rest of the proof, we always use the same way to count  $\alpha$ -switchings by counting the choices of  $e_i, e'_i$ . We use the parameter  $d_1$  to control the choices from  $e_i, 1 \leq i \leq k$  and the parameter  $d_2$  for the choices from  $e'_i, 1 \leq i \leq k$ .

To prove the lower bound in (2.6), we choose a subset of configurations that are guaranteed to have a valid forward  $\alpha$ -switching. Consider  $e_i, e'_i, 1 \leq i \leq k$  such that the following holds:

- (1)  $e_i$  and  $e'_i$  are not contained in any short cycle in  $G$  for  $1 \leq i \leq k$ .
- (2) The distance from any vertex in  $\{e_i, e'_i\}$  to any vertex in  $\tilde{e}_i$  is at least  $2r$  for any  $1 \leq i \leq k$ .
- (3) The distance between any two different edges among the  $2k$  edges  $\{e_i, e'_i, 1 \leq i \leq k\}$  is at least  $r$ .
- (4) For all  $1 \leq i \leq k$ , the distance between  $v_i$  and  $v'_i$  is at least  $2r$ , and the distance between  $u_i$  and  $u'_i$  is at least  $2r$ .

Recall the definition of  $\tilde{e}_i$  in (2.7). By Condition (2), for all  $1 \leq i \leq k$ ,  $u_i, u'_i$  are not adjacent to  $y_i$ , also  $v_i, v'_i$  are not adjacent to  $x_i$ , which satisfies the definition of a forward  $\alpha$ -switching. Let  $G'$  be the graph obtained by applying the forward  $\alpha$ -switching from  $G$ . We need to check that  $\alpha$  is the only cycle deleted in  $G$  by this switching, and no other short cycles are created in  $G'$ .

Since  $\alpha$  shares no edges with other short cycles by our assumptions, deleting  $\alpha$  will not destroy other short cycles. From Condition (1), deleting  $e_i, e'_i$  will not destroy any short cycles either.

Next, we show no other short cycles are created in  $G'$ . Suppose there exists a new short cycle  $\beta$  in  $G'$  created by the switching. Then  $\beta$  contains paths in  $G \cap G'$  separated by edges created in the forward switching in  $G'$  ( $\beta$  must contain at least such edge because it is created). Any such path in  $G \cap G'$  must have a length at least  $r$ , because

- if it starts and ends at vertices in  $\alpha$  and has length less than  $r$ , then combining this path with a path in  $\alpha$  gives a short cycle in  $G$  that intersects  $\alpha$ , which is a contradiction to our assumption on  $\alpha$ ;
- if it starts in  $\alpha$  and ends in  $\{u_i, v_i, u'_i, v'_i\}$  for some  $i$  and has length less than  $r$ , then combining this path with a path in  $\alpha$  gives a path between  $\tilde{e}_i, e_i$  or between  $\tilde{e}_i, e'_i$  of length less than  $2r$ , which violates Condition (2);
- if it starts and ends in different edges among  $\{e_i, e'_i, 1 \leq i \leq k\}$ , then it must have length at least  $r$  by Condition (3);
- if it starts at some vertex in  $e_i$  and ends at some vertex in  $e_i$ , then the path must start and end at different vertices in  $e_i$ . Otherwise,  $\beta$  is not a cycle in the sense of Definition 2.1. Then the path combined with  $e_i$  is a cycle. By Condition (1), it has a length at least  $r$ , a contradiction. In the same way, it cannot start at some vertex in  $e'_i$  and end at some vertex in  $e'_i$ .

This implies  $\beta$  contains exactly one path in  $G \cap G'$ . If not, the two separated paths together with new edges in  $G'$  have lengths greater than  $2r$ , a contradiction to the condition that  $\beta$  is a short cycle. Given the path in  $G \cap G'$ , the remainder of  $\beta$  has two cases:

- a single edge that can be  $x_i v_i, x_i v'_i, y_i u_i$ , or  $y_i u'_i$  for some  $1 \leq i \leq k$ , then by Condition (2), the path in  $G \cap G'$  connecting the two vertices in the edge has length at least  $2r$ , which is a contradiction to the fact that  $\beta$  is a short cycle;
- a single path  $v_i x_i v'_i$  or  $u_i y_i u'_i$ , which is impossible by Condition (4).

From the analysis above, no such  $\beta$  can exist, hence any  $\alpha$ -switching satisfying Conditions (1)-(4) is valid.

Next, we find the number of all switchings satisfying Conditions (1) to (4) to have a lower bound on  $F_\alpha$ . We will do this by bounding from above the number of switchings out of the  $[n]_k [m]_k d_1^k d_2^k$  many choices counted in (2.5) that fail one of the Conditions (1)-(4). We treat the 4 conditions in the following (a)-(d) parts.

(a) There are a total of at most  $\sum_{i=2}^r 2jC_j$  edges in all short cycles of  $G$ . For some  $1 \leq i \leq k$ , if we choose one edge  $e_i$  from a short cycle and the other  $(2k-1)$  edges arbitrarily, we obtain a

forward  $\alpha$ -switching that fails Condition (1). The number of all possible choices is at most

$$k \sum_{j=2}^r 2jC_j \cdot [n-1]_{k-1}[m]_k d_1^{k-1} d_2^k.$$

And if we choose  $e'_i$  from a short cycle and the other  $(2k-1)$  edges arbitrarily, the number of all possible choices is at most

$$k \sum_{j=2}^r 2jC_j \cdot [n]_k [m-1]_{k-1} d_1^k d_2^{k-1}.$$

Altogether the number of choices is at most

$$(2.8) \quad \frac{4}{nd_1} [n]_k [m]_k (d_1 d_2)^k k \sum_{j=2}^r j C_j.$$

(b) To fail Condition (2), we can obtain  $\alpha$ -forward switchings by choosing  $(2k-1)$  edges arbitrarily, and then choose one edge  $e_i$  or  $e'_i$  that is at most of distance  $2r-1$  from  $\tilde{e}_i$  for some  $1 \leq i \leq k$ . From the degree constraints, the number of edges of distance less than  $2r$  from some edge is at most  $O((d_1-1)^r (d_2-1)^r)$ . Similar to Part (a), by considering whether  $e_i$  or  $e'_i$  is chosen for  $1 \leq i \leq k$ , the number of such switchings is at most

$$(2.9) \quad \begin{aligned} & \left( [n-1]_{k-1} [m]_k d_1^{k-1} d_2^k + [n]_k [m-1]_k d_1^k d_2^{k-1} \right) \cdot k \cdot O((d_1-1)^r (d_2-1)^r) \\ & = \frac{1}{nd_1} [n]_k [m]_k (d_1 d_2)^k k O((d_1-1)^r (d_2-1)^r). \end{aligned}$$

(c) For Condition (3), there are three cases to consider depending on whether the pair is  $(e_i, e_j)$ ,  $(e'_i, e'_j)$  or  $(e_i, e'_j)$ .

Suppose the pair  $e_i, e_j$  violates Condition (3). We pick the pair of edges that are within distance  $r-1$  and pick the remaining  $(2k-2)$  edges arbitrarily. There are  $(nd_1)$  many ways to choose  $e_i$ . When  $e_i$  is fixed, there are at most  $O((d_1-1)^{(r+1)/2} (d_2-1)^{(r+1)/2})$  choices for  $e_j$ . Hence the number of switchings that fail Condition (3) is at most

$$\begin{aligned} & (nd_1) \cdot [O((d_1-1)^{(r+1)/2} (d_2-1)^{(r+1)/2})] \cdot k(k-1) \cdot ([n-2]_{k-2} [m]_k d_1^{k-2} d_2^k) \\ & = \frac{1}{nd_1} [n]_k [m]_k d_1^k d_2^k \cdot k^2 \cdot O((d_1-1)^{(r+1)/2} (d_2-1)^{(r+1)/2}). \end{aligned}$$

By the same argument, if the pair is  $(e'_i, e'_j)$ , the number of switchings that fail Condition (3) is at most

$$\frac{1}{nd_1} [n]_k [m]_k d_1^k d_2^k k^2 \cdot O((d_1-1)^{(r+1)/2} (d_2-1)^{(r+1)/2}).$$

When the two edges of the pair violating Condition (3) are  $e_i, e'_j$  for some  $i, j$ , the number is at most

$$\begin{aligned} & (nd_1) \cdot [O((d_1-1)^{(r+1)/2} (d_2-1)^{(r+1)/2})] \cdot (2k^2) \cdot ([n-1]_{k-1} [m-1]_{k-1} d_1^{k-1} d_2^{k-1}) \\ & = \frac{1}{nd_1} [n]_k [m]_k d_1^k d_2^k k^2 O((d_1-1)^{(r+1)/2} (d_2-1)^{(r+1)/2}). \end{aligned}$$

Combining the three cases in Part (c), the number of switchings that violate Condition (3) is at most

$$(2.10) \quad \frac{1}{nd_1} [n]_k [m]_k (d_1 d_2)^k k^2 \cdot O((d_1-1)^{(r+1)/2} (d_2-1)^{(r+1)/2}).$$

(d) Since the distance between a pair of vertices in  $V_1$  or  $V_2$  must be even, to violate Condition (4), we can choose a pair  $u_i, u'_i \in V_1$  or  $v_i, v'_i \in V_2$  that are within distance  $2r-2$  first, then choose

other edges arbitrarily. Similar to the cases above, the number of switchings that fail Condition (4) is at most

$$(2.11) \quad \frac{1}{nd_1} [n]_k [m]_k d_1^k d_2^k k O((d_1 - 1)^r (d_2 - 1)^r).$$

Combining the 4 Cases (a)-(d) above, from (2.8), (2.9), (2.10), and (2.11), we have at most

$$\frac{4}{nd_1} [n]_k [m]_k (d_1 d_2)^k \left( k \sum_{j=2}^r j C_j + O(k(d_1 - 1)^r (d_2 - 1)^r) \right)$$

many switchings that fail one of the Conditions (1)-(4) among the  $[n]_k [m]_k (d_1 d_2)^k$  possible switchings. Then for an absolute constant  $c_1 > 0$ ,

$$F_\alpha \geq [n]_k [m]_k d_1^k d_2^k \left( 1 - \frac{4k \sum_{j=2}^r j C_j + c_1 k (d_1 - 1)^r (d_2 - 1)^r}{nd_1} \right).$$

Therefore (2.6) holds.  $\square$

For the number of backward switchings, we obtain a similar upper bound, but the lower bound is only in expectation.

**Lemma 2.8.** *Let  $G$  be a random  $(d_1, d_2)$ -biregular bipartite graph and let  $\alpha$  be a cycle of length  $2k \leq 2r$  in the complete bipartite graph  $K_{n,m}$ . Let  $B_\alpha$  be the number of valid backward switchings from  $G$  that create  $\alpha$ . Then*

$$(2.12) \quad B_\alpha \leq (d_1(d_1 - 1))^k (d_2(d_2 - 1))^k,$$

and there is an absolute constant  $c_2 > 0$  such that

$$(2.13) \quad \mathbb{E} B_\alpha \geq (d_1(d_1 - 1))^k (d_2(d_2 - 1))^k \left( 1 - \frac{c_2 k (d_1 - 1)^r (d_2 - 1)^r}{nd_1} \right).$$

*Proof.* Given  $\alpha$ , from the degree constraints, the number of choices for  $u_i, u'_i, v_i, v'_i, 1 \leq i \leq k$  that yield a valid backward  $\alpha$  switching is at most  $(d_1(d_1 - 1))^k (d_2(d_2 - 1))^k$ , which gives (2.12).

For the lower bound, we consider the quantity  $B := \sum_\beta B_\beta$ , where  $\beta$  is summing over all possible cycles of length  $2k$  in the complete bipartite graph  $K_{n,m}$ . As in the proof of Lemma 2.7, we give conditions that guarantee a valid backward switching.

Assume  $\beta = (x_1, y_1, \dots, x_k, y_k)$ . We first consider backward switchings that create  $\beta$ . Suppose the paths  $v_i x_i v'_i, u_i y_i u'_i, 1 \leq i \leq k$  in  $G$  satisfy the following conditions:

- (1) The edges  $x_i v_i, x_i v'_i, y_i u_i$ , and  $y_i u'_i$  are not contained in any short cycles.
- (2) For  $1 \leq i \leq k$ , the distance between any vertex in the path  $v_i x_i v'_i$  and any vertex in the path  $u_i y_i u'_i$  is at least  $2r$ .
- (3) For all  $1 \leq i \leq k$  and  $1 \leq j \leq k/2$ , the distance between the paths  $v_i x_i v'_i$  and  $v_{i+j} x_{i+j} v'_{i+j}$  (the index  $i+j$  is calculated modulo  $k$ ) and the distance between  $u_i y_i u'_i$  and  $u_{i+j} y_{i+j} u'_{i+j}$  are at least  $2r - 2j + 1$ .
- (4) For  $1 \leq i \leq k, 1 \leq j \leq k/2$ , the distance between  $v_i x_i v'_i$  and  $u_{i+j} y_{i+j} u'_{i+j}$ , and the distance between  $u_i y_i u'_i$  and  $v_{i+j} x_{i+j} v'_{i+j}$  are at least  $2r - 2j + 2$ .

We will show the four conditions above guarantee a valid backward  $\beta$ -switching.

By Condition (1), no short cycles are deleted. We denote  $x \not\sim y$  if two vertices  $x, y$  are not connected in  $G$ . An immediate consequence of Condition (2) ensures that  $x_i \not\sim y_i$  and  $u_i \not\sim v_i, u'_i \not\sim v'_i$ , and Condition (4) ensures that  $y_i \not\sim x_{i+1}$ . Therefore such switching can be applied.

Let  $G'$  be the graph obtained by applying the backward  $\beta$ -switching. We need to check that no short cycles other than  $\beta$  are created in  $G'$ .

Suppose a short cycle  $\beta' \neq \beta$  is created. Then  $\beta'$  possibly consists of paths in  $G \cap G'$ , portions of  $\beta$ , and edges  $u_i v_i, u'_i v'_i$  for some  $1 \leq i \leq k$ . Any such path in  $G \cap G'$  must have length at least  $r$  because

- if it starts in one of the sets  $\{x_i, v_i, v'_i\}$  or  $\{y_i, u_i, u'_i\}$  for  $1 \leq i \leq k$ , and ends at a different set  $\{x_j, v_j, v'_j\}$  or  $\{y_j, u_j, u'_j\}$  for  $1 \leq j \leq k$ , then Conditions (2), (3) and (4) imply this;
- if it starts and ends in the same set  $\{x_i, v_i, v'_i\}$  or  $\{y_i, u_i, u'_i\}$ , then it follows from Condition (1) that the path must have length at least  $r$ .

It follows that  $\beta'$  must contain exactly one such path, otherwise, if two such paths are included in  $\beta'$ , the length of  $\beta'$  is greater than  $2r$ , a contradiction to the fact that  $\beta'$  is a short cycle.

Besides this path in  $G \cap G'$ , the remainder of  $\beta'$  must either be an edge  $u_i v_i$  or  $u'_i v'_i$ , or a portion of  $\beta$ . If the remainder is some  $u_i v_i$ , then the distance between  $u_i$  and  $v_i$  in  $G$  is at most  $2r - 1$ , a contradiction to Condition (2). The same holds if the remainder is some  $u'_i v'_i$ .

If the remainder is a portion of  $\beta$ , then there exist two vertices in  $\beta$  connected by the path in  $G \cap G'$  contained in  $\beta'$ . If the two vertices are  $x_i, x_{i+j}$  for some  $1 \leq i \leq k, 1 \leq j \leq k/2$ , then from Condition (3), the path in  $G \cap G'$  contained in  $\beta'$  that connects the two vertices has length at least  $2r - 2j + 1$ . Since the path in  $\beta$  connecting  $x_i, x_{i+j}$  has length  $2j$ , this implies  $\beta'$  has length at least  $(2r - 2j + 1) + 2j = 2r + 1$ , a contradiction. In the same way, if the two vertices are  $y_i, y_{i+j}$  for some  $1 \leq i \leq k, 1 \leq j \leq k/2$ , we can find a contradiction for  $\beta'$  from Condition (3).

If the two vertices connected by the path are  $x_i, y_{i+j}$  with  $1 \leq i \leq k, 1 \leq j \leq k/2$ , then the path in  $\beta$  connecting the two vertices has length at least  $2j - 1$ . Combining the path in  $G \cap G'$  contained in  $\beta'$ , from Condition (4), we conclude that  $\beta'$  has length at least  $(2r - 2j + 2) + (2j - 1) = 2r + 1$ , a contradiction. By the same argument, if the two vertices connected by the path are  $y_i, x_{i+j}$  for some  $1 \leq i \leq k, 1 \leq j \leq k/2$ , we can find a contradiction that  $\beta'$  is not a short cycle.

Therefore such  $\beta'$  does not exist, and all backward switchings satisfying Conditions (1)-(4) are valid.

There are  $[n]_k [m]_k / (2k)$  choices for the  $2k$ -cycle  $\beta$  in the complete bipartite graph  $K_{n,m}$ , and at most  $(d_1(d_1 - 1))^k (d_2(d_2 - 1))^k$  choices for  $u_i, u'_i, v_i, v'_i, 1 \leq i \leq k$  given  $\beta$ . We now count how many possible backward switchings violate one of the four Conditions (1)-(4) to get a lower bound on  $B$ . We treat the Conditions (1)-(4) in 4 parts.

(a) Suppose Condition (1) is violated. We estimate the number of switchings by choosing one edge from the set of edges in short cycles and the other edges arbitrarily. Note that by our definition of switchings, we identify  $2k$  different switchings by applying the cyclic rotation or inversion on  $[k]$ . Suppose we choose an edge  $x_i v_i$  or  $x_i v'_i$  from short cycles, similar to the analysis in Lemma (2.7), the number of switchings is at most

$$\begin{aligned} & \left( 2k \sum_{j=2}^r 2j C_j \cdot (d_1 - 1) \right) \cdot \left( \frac{1}{2k} [n-1]_{k-1} (d_1(d_1 - 1))^{k-1} \cdot [m]_k (d_2(d_2 - 1))^k \right) \\ &= \frac{2}{nd_1} [n]_k [m]_k [d_1(d_1 - 1) d_2(d_2 - 1)]^{2k} \sum_{j=2}^r j C_j. \end{aligned}$$

Similarly, if we choose an edge  $y_i u_i$  or  $y_i u'_i$  from short cycles, the number of switchings is at most

$$\begin{aligned} & \left( 2k \sum_{j=2}^r 2j C_j \cdot (d_2 - 1) \right) \cdot \left( \frac{1}{2k} [m-1]_{k-1} (d_2(d_2 - 1))^{k-1} \cdot [n]_k (d_1(d_1 - 1))^k \right) \\ &= \frac{2}{nd_1} [n]_k [m]_k [d_1(d_1 - 1) d_2(d_2 - 1)]^{2k} \sum_{j=2}^r j C_j. \end{aligned}$$

Combining two parts, the number of switchings that violate Condition (1) is at most

$$(2.14) \quad \frac{8k}{nd_1} [n]_k [m]_k [d_1(d_1 - 1)d_2(d_2 - 1)]^k \sum_{j=2}^r jC_j.$$

(b) Suppose for some  $1 \leq i \leq k$ , two paths  $v_i x_i v'_i$  and  $u_i y_i u'_i$  are within distance  $2r - 1$ . The number of switching is at most

$$(2.15) \quad \begin{aligned} & \frac{[n-1]_{k-1} [m-1]_{k-1}}{2k} [d_1(d_1 - 1)d_2(d_2 - 1)]^{k-1} \cdot (knd_1(d_1 - 1)) \cdot O((d_1 - 1)^r (d_2 - 1)^{r+1}) \\ & = \frac{1}{nd_1} [n]_k [m]_k (d_1(d_1 - 1))^k (d_2(d_2 - 1))^k O((d_1 - 1)^r (d_2 - 1)^r). \end{aligned}$$

(c) Suppose for some  $1 \leq i \leq k, 1 \leq j \leq k/2$ , two paths  $\{v_i x_i v'_i, v_{i+j} x_{i+j} v'_{i+j}\}$  are within distance  $2r - 2j$ . The number of switchings is at most

$$\begin{aligned} & \frac{[n-2]_{k-2} [m]_k}{2k} (d_1(d_1 - 1))^{k-2} (d_2(d_2 - 1))^k \cdot nd_1(d_1 - 1) \sum_{i=1}^k \sum_{j=1}^{\lfloor k/2 \rfloor} O((d_1 - 1)^{r-j+2} (d_2 - 1)^{r-j}) \\ & = \frac{1}{nd_1} [n]_k [m]_k (d_1(d_1 - 1))^k (d_2(d_2 - 1))^k O((d_1 - 1)^r (d_2 - 1)^{r-1}). \end{aligned}$$

Suppose for some  $1 \leq i \leq k, 1 \leq j \leq k/2$ , two paths  $\{u_i y_i u'_i, u_{i+j} y_{i+j} u'_{i+j}\}$  are within distance  $2r - 2j$ . Similarly, the number of switchings is bounded by

$$\frac{1}{nd_1} [n]_k [m]_k (d_1(d_1 - 1))^k (d_2(d_2 - 1))^k O((d_1 - 1)^r (d_2 - 1)^{r-1}).$$

Therefore the number of switchings that violate Condition (3) is at most

$$(2.16) \quad \frac{1}{nd_1} [n]_k [m]_k (d_1(d_1 - 1))^k (d_2(d_2 - 1))^k O((d_1 - 1)^r (d_2 - 1)^{r-1}).$$

(d) Suppose two paths  $v_i x_i v'_i, u_{i+j} y_{i+j} u'_{i+j}$  for some  $1 \leq i \leq k, 1 \leq j \leq k/2$  are within distance  $2r - 2j + 1$ . The number of choices is at most

$$\begin{aligned} & \frac{[n]_k [m-1]_{k-1}}{2k} (d_1(d_1 - 1))^k (d_2(d_2 - 1))^{k-1} \sum_{i=1}^k \sum_{j=1}^{\lfloor k/2 \rfloor} O((d_1 - 1)^{r-j+1} (d_2 - 1)^{r-j+2}) \\ & = \frac{1}{nd_1} [n]_k [m]_k (d_1(d_1 - 1))^k (d_2(d_2 - 1))^k O((d_1 - 1)^r (d_2 - 1)^r). \end{aligned}$$

Suppose two paths  $u_i y_i u'_i, v_{i+j} x_{i+j} v'_{i+j}$  for some  $1 \leq i \leq k, 1 \leq j \leq k/2$  are within distance  $2r - 2j + 1$ . By the same argument, the number of choices is at most

$$\frac{1}{nd_1} [n]_k [m]_k (d_1(d_1 - 1))^k (d_2(d_2 - 1))^k O((d_1 - 1)^r (d_2 - 1)^r).$$

Then the number of switchings that violate Condition (4) is at most

$$(2.17) \quad \frac{1}{nd_1} [n]_k [m]_k (d_1(d_1 - 1))^k (d_2(d_2 - 1))^k O((d_1 - 1)^r (d_2 - 1)^r).$$

From (2.14), (2.15), (2.16) and (2.17), the lower bound of  $B$  is given by

$$(2.18) \quad B \geq \frac{[n]_k [m]_k}{2k} (d_1(d_1 - 1))^k (d_2(d_2 - 1))^k \left( 1 - \frac{8k \sum_{j=2}^r jC_j + O(k(d_1 - 1)^r (d_2 - 1)^r)}{nd_1} \right).$$

By Lemma 2.3 (b),

$$\mathbb{E}C_k \leq \frac{[n]_k[m]_k}{2k} \frac{c_1(d_1-1)^k(d_2-1)^k}{n^k m^k} \leq \frac{c_1(d_1-1)^k(d_2-1)^k}{2k}.$$

Applying the inequality above to (2.18), we obtain

$$\mathbb{E}B \geq \frac{[n]_k[m]_k}{2k} (d_1(d_1-1))^k (d_2(d_2-1))^k \left(1 - \frac{O(k(d_1-1)^r(d_2-1)^r)}{nd_1}\right).$$

By the exchangeability of the vertex labels in the uniformly distributed RBBG model, the law of  $B_\beta$  is the same for any  $2k$ -cycle  $\beta$ . Then

$$\mathbb{E}B_\alpha = \frac{2k}{[n]_k[m]_k} \mathbb{E}B \geq (d_1(d_1-1)d_2(d_2-1))^k \left(1 - \frac{c_2 k (d_1-1)^r (d_2-1)^r}{nd_1}\right),$$

for an absolute constant  $c_2 > 0$ . This completes the proof.  $\square$

**2.2. Poisson approximation of cycle counts.** In this section, we prove the cycle counts in RBBGs are asymptotically distributed as Poisson random variables. The main tool we will use is the following total variation distance bound from [23].

**Lemma 2.9** (Proposition 10 in [23]). *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  $\mathbb{E}Z_k = \mu_k$ . 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  $c_k$ , we have*

$$(2.19) \quad d_{\text{TV}}(W, Z) \leq \sum_{k=1}^r \xi_k (\mathbb{E}|\mu_k - c_k \mathbb{P}(\Delta_k^+ | \mathcal{F})| + \mathbb{E}|W_k - c_k \mathbb{P}(\Delta_k^- | \mathcal{F})|),$$

where  $\xi_k := \min\{1, 1.4\mu_k^{-1/2}\}$  and

$$(2.20) \quad \Delta_k^+ := \{W'_k = W_k + 1, W_j = W'_j, k < j \leq r\},$$

$$(2.21) \quad \Delta_k^- := \{W'_k = W_k - 1, W_j = W'_j, k < j \leq r\}.$$

We apply Stein's method to obtain the following Poisson approximation in total variation distance.

**Theorem 2.10.** *Let  $G$  be a random  $(d_1, d_2)$ -biregular bipartite graph with cycle counts  $(C_k, k \geq 2)$ . Let  $(Z_k, k \geq 2)$  be independent Poisson random variables with*

$$\mu_k := \mathbb{E}Z_k = \frac{(d_1-1)^k(d_2-1)^k}{2k}.$$

For any  $n, m \geq 1$  and  $r \geq 2, d_1 \geq 3$ , there exists an absolute constant  $c_6 > 0$  such that

$$d_{\text{TV}}((C_2, \dots, C_r), (Z_2, \dots, Z_r)) \leq \frac{c_6 \sqrt{r} (d_1-1)^{3r/2} (d_2-1)^{3r/2}}{nd_1}.$$

*Proof.* If  $d_1 > n^{1/3}$  or  $r > n^{1/10}$ , then

$$\frac{c_6 \sqrt{r} (d_1-1)^{3r/2} (d_2-1)^{3r/2}}{nd_1} > 1$$

for a sufficiently large choice of  $c_6$  and the theorem holds trivially. Thus we assume  $d_1 \leq n^{1/3}$  and  $r \leq n^{1/10}$ . We now construct an exchangeable pair of random biregular bipartite graphs by taking a step in a reversible Markov chain.

Define a graph  $\mathcal{G}$  whose vertex set consists of all  $(d_1, d_2)$ -biregular bipartite graphs. If there is a valid forward or backward  $\alpha$ -switching from a  $(d_1, d_2)$ -biregular bipartite graph  $G_0$  to another

graph  $G_1$  with the length of  $\alpha$  being  $2k$ , we make an undirected edge in  $\mathcal{G}$  between  $G_0, G_1$  and place a weight of

$$\frac{1}{[n]_k [m]_k (d_1 d_2)^k}$$

on each such edge. Define the degree of a vertex in  $\mathcal{G}$  to be the sum of weights from all adjacent edges. Let  $d_0$  be the largest degree in  $\mathcal{G}$ . To make  $\mathcal{G}$  regular, we add a weighted loop to each vertex if necessary to increase the degree of all vertices to  $d_0$ .

Now consider the simple random walk on  $\mathcal{G}$ . This is a reversible Markov chain with respect to the uniform distribution on  $(d_1, d_2)$ -biregular bipartite graphs. Thus suppose  $G$  is a uniformly chosen random biregular bipartite graph, we can obtain another random biregular bipartite graph  $G'$  by taking an extra step in the random walk from  $G$ , and the pair  $(G, G')$  is exchangeable.

Let  $\mathcal{J}_k$  be the collection of cycles of length  $2k$  in  $K_{n,m}$  with  $k \leq r$ . We have  $|\mathcal{J}_k| = [n]_k [m]_k / 2k$ . Define  $I_\alpha = \mathbf{1}\{\alpha \subseteq G\}$ . Then  $C_k = \sum_{\alpha \in \mathcal{J}_k} \mathbf{1}_\alpha$ . Let  $I'_\alpha, C'_k$  be defined on  $G'$  in the same way. Since  $G$  and  $G'$  are exchangeable, the vectors  $(C_2, \dots, C_r)$  and  $(C'_2, \dots, C'_r)$  are also exchangeable. We can then apply Lemma 2.9 to this exchangeable pair of vectors. Now define two events

$$\begin{aligned} \Delta_k^+ &:= \{C'_k = C_k + 1, C_j = C'_j, k < j \leq r\}, \\ \Delta_k^- &:= \{C_k = C'_k + 1, C_j = C'_j, k < j \leq r\}. \end{aligned}$$

Through our construction of the exchangeable pair,

$$\begin{aligned} \mathbb{P}(\Delta_k^+ \mid G) &= \sum_{\alpha \in \mathcal{J}_k} \frac{B_\alpha}{d_0 [n]_k [m]_k (d_1 d_2)^k}, \\ \mathbb{P}(\Delta_k^- \mid G) &= \sum_{\alpha \in \mathcal{J}_k} \frac{F_\alpha}{d_0 [n]_k [m]_k (d_1 d_2)^k}. \end{aligned}$$

Applying Lemma 2.9 with all  $c_k = d_0, 1 \leq k \leq r$ , we have

$$\begin{aligned} (2.22) \quad & d_{\text{TV}}((C_3, \dots, C_r), (Z_3, \dots, Z_r)) \\ & \leq \sum_{k=2}^r \xi_k \mathbb{E} \left| \mu_k - \sum_{\alpha \in \mathcal{J}_k} \frac{B_\alpha}{[n]_k [m]_k (d_1 d_2)^k} \right| + \sum_{k=2}^r \xi_k \mathbb{E} \left| C_k - \sum_{\alpha \in \mathcal{J}_k} \frac{F_\alpha}{[n]_k [m]_k (d_1 d_2)^k} \right| \\ & = \sum_{k=2}^r \xi_k \mathbb{E} \left| \sum_{\alpha \in \mathcal{J}_k} \frac{(d_1 - 1)^k (d_2 - 1)^k}{[n]_k [m]_k} - \frac{B_\alpha}{[n]_k [m]_k (d_1 d_2)^k} \right| + \sum_{k=2}^r \xi_k \mathbb{E} \left| \sum_{\alpha \in \mathcal{J}_k} I_\alpha - \frac{F_\alpha}{[n]_k [m]_k (d_1 d_2)^k} \right| \\ & \leq \sum_{k=2}^r \xi_k \left( \sum_{\alpha \in \mathcal{J}_k} \mathbb{E} \left| \frac{(d_1 - 1)^k (d_2 - 1)^k}{[n]_k [m]_k} - \frac{B_\alpha}{[n]_k [m]_k (d_1 d_2)^k} \right| + \sum_{\alpha \in \mathcal{J}_k} \mathbb{E} \left| I_\alpha - \frac{F_\alpha}{[n]_k [m]_k (d_1 d_2)^k} \right| \right). \end{aligned}$$

For the rest of the proof, we estimate the following two sums

$$(2.23) \quad \sum_{\alpha \in \mathcal{J}_k} \mathbb{E} \left| \frac{(d_1 - 1)^k (d_2 - 1)^k}{[n]_k [m]_k} - \frac{B_\alpha}{[n]_k [m]_k (d_1 d_2)^k} \right|,$$

$$(2.24) \quad \sum_{\alpha \in \mathcal{J}_k} \mathbb{E} \left| I_\alpha - \frac{F_\alpha}{[n]_k [m]_k (d_1 d_2)^k} \right|$$

from (2.22) in different ways.

(1) *The upper bound on (2.23)*. From Lemma 2.8, for all  $\alpha \in \mathcal{J}_k$ ,

$$\begin{aligned} \mathbb{E} \left| \frac{(d_1 - 1)^k (d_2 - 1)^k}{[n]_k [m]_k} - \frac{B_\alpha}{[n]_k [m]_k (d_1 d_2)^k} \right| &= \frac{(d_1 - 1)^k (d_2 - 1)^k}{[n]_k [m]_k} - \frac{\mathbb{E} B_\alpha}{[n]_k [m]_k (d_1 d_2)^k} \\ &\leq \frac{c_2 k (d_1 - 1)^{r+k} (d_2 - 1)^{r+k}}{nd_1 [n]_k [m]_k}, \end{aligned}$$

where the first line is from (2.12) and the second line is from (2.13). Therefore (2.23) satisfies

$$(2.25) \quad \sum_{\alpha \in \mathcal{J}_k} \mathbb{E} \left| \frac{(d_1 - 1)^k (d_2 - 1)^k}{[n]_k [m]_k} - \frac{B_\alpha}{[n]_k [m]_k (d_1 d_2)^k} \right| \leq \frac{c_2 [(d_1 - 1)(d_2 - 1)]^{r+k}}{2nd_1}.$$

(2) *The upper bound on (2.24)*. To bound the summation in (2.24), for a given short cycle  $\alpha$ , we consider a partition of  $\mathcal{G}$  in the following way:

$$\begin{aligned} A_1^\alpha &= \{G \text{ does not contain } \alpha\}, \\ A_2^\alpha &= \{G \text{ contains } \alpha, \text{ which does not share an edge with another short cycle in } G\}, \\ A_3^\alpha &= \{G \text{ contains } \alpha, \text{ which shares an edge with another short cycle in } G\}. \end{aligned}$$

Conditioned on  $A_1^\alpha$ , we have  $I_\alpha = F_\alpha = 0$ . Conditioned on  $A_2^\alpha$ , both the upper and lower bounds in Lemma 2.7 can apply, which yield the following inequality:

$$(2.26) \quad \left| I_\alpha - \frac{F_\alpha}{[n]_k [m]_k (d_1 d_2)^k} \right| \leq \frac{4k \sum_{j=2}^r j C_j + c_1 k (d_1 - 1)^r (d_2 - 1)^r}{nd_1}.$$

Conditioned on  $A_3^\alpha$ , we have  $I_\alpha = 1, F_\alpha = 0$ .

With the partition of  $\mathcal{G}$ , the following inequality holds:

$$\begin{aligned} \mathbb{E} \left| I_\alpha - \frac{F_\alpha}{[n]_k [m]_k (d_1 d_2)^k} \right| &= \mathbb{E} \left[ \mathbf{1}_{A_2^\alpha} \left| I_\alpha - \frac{F_\alpha}{[n]_k [m]_k (d_1 d_2)^k} \right| \right] + \mathbb{P}(A_3^\alpha) \\ (2.27) \quad &\leq \frac{2k}{nd_1} \mathbb{E} \left[ \mathbf{1}_{A_2^\alpha} \sum_{j=2}^r 2j C_j \right] + \frac{c_1 k (d_1 - 1)^r (d_2 - 1)^r}{nd_1} \mathbb{P}(A_2^\alpha) + \mathbb{P}(A_3^\alpha). \end{aligned}$$

Let  $\mathcal{J}_\alpha$  be the set of all short cycles in  $K_{n,m}$  that share no edges with  $\alpha$ . On the event  $A_2^\alpha$ , the graph  $G$  contains no short cycles outside  $\mathcal{J}_\alpha$  except for  $\alpha$ . Define  $|\beta|$  be the length of the cycle  $\beta$ . Then

$$\sum_{j=2}^r 2j C_j = 2k + \sum_{\beta \in \mathcal{J}_\alpha} |\beta| I_\beta.$$

Therefore the right-hand side of (2.27) can be bounded by

$$\begin{aligned} (2.28) \quad &\frac{4k^2}{nd_1} \mathbb{P}(A_2^\alpha) + \frac{2k}{nd_1} \sum_{\beta \in \mathcal{J}_\alpha} |\beta| \mathbb{E} I_\alpha I_\beta + \frac{c_1 k (d_1 - 1)^r (d_2 - 1)^r}{nd_1} \mathbb{P}(A_2^\alpha) + \mathbb{P}(A_3^\alpha) \\ &\leq \frac{4k^2}{nd_1} \mathbb{P}(\alpha \subseteq G) + \frac{c_1 k (d_1 - 1)^r (d_2 - 1)^r}{nd_1} \mathbb{P}(\alpha \subseteq G) + \frac{2k}{nd_1} \sum_{\beta \in \mathcal{J}_\alpha} |\beta| \mathbb{E} I_\alpha I_\beta + \mathbb{P}(A_3^\alpha). \end{aligned}$$

By Lemma 2.3(1),

$$\begin{aligned} \frac{4k^2}{nd_1} \mathbb{P}(\alpha \subseteq G) &= O \left( \frac{k^2 [(d_1 - 1)(d_2 - 1)]^k}{nd_1 (nm)^k} \right), \\ \frac{c_1 k (d_1 - 1)^r (d_2 - 1)^r}{nd_1} \mathbb{P}(\alpha \subseteq G) &= O \left( \frac{k [(d_1 - 1)(d_2 - 1)]^{k+r}}{nd_1 (nm)^k} \right). \end{aligned}$$

Hence the first and the second term in (2.28) combine to yield a corresponding upper bound in (2.24) of

$$(2.29) \quad \sum_{\alpha \in \mathcal{J}_k} \left( \frac{4k^2}{nd_1} \mathbb{P}(\alpha \subseteq G) + \frac{c_1 k (d_1 - 1)^r (d_2 - 1)^r}{nd_1} \mathbb{P}(\alpha \subseteq G) \right) = O \left( \frac{[(d_1 - 1)(d_2 - 1)]^{k+r}}{nd_1} \right).$$

From Lemma 2.3 (3), we have for any  $\beta \in \mathcal{J}_\alpha$  with  $|\beta| = 2j$ ,

$$\mathbb{E} I_\alpha I_\beta = \mathbb{P}(\alpha \cup \beta \in G) \leq \frac{c_1 [(d_1 - 1)(d_2 - 1)]^{j+k}}{(nm)^{j+k}}.$$

For  $2 \leq j \leq r$ , there are at most  $[n]_j [m]_j / (2j)$  cycles in  $\mathcal{J}_\alpha$  of length  $2j$ . The third term in (2.28) then satisfies

$$\begin{aligned} \frac{2k}{nd_1} \sum_{\beta \in \mathcal{J}_\alpha} |\beta| \mathbb{E} I_\alpha I_\beta &\leq \frac{2k}{nd_1} \sum_{j=2}^r \frac{[n]_j [m]_j}{2j} \cdot 2j \cdot \frac{c_1 [(d_1 - 1)(d_2 - 1)]^{j+k}}{(nm)^{j+k}} \\ &= O \left( \frac{k [(d_1 - 1)(d_2 - 1)]^{r+k}}{nd_1 (nm)^k} \right). \end{aligned}$$

Summing over all possible  $\alpha \in \mathcal{J}_k$ , we obtain a corresponding term in (2.24) of

$$(2.30) \quad \sum_{\alpha \in \mathcal{J}_k} \frac{2k}{nd_1} \sum_{\beta \in \mathcal{J}_\alpha} |\beta| \mathbb{E} I_\alpha I_\beta = O \left( \frac{[(d_1 - 1)(d_2 - 1)]^{r+k}}{nd_1} \right).$$

Now given (2.29) and (2.30), to control (2.24), it remains to estimate  $\sum_{\alpha \in \mathcal{J}_k} \mathbb{P}(A_3^\alpha)$ . Let  $\mathcal{K}_\alpha$  be the set of all short cycles in  $K_{n,m}$  that share an edge with  $\alpha$ , not including  $\alpha$  itself. By a union bound,

$$(2.31) \quad \sum_{\alpha \in \mathcal{J}_k} \mathbb{P}(A_3^\alpha) \leq \sum_{\alpha \in \mathcal{J}_k} \sum_{\beta \in \mathcal{K}_\alpha} \mathbb{P}(\alpha \cup \beta \subset G).$$

From (2.3) in Lemma 2.3, the upper bound for  $\mathbb{P}(\alpha \cup \beta \subset G)$  depends on the lengths of  $\alpha, \beta$ , and the number of edges that  $\alpha, \beta$  share. To get an upper bound on (2.31), we will classify and count the number of pairs  $(\alpha, \beta)$  based on the structure of  $\alpha \cup \beta$ .

Recall  $\alpha$  has length  $2k$ . Suppose  $\beta$  has length  $2j$ . Let  $H = (V(\alpha) \cap V(\beta), E(\alpha) \cap E(\beta))$  be the intersection of  $\alpha$  and  $\beta$ . Suppose  $H$  has  $p$  components and  $f$  edges. Since  $H$  is the intersection of two different cycles,  $H$  must be a forest with  $p + f$  vertices. So  $\alpha \cup \beta$  has  $2j + 2k - p - f$  vertices and  $2j + 2k - f$  edges. Let  $a, b$  be the number of vertices in  $\alpha \cup \beta$  that are from  $V_1$  and  $V_2$ , respectively. Then

$$(2.32) \quad a + b = 2j + 2k - p - f.$$

Let  $v_1, v_2$  be the number of vertices in  $V_1$  and  $V_2$  for  $H$ , respectively. Then we have  $a = j + k - v_1, b = j + k - v_2$ , and  $|a - b| = |v_1 - v_2|$ . Note that each component in  $H$  is a path. For each path, the difference between the number of vertices from  $V_1$  and  $V_2$  is at most 1. This implies

$$(2.33) \quad |a - b| = |v_1 - v_2| \leq p.$$

From the proof of Corollary 21 in [29], the number of all possible isomorphism types of  $\alpha \cup \beta$  given  $|\alpha|, |\beta| \leq 2r$  and  $p, f \leq 2r$  is at most

$$\frac{(16r^3)^{p-1}}{((p-1)!)^2}.$$

For each isomorphism type, as a subgraph in  $K_{n,m}$ , the number of ways to label it is at most  $[n]_a [m]_b + [n]_b [m]_a$ , where the two terms come from assigning vertices in  $V_1, V_2$  in two ways (pick an arbitrary starting vertex, decide whether it is from  $V_1$  or  $V_2$ , then choose labels accordingly).

From (2.32), (2.33), and the assumption that  $n \leq m$ , we have that when  $f$  is even,

$$[n]_a[m]_b + [n]_b[m]_a \leq 2n^{j+k-p-f/2}m^{j+k-f/2} = 2n^{-p}(nm)^{j+k-f/2}.$$

And when  $f$  is odd,

$$[n]_a[m]_b + [n]_b[m]_a \leq 2n^{-p+1}(nm)^{j+k-f/2-1/2}.$$

By (2.3) in Lemma 2.3, the probability of any realization of the isomorphism type as a subgraph in  $G$  is bounded by

$$\frac{c_1[(d_1-1)(d_2-1)]^{j+k-f/2}}{(nm)^{j+k-f/2}}.$$

With all the estimates above, the right-hand side of (2.31) is now bounded by

$$\begin{aligned} & \sum_{j=2}^r \sum_{1 \leq p, f \leq 2r} \frac{(16r^3)^{p-1}}{((p-1)!)^2} \cdot ([n]_a[m]_b + [n]_b[m]_a) \cdot \frac{c_1[(d_1-1)(d_2-1)]^{j+k-f/2}}{(nm)^{j+k-f/2}} \\ & \leq \sum_{j=2}^r \sum_{1 \leq p, f \leq 2r} \frac{(16r^3)^{p-1}}{((p-1)!)^2} \cdot (2n^{-p}(nm)^{j+k-f/2}) \cdot \frac{c_1[(d_1-1)(d_2-1)]^{j+k-f/2}}{(nm)^{j+k-f/2}} \mathbf{1}\{f \text{ is even}\} \\ & \quad + \sum_{j=2}^r \sum_{1 \leq p, f \leq 2r} \frac{(16r^3)^{p-1}}{((p-1)!)^2} \cdot (2n^{-p+1}(nm)^{j+k-f/2-1/2}) \cdot \frac{c_1[(d_1-1)(d_2-1)]^{j+k-f/2}}{(nm)^{j+k-f/2}} \mathbf{1}\{f \text{ is odd}\} \\ & = O\left(\frac{[(d_1-1)(d_2-1)]^{k+r-1}}{n}\right) + O\left(\frac{[(d_1-1)(d_2-1)]^{r+k-1/2}}{(nm)^{1/2}}\right) \\ (2.34) \quad & = O\left(\frac{[(d_1-1)(d_2-1)]^{k+r}}{nd_1}\right). \end{aligned}$$

Combining all estimates from (2.29), (2.30) and (2.34), we finally obtain

$$(2.35) \quad \sum_{\alpha \in \mathcal{J}_k} \mathbb{E} \left| I_\alpha - \frac{F_\alpha}{[n]_k[m]_k(d_1 d_2)^k} \right| = O\left(\frac{[(d_1-1)(d_2-1)]^{r+k}}{nd_1}\right).$$

This provides an upper bound for (2.24).

(3) *The upper bound on (2.22).* Now the upper bounds on (2.23) and (2.24) have been provided in (2.25) and (2.35), respectively. We are ready to estimate (2.22). Recall

$$\xi_k = \min\{1, 1.4\mu_k^{-1/2}\} = \frac{2.8\sqrt{k}}{[(d_1-1)(d_2-1)]^{k/2}}.$$

Then from (2.25) and (2.35), there is an absolute constant  $c_7 > 0$  such that (2.22) is bounded by

$$(2.36) \quad \sum_{k=2}^r \frac{c_7\sqrt{k}[(d_1-1)(d_2-1)]^{r+k/2}}{nd_1} = O\left(\frac{\sqrt{r}[(d_1-1)(d_2-1)]^{3r/2}}{nd_1}\right).$$

This completes the proof.  $\square$

**2.3. Cyclically non-backtracking walks and the Chebyshev polynomials.** In this section, we study non-backtracking walks in biregular bipartite graphs and relate them to the Chebyshev polynomials. The relation will be used in Section 4 to study eigenvalue fluctuations for random biregular bipartite graphs.

**Definition 2.11** (non-backtracking walk). We define a *non-backtracking walk* of length  $2k$  in a biregular bipartite graph to be a walk  $(u_1, v_1, \dots, u_k, v_k, u_{k+1})$  such that  $u_i \in V_1, v_i \in V_2, u_{i+1} \neq u_i$ , for all  $1 \leq i \leq k$  and  $v_{i+1} \neq v_i$  for all  $1 \leq i \leq k-1$ . Note that in our definition, all such walks start and end at some vertices from  $V_1$ .

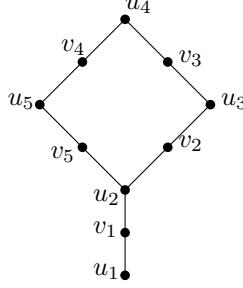


FIGURE 2.  $(u_2, v_2, u_3, v_3, u_4, v_4, u_5, v_5, u_2)$  is a cyclically non-backtracking walk.  $(u_1, v_1, u_2, v_2, u_3, v_3, u_4, v_4, u_5, v_5, u_2, v_1, u_1)$  is a closed non-backtracking but is not cyclically non-backtracking.

**Definition 2.12** (cyclically non-backtracking walk). A walk of length  $2k$  denoted by

$$(u_1, v_1, \dots, u_k, v_k, u_{k+1})$$

is closed if  $u_{k+1} = u_1$ . A *cyclically non-backtracking walk* is a closed non-backtracking walk such that its last two steps are not the reverse of its first two steps. Namely,  $(u_1, v_1, u_2) \neq (u_{k+1}, v_k, u_k)$ . Figure 2 gives an example of a closed non-backtracking walk that is not cyclic non-backtracking.

Let  $G_n$  be a random  $(d_1, d_2)$ -biregular bipartite graph and  $C_k^{(n)}$  be the number of cycles of length  $2k$  in  $G_n$ . Denote  $\text{NBW}_k^{(n)}$  to be the number of non-backtracking walk of length  $2k$ , and  $\text{CNBW}_k^{(n)}$  to be the number of cyclically non-backtracking walks of length  $2k$  in  $G_n$ . Let  $(C_k^{(\infty)}, k \geq 2)$  be independent Poisson random variables with mean

$$\mu_k = \frac{[(d_1 - 1)(d_2 - 1)]^k}{2k}.$$

We also define  $C_1^{(\infty)} = C_1^{(n)} = 0$ . For  $k \geq 1$ , denote

$$(2.37) \quad \text{CNBW}_k^{(\infty)} = \sum_{j|k} 2jC_j^{(\infty)}.$$

For any cycle of length  $2j$  in  $G_n$  with  $j \mid k$ , we can obtain  $2j$  cyclically non-backtracking walks by choosing a starting point from  $V_1$ , fixing a direction and then walking around the cycle of length  $2k$  repeatedly. The next lemma shows that  $\text{CNBW}_k^{(n)}$  can be approximated by the count of those repeated walks around cycles.

**Lemma 2.13.** Let  $G_n$  be a random  $(d_1, d_2)$ -biregular bipartite graph. Suppose  $d_1 \leq n^{1/3}, k \leq n^{1/10}$ , define

$$(2.38) \quad B_k^{(n)} = \text{CNBW}_k^{(n)} - \sum_{j|k} 2jC_j^{(n)}$$

to be the number of cyclically non-backtracking walks of length  $2k$  in  $G_n$  that are not repeated walks around cycles. Then

$$\mathbb{E}B_k^{(n)} \leq \frac{c_7 k^7 [(d_1 - 1)(d_2 - 1)]^k}{n}.$$

We call a cyclically non-backtracking walk *bad* if it's not a repeated walk on a cycle. Then from (2.38),  $B_k^{(n)}$  counts the number of bad cyclically non-backtracking walks of length  $2k$ .

Let  $(w_0, w_1, \dots, w_{2k})$  with  $w_{2k} = w_0 \in V_1$  be a bad cyclically non-backtracking walk in  $K_{n,m}$  of length  $2k$ . For any  $1 \leq i \leq 2k$ , we say that the  $i$ -th step of the walk is

- *free* if  $w_i$  did not previously occur in the walk;
- *a coincidence* if  $w_i$  previously occurred in the walk, but the edge  $w_{i-1}w_i$  didn't;
- *forced* if the edge  $w_{i-1}w_i$  previously occurred in the walk.

Let  $\chi + 1$  be the number of coincidences and  $f$  be the number of forced steps in the walk. Let  $\chi_1 + 1$  and  $\chi_2$  be the number of coincidence steps ending at a vertex from  $V_1$  and  $V_2$ , respectively. Let  $f_1, f_2$  be the number of forced steps ending at a vertex from  $V_1$  and  $V_2$ , respectively. Denote  $v, e$  the number of distinct vertices and edges in the cyclically non-backtracking walk, respectively. We now have the following relations:

$$\begin{aligned}\chi + 1 &= \chi_1 + \chi_2 + 1, \\ f &= f_1 + f_2, \\ v &= (2k + 1) - (\chi + 1) - f = 2k - \chi - f, \\ e &= 2k - f.\end{aligned}$$

For any repeated walk on a cycle, the number of coincidences is 1 and  $\chi = 0$ . Therefore if the walk is bad, we must have  $\chi \geq 1$ .

The following lemma bounds the number of cyclically non-backtracking walks with given parameters  $\chi_1, \chi_2, f_1$ , and  $f_2$ .

**Lemma 2.14.** *Consider cyclically non-backtracking walks of length  $2k$  on  $K_{n,m}$  such that in the subgraph spanned by this walk, all vertices from  $V_1$  have degrees at most  $d_1$  and vertices from  $V_2$  have degrees at most  $d_2$ . Then the number of such walks with given  $\chi_1, \chi_2, f_1, f_2$  satisfying  $\chi \geq 1$  is at most*

$$(2k)^{3(\chi_1 + \chi_2) + 2} (d_1 - 1)^{f_2} (d_2 - 1)^{f_1} n^{k - \chi_1 - f_1} m^{k - \chi_2 - f_2}.$$

Moreover, we must have  $|f_1 - f_2| \leq \chi + 1$ .

*Proof.* We count the number of such cyclically non-backtracking walks by choosing the coincidences, forced steps, and free steps separately. Given that there are  $\chi + 1$  coincidences, there are  $\binom{2k}{\chi+1}$  many possible subsets of indices in  $\{1, \dots, 2k\}$  where coincidences can happen. The vertices at a coincidence have already occurred in the walk, so there are at most  $2k$  choices for each of them, giving us a total of  $\binom{2k}{\chi+1} (2k)^{\chi+1} \leq (2k)^{2\chi+2}$  many choices.

For forced steps, they can only occur after a coincidence or another forced step. After each coincidence, imagine assigning some number of 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, which is  $\binom{f+\chi}{\chi} \leq (2k)^\chi$ . For each forced step ending at a vertex from  $V_1$ , the walk can only move along an edge that has already been traversed, so there are at most  $(d_2 - 1)$  possible choices at every step due to the non-backtracking property. Similarly, for each forced step ending at a vertex from  $V_2$  there are at most  $d_1 - 1$  possible choices. Altogether this gives us at most  $(2k)^\chi (d_1 - 1)^{f_2} (d_2 - 1)^{f_1}$  choices for all forced steps.

There are  $k - \chi - 1 - f_1$  many free steps ending at a vertex from  $V_1$ , we have at most  $n$  choices for the next vertex, and we have an additional  $n$  choices for  $w_0 \in V_1$ , which gives a total of at most  $n^{k - \chi_1 - f_1}$  many choices. Similarly, the number of free steps ending at a vertex from  $V_2$  is at most  $m^{k - \chi_2 - f_2}$ . Multiplying together every part from coincidences, forced steps, and free steps give us at most

$$(2k)^{3\chi+2} (d_1 - 1)^{f_2} (d_2 - 1)^{f_1} n^{k - \chi_1 - f_1} m^{k - \chi_2 - f_2}$$

many such cyclically non-backtracking walks.

Next, we bound  $|f_1 - f_2|$ . Recall that forced steps can only occur after a coincidence or another forced step. Then there are at most  $\chi + 1$  many consecutive forced steps starting from a certain coincidence step. In each consecutive forced step, the number of vertices from  $V_1$  and  $V_2$  differ by at most 1, since the subgraph spanned by any consecutive forced steps is a path. Hence we have  $|f_1 - f_2| \leq \chi + 1$ .  $\square$

Equipped with Lemma 2.14, we continue to prove Lemma 2.13.

*Proof of Lemma 2.13.* By Part (a) in Lemma 2.3, the probability that a given bad walk appears in  $G_n$  is at most

$$c_1 \left( \frac{(d_1 - 1)(d_2 - 1)}{nm} \right)^{k-f/2}.$$

From the upper bound on the number of such walks in Lemma 2.14, summing over all possibilities of  $\chi_1, \chi_2, f_1, f_2$ , we have

$$\begin{aligned} & \mathbb{E} B_k^{(n)} \\ & \leq \sum_{\substack{\chi_1, \chi_2: \\ \chi_1 + \chi_2 \geq 1}} \sum_{\substack{0 \leq f_1, f_2 \leq k-1 \\ |f_1 - f_2| \leq \chi + 1}} (2k)^{3\chi+2} (d_1 - 1)^{f_2} (d_2 - 1)^{f_1} n^{k-\chi_1-f_1} m^{k-\chi_2-f_2} c_1 \left( \frac{(d_1 - 1)(d_2 - 1)}{nm} \right)^{k-f/2} \\ & = c_1 [(d_1 - 1)(d_2 - 1)]^k \sum_{\chi_1 + \chi_2 \geq 1} n^{-\chi_1} m^{-\chi_2} (2k)^{3(\chi_1 + \chi_2) + 2} \sum_{\substack{0 \leq f_1, f_2 \leq k-1 \\ |f_1 - f_2| \leq \chi + 1}} \left( \frac{(d_2 - 1)m}{(d_1 - 1)n} \right)^{(f_1 - f_2)/2}. \end{aligned}$$

Since  $(d_2 - 1)d_1 \leq (d_1 - 1)d_2$ , the following inequality holds:

$$\begin{aligned} (2.39) \quad & \sum_{\substack{0 \leq f_1, f_2 \leq k-1 \\ |f_1 - f_2| \leq \chi + 1}} \left( \frac{(d_2 - 1)m}{(d_1 - 1)n} \right)^{(f_1 - f_2)/2} = \sum_{\substack{0 \leq f_1, f_2 \leq k-1 \\ |f_1 - f_2| \leq \chi + 1}} \left( \frac{(d_2 - 1)d_1}{(d_1 - 1)d_2} \right)^{(f_1 - f_2)/2} \\ & \leq k^2 \left( \frac{(d_1 - 1)d_2}{(d_2 - 1)d_1} \right)^{(\chi+1)/2}. \end{aligned}$$

Since  $d_1 \leq n^{1/3}, k \leq n^{1/10}$ , (2.39) implies

$$\begin{aligned} \mathbb{E} B_k^{(n)} & \leq c_1 k^2 [(d_1 - 1)(d_2 - 1)]^k \sum_{\chi_1 + \chi_2 \geq 1} n^{-\chi_1} m^{-\chi_2} (2k)^{3(\chi_1 + \chi_2) + 2} \left( \frac{(d_1 - 1)d_2}{(d_2 - 1)d_1} \right)^{(\chi+1)/2} \\ & = k^2 [(d_1 - 1)(d_2 - 1)]^k O \left( \frac{(2k)^5 (d_1 - 1)d_2}{n(d_2 - 1)d_1} \right) \\ & = O \left( \frac{k^7 (d_1 - 1)^k (d_2 - 1)^k}{n} \right). \end{aligned}$$

This completes the proof of Lemma 2.13.  $\square$

Recall the definition of  $\text{CNBW}_k^{(\infty)}$  from (2.37). The following corollary holds.

**Corollary 2.15.** *Suppose  $d_1 \leq n^{1/3}$  and  $r \leq n^{1/10}$ . There exists a constant  $c_8 > 0$  such that*

$$(2.40) \quad d_{\text{TV}} \left( (\text{CNBW}_k^{(n)}, 2 \leq k \leq r), (\text{CNBW}_k^{(\infty)}, 2 \leq k \leq r) \right) \leq \frac{c_8 \sqrt{r} [(d_1 - 1)(d_2 - 1)]^{3r/2}}{nd_1}.$$

*Proof.* By the definition of total variation distance, for any measurable map  $f$  and random variable  $X, Y$ , we have

$$(2.41) \quad d_{\text{TV}}(f(X), f(Y)) \leq d_{\text{TV}}(X, Y).$$

It follows from Theorem 2.10 that

$$(2.42) \quad d_{\text{TV}} \left( \left( \sum_{j|k} 2jC_j^{(n)}, 2 \leq k \leq r \right), \left( \text{CNBW}_k^{(\infty)}, 2 \leq k \leq r \right) \right) \leq \frac{c_6 \sqrt{r} [(d_1 - 1)(d_2 - 1)]^{3r/2}}{nd_1}.$$

By Markov's inequality and Lemma 2.13,

$$(2.43) \quad \mathbb{P}(B_k^{(n)} \geq 1) \leq \frac{c_7 k^7 [(d_1 - 1)(d_2 - 1)]^k}{n}.$$

Summing these probabilities for  $k = 2, \dots, r$  implies

$$(2.44) \quad \left( \sum_{j|k} 2jC_j^{(n)}, 2 \leq k \leq r \right) = (\text{CNBW}_k^{(n)}, 2 \leq k \leq r)$$

with probability  $1 - O\left(\frac{r^7[(d_1 - 1)(d_2 - 1)]^r}{n}\right)$ . Therefore by the coupling inequality,

$$(2.45) \quad d_{\text{TV}} \left( \left( \sum_{j|k} 2jC_j^{(n)}, 2 \leq k \leq r \right), (\text{CNBW}_k^{(n)}, 2 \leq k \leq r) \right) = O\left(\frac{r^7[(d_1 - 1)(d_2 - 1)]^r}{n}\right).$$

From (2.42) and (2.45),

$$\begin{aligned} & d_{\text{TV}} \left( (\text{CNBW}_k^{(n)}, 2 \leq k \leq r), (\text{CNBW}_k^{(\infty)}, 2 \leq k \leq r) \right) \\ & \leq \frac{c_6 \sqrt{r} [(d_1 - 1)(d_2 - 1)]^{3r/2}}{nd_1} + O\left(\frac{r^7[(d_1 - 1)(d_2 - 1)]^r}{n}\right) = O\left(\frac{\sqrt{r} [(d_1 - 1)(d_2 - 1)]^{3r/2}}{nd_1}\right). \end{aligned}$$

□

Let  $\lambda_1 \geq \dots \geq \lambda_n$  be the eigenvalues of  $\frac{XX^\top - d_1 I}{\sqrt{(d_1 - 1)(d_2 - 1)}}$ . For the rest of this section, we connect the spectrum of  $\frac{XX^\top - d_1 I}{\sqrt{(d_1 - 1)(d_2 - 1)}}$  with Chebyshev polynomials and cyclically non-backtracking walks. Define

$$(2.46) \quad \begin{aligned} \Gamma_0(x) &= 1, \\ \Gamma_{2k}(x) &= 2T_{2k}\left(\frac{x}{2}\right) + \frac{d_1 - 2}{(d_1 - 1)^k}, \end{aligned}$$

$$(2.47) \quad \Gamma_{2k+1}(x) = 2T_{2k+1}\left(\frac{x}{2}\right).$$

Here  $\{T_k(x)\}$  are the Chebyshev polynomials of the first kind on  $[-1, 1]$  which satisfy

$$(2.48) \quad \begin{aligned} T_0(x) &= 1, & T_1(x) &= x, \\ T_{k+1}(x) &= 2xT_k(x) - T_{k-1}(x). \end{aligned}$$

Let  $\{U_k(x)\}$  be the Chebyshev polynomials of the second kind on  $[-1, 1]$  such that

$$\begin{aligned} U_{-1}(x) &= 0, & U_0(x) &= 1, \\ U_{k+1}(x) &= 2xU_k(x) - U_{k-1}(x). \end{aligned}$$

Define

$$(2.49) \quad p_k(x) = U_k\left(\frac{x}{2}\right) - \frac{1}{d_1 - 1}U_{k-2}\left(\frac{x}{2}\right).$$

We begin with representing closed non-backtracking walks with  $p_k(x)$ . The following lemma gives a deterministic identity. Recall in our Definition 2.11, all closed non-backtracking walks start and end at vertices in  $V_1$ .

**Lemma 2.16.** *Let  $\text{NBW}_k^{(n)}$  be the number of closed non-backtracking walks of length  $2k$  in a  $(d_1, d_2)$ -biregular bipartite graph  $G$ . Let  $\lambda_1 \geq \dots \geq \lambda_n$  be the eigenvalues of  $\frac{XX^\top - d_1 I}{\sqrt{(d_1-1)(d_2-1)}}$ . We have*

$$(2.50) \quad \sum_{i=1}^n p_k(\lambda_i) = (d_1 - 1)^{-k/2}(d_2 - 1)^{-k/2} \text{NBW}_k^{(n)}.$$

*Proof.* Let  $A^{(k)}$  be the  $n \times n$  matrix such that  $A_{ij}^{(k)}$  is the number of non-backtracking walks of length  $2k$  from  $i$  to  $j$ , where  $i, j \in V_1$ . We have the following relations:

$$(2.51) \quad \begin{aligned} A^{(1)} &= XX^\top - d_1 I, & A^{(2)} &= (A^{(1)})^2 - d_1(d_2 - 1)I, \\ A^{(k+1)} &= A^{(1)}A^{(k)} - (d_1 - 1)(d_2 - 1)A^{(k-1)}, & \forall k \geq 2. \end{aligned}$$

The expressions of  $A^{(1)}$  and  $A^{(2)}$  follow from the definition of non-backtracking walks. Since a non-backtracking walk of length  $2k+2$  can be decomposed as a non-backtracking walk of length  $2k$  and a non-backtracking walk of length 2 which avoid backtracking at the  $2k$ -th step, the expression (2.51) holds. We now claim that for  $k \geq 1$ ,

$$(2.52) \quad p_k\left(\frac{XX^\top - d_1 I}{\sqrt{(d_1-1)(d_2-1)}}\right) = [(d_1 - 1)(d_2 - 1)]^{-k/2} A^{(k)},$$

and prove it by induction. Note that from (2.49),

$$p_1(x) = x, \quad p_2(x) = x^2 - 1 - \frac{1}{d_1 - 1}.$$

It is easy to check (2.52) holds for  $k = 1, 2$ . Since  $p_k(x)$  is a linear combination of  $U_k(x/2)$  and  $U_{k-2}(x/2)$ , it satisfies the recursive relation for  $U_k(x/2)$ , which is

$$p_k(x) = xp_k(x) - p_{k-1}(x).$$

Assume (2.52) holds for  $k \leq s$ . Let  $M = XX^\top - d_1 I$ . Then

$$\begin{aligned} p_{s+1}\left(\frac{M}{\sqrt{(d_1-1)(d_2-1)}}\right) &= M[(d_1 - 1)(d_2 - 1)]^{-(s+1)/2} A^{(s)} - [(d_1 - 1)(d_2 - 1)]^{-(s-1)/2} A^{(s-1)} \\ &= [(d_1 - 1)(d_2 - 1)]^{-(s+1)/2} \left( MA^{(s)} - (d_1 - 1)(d_2 - 1)A^{(s-1)} \right) \\ &= [(d_1 - 1)(d_2 - 1)]^{-(s+1)/2} A^{(s+1)}, \end{aligned}$$

where the last equality is from (2.51). Therefore (2.52) holds. Taking trace on both sides in (2.52), we obtain (2.50).  $\square$

The next theorem is an algebraic relation between  $\Gamma_k$  and the number of cyclic non-backtracking walks. Together with Lemma 2.15, it implies the polynomials  $\Gamma_k(x)$  of the eigenvalues for RBBGs converges in distribution to a sum of Poisson random variables.

**Theorem 2.17.** Let  $G$  be a  $(d_1, d_2)$ -biregular bipartite graph and  $\lambda_1 \geq \dots \geq \lambda_n$  be the eigenvalues of  $\frac{XX^\top - d_1 I}{\sqrt{(d_1-1)(d_2-1)}}$ . Then for any  $k \geq 1$ , we have

$$(2.53) \quad \sum_{i=1}^n \Gamma_k(\lambda_i) = (d_1 - 1)^{-k/2} (d_2 - 1)^{-k/2} \text{CNBW}_k^{(n)}.$$

*Proof.* We first relate the number of cyclically non-backtracking closed walks  $\text{CNBW}_k^{(n)}$  to the number of closed non-backtracking walks  $\text{NBW}_k^{(n)}$ .

A closed non-backtracking walk of length  $2k$  is either cyclically non-backtracking, or it can be obtained from a closed non-backtracking walk of length  $2(k-2)$  by adding a new walk of length 2 (which we call a *tail*) to the beginning of the walk and its reverse to the end (see Figure 2 for an example). For any cyclically non-backtracking walk of length  $2(k-2)$ , we can add a tail in  $(d_1-2)(d_2-1)$  many ways. For any closed non-backtracking walk of length  $2(k-2)$  that is not cyclically non-backtracking, we can add a tail in  $(d_1-1)(d_2-1)$  many ways. Therefore for  $k \geq 3$ , we have the following equation

$$\begin{aligned} \text{NBW}_k^{(n)} &= \text{CNBW}_k^{(n)} + (d_1 - 2)(d_2 - 1)\text{CNBW}_{k-2}^{(n)} + (d_1 - 1)(d_2 - 1)(\text{NBW}_{k-2}^{(n)} - \text{CNBW}_{k-2}^{(n)}) \\ &= \text{CNBW}_k^{(n)} + (d_1 - 1)(d_2 - 1)\text{NBW}_{k-2}^{(n)} - (d_2 - 1)\text{CNBW}_{k-2}^{(n)}, \end{aligned}$$

which can be written as

$$(2.54) \quad \text{CNBW}_k^{(n)} - (d_2 - 1)\text{CNBW}_{k-2}^{(n)} = \text{NBW}_k^{(n)} - (d_1 - 1)(d_2 - 1)\text{NBW}_{k-2}^{(n)}.$$

Note that  $\text{CNBW}_k^{(n)} = \text{NBW}_k^{(n)}$  for  $k = 1, 2$ . Applying (2.54) recursively, we have when  $k$  is even,

$$(2.55) \quad \begin{aligned} \text{CNBW}_k^{(n)} &= \text{NBW}_k^{(n)} - (d_1 - 2)[(d_2 - 1)\text{NBW}_{k-2}^{(n)} + (d_2 - 1)^2\text{NBW}_{k-4}^{(n)} + \dots + (d_2 - 1)^{\frac{k-2}{2}}\text{NBW}_2^{(n)}]. \end{aligned}$$

And when  $k$  is odd,

$$(2.56) \quad \begin{aligned} \text{CNBW}_k^{(n)} &= \text{NBW}_k^{(n)} - (d_1 - 2)[(d_2 - 1)\text{NBW}_{k-2}^{(n)} + (d_2 - 1)^2\text{NBW}_{k-4}^{(n)} + \dots + (d_2 - 1)^{\frac{k-3}{2}}\text{NBW}_3^{(n)}]. \end{aligned}$$

Denote

$$\overline{\text{NBW}}_k^{(n)} := (d_2 - 1)^{-k/2} \text{NBW}_k^{(n)}, \quad \overline{\text{CNBW}}_k^{(n)} := (d_2 - 1)^{-k/2} \text{CNBW}_k^{(n)}.$$

We can simplify the above equations (2.55) and (2.56) as

$$(2.57) \quad \overline{\text{CNBW}}_k^{(n)} = \overline{\text{NBW}}_k^{(n)} - (d_1 - 2) \left( \overline{\text{NBW}}_{k-2}^{(n)} + \overline{\text{NBW}}_{k-4}^{(n)} + \dots + \overline{\text{NBW}}_a^{(n)} \right),$$

where  $a = 2$  if  $k$  is even and  $a = 1$  if  $k$  is odd. Also (2.50) can be written as

$$(2.58) \quad \sum_{i=1}^n p_k(\lambda_i) = (d_1 - 1)^{-k/2} \overline{\text{NBW}}_k^{(n)}.$$

From the proof of Proposition 32 in [30], we have the following relation between  $\Gamma_k(x)$  and  $p_k(x)$  for  $k \geq 1$ :

$$(2.59) \quad \Gamma_{2k}(x) = p_{2k}(x) - (d_1 - 2) \left( \frac{p_{2k-2}(x)}{d_1 - 1} + \frac{p_{2k-4}(x)}{(d_1 - 1)^2} + \dots + \frac{p_2(x)}{(d_1 - 1)^{k-1}} \right),$$

$$(2.60) \quad \Gamma_{2k-1}(x) = p_{2k-1}(x) - (d_1 - 2) \left( \frac{p_{2k-3}(x)}{d_1 - 1} + \frac{p_{2k-5}(x)}{(d_1 - 1)^2} + \dots + \frac{p_1(x)}{(d_1 - 1)^{k-1}} \right).$$

Then from (2.58) and (2.57),

$$(d_1 - 1)^{-k} \overline{\text{CNBW}}_{2k}^{(n)} = \sum_{i=1}^n \left( p_{2k}(\lambda_i) - (d_1 - 2) \left( \frac{p_{2k-2}(\lambda_i)}{d_1 - 1} + \cdots + \frac{p_2(\lambda_i)}{(d_1 - 1)^{k-1}} \right) \right) = \sum_{i=1}^n \Gamma_{2k}(\lambda_i),$$

where the last equality is from (2.59). Similarly, from (2.60),

$$(d_1 - 1)^{(2k-1)/2} \overline{\text{CNBW}}_{2k-1}^{(n)} = \sum_{i=1}^n \Gamma_{2k-1}(\lambda_i).$$

Therefore for all  $k \geq 1$ ,

$$\sum_{i=1}^n \Gamma_k(\lambda_i) = (d_1 - 1)^{-k/2} \overline{\text{CNBW}}_k^{(n)} = [(d_1 - 1)(d_2 - 1)]^{-k/2} \text{CNBW}_k^{(n)}.$$

This completes the proof of Theorem 2.17.  $\square$

### 3. SPECTRAL GAP

In this section, we provide some estimates on the second largest eigenvalue of the random biregular bipartite graphs that will be used to study eigenvalue fluctuations in Section 4. Note that the largest eigenvalue of  $XX^\top - d_1 I$  is  $\lambda_1 = d_1(d_2 - 1)$ . In the next theorem, we provide upper bounds on  $|\lambda|$  for all eigenvalues  $\lambda \neq \lambda_1$ .

**Theorem 3.1.** *Let  $G$  be a  $(d_1, d_2)$ -random biregular bipartite graph with  $d_1 \geq d_2$ . Let  $\lambda_1 \geq \cdots \geq \lambda_n$  be the eigenvalues of  $XX^\top - d_1 I$ .*

(1) *For fixed  $d_1, d_2$ , there exists a sequence  $\varepsilon_n \rightarrow 0$  such that for any eigenvalue  $\lambda \neq \lambda_1$ ,*

$$(3.1) \quad \mathbb{P}(|\lambda - (d_2 - 2)| \geq 2\sqrt{(d_1 - 1)(d_2 - 1)} + \varepsilon_n) \rightarrow 0$$

*as  $n \rightarrow \infty$ .*

(2) *Suppose  $d_2 \leq \frac{1}{2}n^{2/3}$ ,  $d_1 \geq d_2 \geq cd_1$  for some constant  $c \in (0, 1)$ . Then for some constant  $\alpha_1 > 0$  depending on  $c$  and any eigenvalue  $\lambda \neq \lambda_1$ ,*

$$(3.2) \quad \mathbb{P}(|\lambda| \geq \alpha_1 \sqrt{(d_1 - 1)(d_2 - 1)}) \leq \frac{1}{n^2}.$$

(3) *Suppose  $d_2 \leq C_1$ ,  $d_1 \leq n^2$ , there exists a constant  $\alpha_2$  depending on  $C_1$  such that for any eigenvalue  $\lambda \neq \lambda_1$ ,*

$$(3.3) \quad \mathbb{P}(|\lambda| \geq \alpha_2 \sqrt{(d_1 - 1)(d_2 - 1)}) \leq \frac{1}{n^2}.$$

*Remark 3.2.* The probability estimates in (3.2) and (3.3) can be improved, see [66]. In order to prove the main theorems in Section 4, we only include a weaker version for simplicity.

*proof of Theorem 3.1.* Theorem 4 in [16] states that for a random biregular bipartite graph with  $d_1 \geq d_2$ , the eigenvalues of the adjacency matrix  $A$  satisfy the following estimates with high probability:

- (1) the second eigenvalue of  $A$  satisfies  $\lambda_2(A) \leq \sqrt{d_1 - 1} + \sqrt{d_2 - 1} + o(1)$ ,
- (2) the smallest positive eigenvalue of  $A$  satisfies  $\lambda_{\min}^+(A) \geq \sqrt{d_1 - 1} - \sqrt{d_2 - 1} - o(1)$ .

Since eigenvalues of  $XX^\top$  are the squares of the eigenvalues for  $A$ , we have with high probability,

$$\begin{aligned} \lambda_2(XX^\top) - d_1 - (d_2 - 2) &\leq 2\sqrt{(d_1 - 1)(d_2 - 1)} + o(1), \\ \lambda_n(XX^\top) - d_1 - (d_2 - 2) &\geq -2\sqrt{(d_1 - 1)(d_2 - 1)} - o(1), \end{aligned}$$

therefore (3.1) holds.

Theorem 1.1 in [66] states that if  $d_2 \leq \frac{1}{2}n^{2/3}$  and  $d_1 \geq d_2$ , there exists a constant  $\alpha > 0$  such that  $\lambda_2(A) \leq \alpha\sqrt{d_1}$  with probability at least  $1 - m^{-2}$ . This implies for any eigenvalue  $\lambda$  of  $XX^\top - d_1I$  with  $\lambda \neq d_1(d_2 - 1)$ , we have

$$\mathbb{P}(-d_1 \leq \lambda \leq \alpha^2 d_1 - d_1) \geq 1 - m^{-2} \geq 1 - n^{-2}.$$

Since  $d_1 \geq d_2 \geq cd_1$ , we can find a constant  $\alpha_1 > 0$  depending on  $\alpha$  and  $c$  such that

$$\mathbb{P}(|\lambda| \leq \alpha_1 \sqrt{(d_1 - 1)(d_2 - 1)}) \geq 1 - n^{-2}.$$

Therefore (3.2) holds. Theorem 1.5 in [66] states that if  $d_2 \leq C_1, d_1 \leq n^2$ , there exists a constant  $\alpha_2$  depending on  $C_1$  such that

$$\mathbb{P}\left(\max_{2 \leq i \leq m+n-1} |\lambda_i^2(A) - d_1| \geq \alpha_2 \sqrt{(d_1 - 1)(d_2 - 1)}\right) \leq n^{-2}.$$

Then (3.3) follows from the algebraic relation between the spectra of  $A$  and  $XX^\top - d_1I$ .  $\square$

#### 4. EIGENVALUE FLUCTUATIONS

Lemma 2.17 and Corollary 2.15 in Section 2.3 imply the limiting laws for  $\sum_{i=1}^n \Gamma_k(\lambda_i)$  are given by a sum of Poisson random variables. In this section, we extend the results to a more general class of function  $f$  and study the behavior of  $\sum_{i=1}^n f(\lambda_i)$  for RBBGs with fixed and growing degrees.

The following set-up for weak convergence will be used in Section 4.2 to prove Theorem 4.7. We will closely follow the definitions and notations used in [30]. See Section 2 in [30] for more details.

Denote  $\mathbb{N} := \{1, 2, \dots\}$ . Let  $\vec{w} = (w_m)_{m \in \mathbb{N}}$  be a sequence of positive weights. Let  $L^2(\vec{w})$  be the space of sequences  $(x_m)_{m \in \mathbb{N}}$  that are square-integrable with respect to  $\vec{w}$ , i.e.,  $\sum_{m=1}^{\infty} x_m^2 w_m < \infty$ . We define a complete separable metric space  $\mathcal{X} = (L^2(\vec{w}), \|\cdot\|)$ , where for any sequence  $(x_m)_{m \in \mathbb{N}}$ ,

$$\|x\| = \left( \sum_{m=1}^{\infty} x_m^2 w_m \right)^{1/2}.$$

Denote the space of probability measures on the Borel  $\sigma$ -algebra of  $\mathcal{X}$  by  $\mathcal{P}(\mathcal{X})$ . We use the Prokhorov metric for weak convergence as the metric on  $\mathcal{P}(\mathcal{X})$ . The following results are proved in Section 2 of [30].

**Proposition 4.1** (Lemma 2-4 in [30]). *The following holds for the complete separable metric space  $\mathcal{X}$ .*

(1) *Let  $(a_m)_{m \in \mathbb{N}} \in L^2(\vec{w})$  be such that  $a_m \geq 0$  for every  $m$ . Then the set*

$$\{(b_m)_{m \in \mathbb{N}} \in L^2(\vec{w}) : 0 \leq |b_m| \leq a_m, \forall m \in \mathbb{N}\}$$

*is compact in  $(L^2(\vec{w}), \|\cdot\|)$ .*

(2) *Suppose  $\{X_n\}$  and  $X$  are random sequences taking values in  $L^2(\vec{w})$  such that  $X_n$  converges in distribution to  $X$ . Then for any  $b \in L^2(\vec{w})$ , the random variables  $\langle b, X_n \rangle$  converges in distribution to  $\langle b, X \rangle$ .*

(3) *Let  $x \in \mathcal{X}$  and  $P, Q$  be two probability measures in  $\mathcal{P}(\mathcal{X})$ . Suppose for any finite collection of indices  $(i_1, \dots, i_k)$ , the law of random vector  $(x_{i_1}, \dots, x_{i_k})$  is the same under both  $P$  and  $Q$ . Then  $P = Q$  on the entire Borel  $\sigma$ -algebra of  $\mathcal{X}$ .*

We also need the following results from the approximation theory.

**Definition 4.2** (Bernstein ellipse). For  $\rho > 1$ , let  $E_\rho$  be the image of the map  $z \mapsto (z + z^{-1})/2$  of the open disc of radius  $\rho$  in the complex plain centered at the origin. We can  $E_\rho$  the *Bernstein ellipse* of radius  $\rho$ . The ellipse has foci at  $\pm 1$ , and the sum of the major semi-axis and minor semi-axis is exactly  $\rho$ .

**Proposition 4.3** ([64], Theorem 8.1). *Suppose  $f : [-1, 1] \rightarrow \mathbb{R}$  can be analytically extended to  $E_\rho$  and is bounded by  $M$  on  $E_\rho$ . Then  $f$  has a unique expansion on  $[-1, 1]$  as*

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

where  $T_k(x)$  is the Chebyshev polynomial of the first kind defined in (2.48), and the coefficients of this expansion satisfy

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

Define  $f_k(x) = \sum_{i=0}^k a_i T_i(x)$ . Applying the bound  $|T_k(x)| \leq 1$  when  $x \in [-1, 1]$  and Proposition 4.3, we obtain for all  $x \in [-1, 1]$ ,

$$(4.1) \quad |f(x) - f_k(x)| \leq \frac{2M}{\rho^k(\rho - 1)}.$$

**4.1. Poisson fluctuations with fixed degrees.** Now fix  $d_1$  and  $d_2$  as constants. We are ready to extend our results in Section 2.3 to a more general class of functions as follows. Note that the following theorem is given for a sequence of RBBGs with growing  $n$ . For ease of notation, we drop the dependence on  $n$  when writing the matrix  $X$  and eigenvalues  $\lambda_1, \dots, \lambda_n$ .

**Theorem 4.4.** *For fixed  $d_1 \geq d_2 \geq 2$  and  $(d_1, d_2) \neq (2, 2)$ , let  $G_n$  be a sequence of random  $(d_1, d_2)$ -biregular bipartite graph. Let  $\lambda_1 \geq \dots \geq \lambda_n$  be the eigenvalues of  $\frac{XX^\top - d_1 I}{\sqrt{(d_1-1)(d_2-1)}}$ . Suppose  $f$  is a function such that  $f(2z)$  is analytic on  $E_\rho$ , where  $\rho = [(d_1-1)(d_2-1)]^\alpha$  for some  $\alpha > \frac{7}{2}$ . Then  $f(x)$  can be expanded on  $[-2, 2]$  as*

$$(4.2) \quad f(x) = \sum_{k=0}^{\infty} a_k \Gamma_k(x),$$

and the random variable

$$(4.3) \quad Y_f^{(n)} := \sum_{i=1}^n f(\lambda_i) - n a_0$$

converges in distribution as  $n \rightarrow \infty$  to the infinitely divisible random variable

$$(4.4) \quad Y_f := \sum_{k=2}^{\infty} \frac{a_k}{[(d_1-1)(d_2-1)]^{k/2}} \text{CNBW}_k^{(\infty)},$$

where  $\text{CNBW}_k^{(\infty)}$  is defined in (2.37).

*Proof.* Define

$$f_k(x) := \sum_{i=0}^k a_i \Gamma_i(x).$$

We first show that  $f_k(x)$  is a good approximation of  $f(x)$ . Applying Proposition 4.3 to  $f(2x)$  gives an expansion (4.2) with

$$(4.5) \quad |a_k| \leq C[(d_1-1)(d_2-1)]^{-\alpha k}$$

for some constant  $C$  that depends only on  $d_1, d_2$  and the constant  $M$  given in Proposition 4.3.

By the properties of Chebyshev polynomials, on any interval  $[-K, K]$ , we have

$$(4.6) \quad \max_{|x| \leq K} |T_k(x)| = \frac{(K - \sqrt{K^2 - 1})^k + (K + \sqrt{K^2 - 1})^k}{2}.$$

From (2.46) and (2.47), we have  $\Gamma_1(x) = x$ , and for any  $k \geq 2$ ,

$$(4.7) \quad |\Gamma_k(x)| \leq 2 \left| T_k \left( \frac{x}{2} \right) \right| + \frac{d_1 - 2}{(d_1 - 1)^{k/2}} \leq 2 \left| T_k \left( \frac{x}{2} \right) \right| + 1.$$

From (4.6),

$$\max_{|x| \leq 3} 2 \left| T_k \left( \frac{x}{2} \right) \right| = \left( \frac{3}{2} - \frac{\sqrt{5}}{2} \right)^k + \left( \frac{3}{2} + \frac{\sqrt{5}}{2} \right)^k,$$

Then for all  $k \geq 2$ , with (4.7) we obtain

$$(4.8) \quad \sup_{|x| \leq 3} |\Gamma_k(x)| \leq \left( \frac{3}{2} + \frac{\sqrt{5}}{2} \right)^k + 2 \leq 3^{k+1},$$

and the same bound holds when  $k = 1$ . From (4.5) and (4.8), for all  $x \in [-3, 3]$ ,

$$\sum_{k=0}^{\infty} |a_k \Gamma_k(x)| \leq 3C \sum_{k=0}^{\infty} [3((d_1 - 1)(d_2 - 1))^{-\alpha}]^k < \infty,$$

where the last inequality comes from the fact that  $(d_1 - 1)(d_2 - 1) \geq 2$  and  $\alpha > \frac{7}{2}$ . Hence the series  $\sum_{k=0}^{\infty} a_k \Gamma_k(x)$  is absolutely convergent on  $[-3, 3]$ , which implies the expansion of  $f$  in (4.2) is valid on  $[-3, 3]$ . Then we have for a constant  $C_1 > 0$  depending on  $C$ ,

$$(4.9) \quad \sup_{|x| \leq 3} |f(x) - f_k(x)| \leq \sup_{|x| \leq 3} \sum_{i=k+1}^{\infty} |a_i \Gamma_i(x)| \leq C_1 [3(d_1 - 1)(d_2 - 1)^{-\alpha}]^{k+1}.$$

Denote

$$K_1 := \lambda_1 = \frac{d_1 \sqrt{(d_2 - 1)}}{\sqrt{d_1 - 1}}.$$

For sufficiently large  $k$ , from (4.6) and (4.7),

$$\sup_{|x| \leq K_1} |\Gamma_k(x)| \leq (2K_1)^k.$$

And from (4.5) and the assumption  $\alpha > 7/2$ ,

$$\sum_{k=0}^{\infty} |a_k \Gamma_k(x)| \leq C \sum_{k=0}^{\infty} [2K_1((d_1 - 1)(d_2 - 1))^{-\alpha}]^k < \infty.$$

It implies the series  $\sum_{k=0}^{\infty} a_k \Gamma_k(x)$  is also absolutely convergent on  $[-K_1, K_1]$ , and the expansion of  $f$  in (4.2) is valid on  $[-K_1, K_1]$ .

Since  $2K_1 \leq 4[(d_1 - 1)(d_2 - 1)]^{1/2}$  and  $(d_1 - 1)(d_2 - 1) \geq 2$ , for a constant  $C_2 > 0$ ,

$$(4.10) \quad \begin{aligned} \sup_{|x| \leq K_1} |f(x) - f_k(x)| &\leq C_2 [2K_1((d_1 - 1)(d_2 - 1))^{-\alpha}]^{k+1} \\ &\leq C_2 \left[ 4((d_1 - 1)(d_2 - 1))^{-\alpha + \frac{1}{2}} \right]^{k+1} \leq C_2 \left[ ((d_1 - 1)(d_2 - 1))^{-\alpha + \frac{5}{2}} \right]^{k+1}. \end{aligned}$$

Therefore  $f_k$  converges to  $f$  uniformly on  $[-K_1, K_1]$ , and the interval  $[-K_1, K_1]$  deterministically contains all eigenvalues of  $\frac{XX^\top - d_1 I}{\sqrt{(d_1 - 1)(d_2 - 1)}}$ .

By the definition of  $\text{CNBW}_k^{(\infty)}$  in (2.37), Equation (4.4) can be written as

$$Y_f := \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} \frac{a_{ij}}{[(d_1 - 1)(d_2 - 1)]^{ij/2}} 2j C_j^{(\infty)},$$

where  $Y_f$  is a sum of independent random variables, and  $\mathbb{E}|Y_f|^2 < \infty$  by (4.5).

Denote  $\alpha' := \alpha - 2 > \frac{3}{2}$ . Choose  $\beta$  satisfying  $\frac{1}{\alpha'} < \beta < \frac{2}{3}$  and define

$$\begin{aligned} r_n &= \left\lfloor \frac{\beta \log n}{\log[(d_1 - 1)(d_2 - 1)]} \right\rfloor, \\ X_f^{(n)} &= \sum_{k=1}^{r_n} \frac{a_k}{[(d_1 - 1)(d_2 - 1)]^{k/2}} \text{CNBW}_k^{(n)}, \\ \tilde{Y}_f^{(n)} &= \sum_{k=1}^{r_n} \frac{a_k}{[(d_1 - 1)(d_2 - 1)]^{k/2}} \text{CNBW}_k^{(\infty)}. \end{aligned}$$

Note that  $\text{CNBW}_1^{(n)} = 0$ , from (2.53),

$$(4.11) \quad X_f^{(n)} = \sum_{k=2}^{r_n} \frac{a_k}{[(d_1 - 1)(d_2 - 1)]^{k/2}} \text{CNBW}_k^{(n)} = \sum_{i=1}^n f_{r_n}(\lambda_i) - n a_0.$$

By Corollary 2.15,

$$d_{\text{TV}} \left( (\text{CNBW}_k^{(n)}, 2 \leq k \leq r_n), (\text{CNBW}_k^{(\infty)}, 2 \leq k \leq r_n) \right) \leq \frac{c_8 \sqrt{r_n} [(d_1 - 1)(d_2 - 1)]^{3r_n/2}}{n d_1} = o(1).$$

Since  $X_f^{(n)}$  and  $\tilde{Y}_f^{(n)}$  are measurable functions of

$$(\text{CNBW}_k^{(n)}, 2 \leq k \leq r_n) \quad \text{and} \quad (\text{CNBW}_k^{(\infty)}, 2 \leq k \leq r_n),$$

respectively, we have

$$d_{\text{TV}} \left( X_f^{(n)}, \tilde{Y}_f^{(n)} \right) \leq d_{\text{TV}} \left( (\text{CNBW}_k^{(n)}, 2 \leq k \leq r_n), (\text{CNBW}_k^{(\infty)}, 2 \leq k \leq r_n) \right) = o(1).$$

Note that  $\tilde{Y}_f^{(n)}$  converges almost surely to  $Y_f$  by (4.5), so  $X_f^{(n)}$  converges in distribution to  $Y_f$ .

By Slutsky's theorem, to show  $Y_f^{(n)}$  defined in (4.3) converges in distribution to  $Y_f$ , it remains to show that  $Y_f^{(n)} - X_f^{(n)}$  converges to zero in probability. The largest eigenvalue of  $\frac{XX^\top - d_1 I}{\sqrt{(d_1 - 1)(d_2 - 1)}}$  is  $K_1$ , so from (4.10) we have

$$\lim_{k \rightarrow \infty} f_k(\lambda_1) = f(\lambda_1).$$

Then for any  $\delta > 0$  and sufficiently large  $n$ ,

$$(4.12) \quad |f(\lambda_1) - f_{r_n}(\lambda_1)| \leq \delta/2.$$

From (4.3), (4.11) and (4.12), we have for sufficiently large  $n$ ,

$$(4.13) \quad \left| Y_f^{(n)} - X_f^{(n)} \right| \leq \sum_{i=1}^n |f(\lambda_i) - f_{r_n}(\lambda_i)| \leq \frac{\delta}{2} + \sum_{i=2}^n |f(\lambda_i) - f_{r_n}(\lambda_i)|.$$

Suppose that all the non-trivial eigenvalues  $\lambda \neq \lambda_1$  are contained in  $[-3, 3]$ , from (4.9),

$$\begin{aligned} \sum_{i=2}^n |f(\lambda_i) - f_{r_n}(\lambda_i)| &\leq C_1(n-1)[3(d_1 - 1)(d_2 - 1)^{-\alpha}]^{r_n+1} \\ &\leq C_1 n [(d_1 - 1)(d_2 - 1)^{-\alpha+2}]^{r_n} \leq C_1 n^{1-\alpha'\beta} = o(1), \end{aligned}$$

which combining (4.13) implies for sufficiently large  $n$ ,

$$\left| Y_f^{(n)} - X_f^{(n)} \right| \leq \delta.$$

Recall (3.1) and the assumption  $d_1 \geq d_2$ . With high probability, for a sequence  $\varepsilon_n \rightarrow 0$ , we have the nontrivial eigenvalues of  $\frac{XX^\top - d_1 I}{\sqrt{(d_1-1)(d_2-1)}}$  is contained in

$$\left[ -2 - \varepsilon_n + \frac{d_2 - 2}{\sqrt{(d_1-1)(d_2-1)}}, 2 + \varepsilon_n + \frac{d_2 - 2}{\sqrt{(d_1-1)(d_2-1)}} \right] \subseteq [-3, 3]$$

for sufficiently large  $n$ . Therefore

$$\mathbb{P} \left( \left| Y_f^{(n)} - X_f^{(n)} \right| \geq \delta \right) \leq \mathbb{P} \left( \max_{2 \leq i \leq n} |\lambda_i| \geq 3 \right) = o(1).$$

This finishes the proof.  $\square$

As a corollary of Theorem 4.4, we obtain eigenvalue fluctuations for the adjacency matrices of RBBGs as follows.

**Corollary 4.5.** *For fixed  $d_1 \geq d_2 \geq 2$  and  $(d_1, d_2) \neq (2, 2)$ , let  $G_n$  be a sequence of random  $(d_1, d_2)$ -biregular bipartite graph. Let  $\lambda_1 \geq \dots \geq \lambda_{n+m}$  be the eigenvalues of its adjacency matrix  $A$ . Suppose  $f$  satisfies the same conditions as in Theorem 4.4. Then the random variable*

$$Y_f^{(n)} := \frac{1}{2} \left[ \sum_{i=1}^{n+m} f \left( \frac{\lambda_i^2 - d_1}{\sqrt{(d_1-1)(d_2-1)}} \right) - (m-n)f \left( \frac{-d_1}{\sqrt{(d_1-1)(d_2-1)}} \right) \right] - na_0$$

converges in distribution as  $n \rightarrow \infty$  to the infinitely divisible random variable

$$Y_f := \sum_{k=2}^{\infty} \frac{a_k}{[(d_1-1)(d_2-1)]^{k/2}} \text{CNBW}_k^{(\infty)},$$

where  $\text{CNBW}_k^{(\infty)}$  is defined in (2.37).

*Proof.* Recall from Section 1.2 that all eigenvalues of  $A$  consist of two parts. There are  $2n$  eigenvalues in pair as  $\{-\lambda, \lambda\}$  where  $\lambda$  is a singular value of  $X$ . In addition, there are  $(m-n)$  extra zero eigenvalues. the result then follows from the algebraic relation between eigenvalues of  $A$  and eigenvalues of  $XX^\top - d_1 I$ .  $\square$

**4.2. Gaussian fluctuations with growing degrees.** In this section, we consider the eigenvalue fluctuations of RBBGs when  $d_1 \cdot d_2 \rightarrow \infty$ .

We first prove the following weak convergence result for a normalized and centered version of  $\text{CNBW}_k^{(\infty)}$ .

**Lemma 4.6.** *Suppose that  $d_1 \cdot d_2 \rightarrow \infty$ ,  $r_n \rightarrow \infty$  as  $n \rightarrow \infty$ . For  $k \geq 2$ , define*

$$(4.14) \quad N_k^{(n)} := \frac{1}{[d_1-1)(d_2-1)]^{k/2}} \left( \text{CNBW}_k^{(\infty)} - \mathbb{E} \text{CNBW}_k^{(\infty)} \right) \mathbf{1}_{\{k \leq r_n\}}.$$

Let  $\{Z_k\}_{k \geq 2}$  be independent Gaussian random variables with  $\mathbb{E} Z_k = 0$  and  $\mathbb{E} Z_k^2 = 2k$ . Define the weight  $w_k = b_k/(k^2 \log(k+1))$ , where  $(b_k)_{k \in \mathbb{N}}$  is any fixed positive summable sequence.

Let  $P_n$  be the law of the sequence  $(N_k^{(n)})_{k \geq 2}$ . Then as an element in  $\mathcal{P}(\mathcal{X})$ ,  $P_n$  converges weakly to the law of the random vector  $(Z_k)_{k \geq 2}$ .

*Proof.* We first prove the following *Claim (1)*: for any fixed  $r$ ,  $(N_k^{(n)})_{2 \leq k \leq r}$  converges in distribution to  $(Z_k)_{2 \leq k \leq r}$ .

For any fixed  $k$ , when  $n$  is sufficiently large, we can write (4.14) as

$$(4.15) \quad \begin{aligned} N_k^{(n)} &= \frac{1}{[d_1 - 1](d_2 - 1)]^{k/2}} \left( 2kC_k^{(\infty)} - [(d_1 - 1)(d_2 - 1)]^k \right) \\ &\quad + \frac{1}{[d_1 - 1](d_2 - 1)]^{k/2}} \sum_{j|k, j < k} \left( 2jC_j^{(\infty)} - [(d_1 - 1)(d_2 - 1)]^j \right). \end{aligned}$$

Recall  $C_k^{(\infty)}$  is a Poisson random variable with mean  $\frac{(d_1 - 1)^k(d_2 - 1)^k}{2^k}$ . The first term in (4.15) converges in distribution to a centered Gaussian random variable  $Z_k$  with variance  $2k$  as  $n \rightarrow \infty$  from the Gaussian approximation of Poisson distribution.

To show the convergence of  $N_k^{(n)}$  for a fixed  $k$ , it remains to show the second term in (4.15) converges to zero in probability. Note that the second term in (4.15) has a zero mean and its variance is given by

$$\text{Var} \left[ \frac{1}{[d_1 - 1](d_2 - 1)]^{k/2}} \sum_{j|k, j < k} \left( 2jC_j^{(\infty)} - [(d_1 - 1)(d_2 - 1)]^j \right) \right] = \sum_{j|k, j < k} 2j[(d_1 - 1)(d_2 - 1)]^{j-k},$$

which goes to 0 as  $n \rightarrow \infty$ . Then by Chebyshev's inequality, this term converges to 0 in probability. Therefore Claim (1) holds.

We further define  $N_1^{(n)} = 0, Z_1 = 0$ , and consider the weak convergence of  $(N_k^{(n)})_{k \in \mathbb{N}}$  as an element in  $L^2(\vec{w})$ . Since

$$\mathbb{E} \sum_{k=1}^{\infty} (Z_k)^2 w_k = \sum_{k=2}^{\infty} \frac{b_k}{k \log(k+1)} < \infty,$$

$(Z_k)_{k \in \mathbb{N}} \in L^2(\vec{w})$  almost surely. From Claim (1), every sub-sequential limit of  $P_n$  has the same finite-dimensional distributions as  $(Z_k)_{k \in \mathbb{N}}$ . From Proposition 4.1 (3), every sub-sequential weak limit of  $P_n$  in  $\mathcal{P}(\mathcal{X})$  is equal to the law of  $(Z_k)_{k \in \mathbb{N}}$ .

By Prokhorov's Theorem (see for example [57, Chapter 14, Theorem 1.5]), if  $\{P_n\}_{n \in \mathbb{N}}$  is tight, and every weakly convergent sub-sequence has the same limit  $\mu$  in  $\mathcal{P}(\mathcal{X})$ , then the sequence  $\{P_n\}_{n \in \mathbb{N}}$  converges weakly to  $\mu$ . Since we have already shown every sub-sequential weak limit of  $P_n$  is the law of  $(Z_k)_{k \in \mathbb{N}}$  in  $\mathcal{P}(\mathcal{X})$ , to finish the proof, it remains to show  $\{P_n\}_{n \in \mathbb{N}}$  is tight.

From the description of compact sets in  $L^2(\vec{w})$  given in Proposition 4.1 (1), it suffices to show for any  $\varepsilon > 0$ , there exists an element  $(a_k)_{k \in \mathbb{N}} \in L^2(\vec{w})$  with  $a_k > 0, \forall k \in \mathbb{N}$ , such that

$$(4.16) \quad \sup_n \mathbb{P} \left[ \bigcup_{k \in \mathbb{N}} \left\{ |N_k^{(n)}| > a_k \right\} \right] = \sup_n \mathbb{P} \left[ \bigcup_{k=1}^{r_n} \left\{ |N_k^{(n)}| > a_k \right\} \right] < \varepsilon,$$

where  $\bigcup_{k \in \mathbb{N}} \left\{ |N_k^{(n)}| > a_k \right\}$  is the complement of a compact set in  $L^2(\vec{w})$ .

For any fixed  $\varepsilon > 0$ , choose  $a_k = \alpha k \sqrt{\log(k+1)}$  for a constant  $\alpha^2 > 32$  depending on  $\varepsilon$ , then  $(a_k)_{k \in \mathbb{N}} \in L^2(\vec{w})$  and  $a_k > 0, \forall k \in \mathbb{N}$ . According to the definition of  $N_k^{(n)}$  in (4.14), the above Condition (4.16) is equivalent to

$$(4.17) \quad \sup_n \mathbb{P} \left[ \bigcup_{k=1}^{r_n} \left\{ |\text{CNBW}_k^{(\infty)} - \mathbb{E} \text{CNBW}_k^{(\infty)}| > a_k [(d_1 - 1)(d_2 - 1)]^{k/2} \right\} \right] < \varepsilon.$$

From the proof of Theorem 22 in [30],  $\text{CNBW}_k^{(\infty)}$ , as a sum of independent Poisson random variables, satisfies the following concentration inequality: for any  $t > 0$ ,

$$(4.18) \quad \mathbb{P} \left( |\text{CNBW}_k^{(\infty)} - \mathbb{E} \text{CNBW}_k^{(\infty)}| > t \right) \leq 2 \exp \left( -\frac{t}{8k} \log \left( 1 + \frac{t}{2k[(d_1 - 1)(d_2 - 1)]^k} \right) \right).$$

Since  $\log(1 + x) \geq x/2$  for  $x \in [0, 1]$ , we have from (4.18), for sufficiently large  $n$  and all  $k \leq r_n$ ,

$$\begin{aligned} & \mathbb{P} \left( |\text{CNBW}_k^{(\infty)} - \mathbb{E} \text{CNBW}_k^{(\infty)}| > a_k[(d_1 - 1)(d_2 - 1)]^{k/2} \right) \\ & \leq 2 \exp \left( -\frac{a_k[(d_1 - 1)(d_2 - 1)]^{k/2}}{8k} \log \left( 1 + \frac{a_k}{2k[(d_1 - 1)(d_2 - 1)]^{k/2}} \right) \right) \\ & \leq 2 \exp \left( -\frac{a_k^2}{32k^2} \right) = 2(k + 1)^{-\alpha^2/32}. \end{aligned}$$

With the assumption  $\alpha^2 > 32$ , we can make  $\sum_{k=1}^{\infty} 2(k + 1)^{-\alpha^2/32} < \varepsilon$  by choosing a sufficiently large constant  $\alpha$  depending on  $\varepsilon$ , which guarantees (4.16). Hence  $\{P_n\}_{n \in \mathbb{N}}$  is tight. This completes the proof.  $\square$

We now continue to study the eigenvalue fluctuation for  $\frac{XX^\top - d_1 I}{\sqrt{(d_1 - 1)(d_2 - 1)}}$  when  $d_1 \cdot d_2 \rightarrow \infty$ . Before stating the main result, we make several assumptions on the test function  $f$ . Define

$$\Phi_0(x) = 1, \quad \Phi_k(x) = 2T_k(x/2), \quad \forall k \geq 1.$$

Assume  $f$  is an entire function on  $\mathbb{C}$ . Let  $K_1 = \max\{\alpha_1, \alpha_2\}$ , where  $\alpha_1$  and  $\alpha_2$  are the constants in (3.2), (3.3), respectively. Then from Proposition 4.3,  $f$  has the expansion

$$(4.19) \quad f(x) = \sum_{i=0}^{\infty} a_i \Phi_i(x)$$

on  $[-K_1, K_1]$ . Denote

$$f_k(x) := \sum_{i=0}^k a_i \Phi_i(x).$$

Suppose the following conditions hold for  $f$ :

(1) For some  $\alpha > 3/2$  and  $M > 0$ ,

$$(4.20) \quad \sup_{|x| \leq K_1} |f(x) - f_k(x)| \leq M \exp(-\alpha kh(k)),$$

where  $h$  is a function such that  $h(r_n) \geq \log[(d_1 - 1)(d_2 - 1)]$  for a sequence

$$(4.21) \quad r_n = \left\lfloor \frac{\beta \log n}{\log[(d_1 - 1)(d_2 - 1)]} \right\rfloor$$

with a constant  $\beta < 1/\alpha$ .

(2)

$$(4.22) \quad \lim_{n \rightarrow \infty} \left| f_{r_n} \left( \frac{d_1(d_2 - 1)}{\sqrt{(d_1 - 1)(d_2 - 1)}} \right) - f \left( \frac{d_1(d_2 - 1)}{\sqrt{(d_1 - 1)(d_2 - 1)}} \right) \right| = 0.$$

Let  $\mu_k(d_1, d_2) := \mathbb{E} \text{CNBW}_k^{(\infty)}$ . We define the following sequence:

$$(4.23) \quad m_f^{(n)} := n a_0 + \sum_{k=1}^{r_n} \frac{a_k}{[(d_1 - 1)(d_2 - 1)]^{k/2}} \left( \mu_k(d_1, d_2) - n(d_1 - 2) \cdot (d_2 - 1)^{k/2} \mathbf{1}\{k \text{ is even}\} \right).$$

Now we are ready to state our results for eigenvalue fluctuations when  $d_1 \cdot d_2 \rightarrow \infty$ . Here  $d_1, d_2, (\lambda_i)_{1 \leq i \leq n}$  and the matrix  $X$  are quantities depending on  $n$ , but for simplicity of notations, we drop the dependence on  $n$ .

**Theorem 4.7.** *Let  $G_n$  be a sequence of random  $(d_1, d_2)$ -biregular bipartite graphs with*

$$d_1 d_2 \rightarrow \infty, \quad d_1 d_2 = n^{o(1)}.$$

*Let  $\lambda_1 \geq \dots \geq \lambda_n$  be the eigenvalues of  $\frac{XX^\top - d_1 I}{\sqrt{(d_1-1)(d_2-1)}}$ . Suppose one of the following two assumptions holds:*

- (1) *There exists a constant  $c \geq 1$  such that  $1 \leq \frac{d_1}{d_2} \leq c$ .*
- (2) *There exists a constant  $c_1$  such that  $d_2 \leq c_1$  for all  $n$ .*

*Let  $f$  be an entire function on  $\mathbb{C}$  satisfying (4.20) and (4.22). Then as  $n \rightarrow \infty$ , the random variable*

$$(4.24) \quad Y_f^{(n)} = \sum_{i=1}^n f(\lambda_i) - m_f^{(n)}$$

*converges in distribution to a centered Gaussian random variable with variance  $\sigma_f^2 = 2 \sum_{k=2}^{\infty} k a_k^2$ .*

*Moreover, for any fixed  $t$ , consider the entire functions  $g_1, \dots, g_t$  satisfying (4.20) and (4.22). The corresponding random vector  $(Y_{g_1}^{(n)}, \dots, Y_{g_t}^{(n)})$  converges in distribution to a centered Gaussian random vector  $(Z_{g_1}, \dots, Z_{g_t})$  with covariance*

$$(4.25) \quad \text{Cov}(Z_{g_i}, Z_{g_j}) = 2 \sum_{k=2}^{\infty} k a_k(g_i) a_k(g_j)$$

*for  $1 \leq i, j \leq t$ , where  $a_k(g_i), a_k(g_j)$  are the  $k$ -th coefficients in the expansion (4.19) for  $g_i, g_j$ , respectively.*

*Proof.* We first prove the CLT for a single test function  $f$ . Define

$$\begin{aligned} X_f^{(n)} &:= \sum_{k=2}^{r_n} \frac{a_k}{[(d_1-1)(d_2-1)]^{k/2}} \text{CNBW}_k^{(n)} - \mathbb{E} \sum_{k=2}^{r_n} \frac{a_k}{[(d_1-1)(d_2-1)]^{k/2}} \text{CNBW}_k^{(\infty)}, \\ \tilde{X}_f^{(n)} &:= \sum_{k=2}^{r_n} \frac{a_k}{[(d_1-1)(d_2-1)]^{k/2}} \text{CNBW}_k^{(\infty)} - \mathbb{E} \sum_{k=2}^{r_n} \frac{a_k}{[(d_1-1)(d_2-1)]^{k/2}} \text{CNBW}_k^{(\infty)}. \end{aligned}$$

Recall the definition of  $m_f(n)$  in (4.23). From (2.46), (2.47), and (2.53),  $X_f^{(n)}$  can be written as

$$\begin{aligned} X_f^{(n)} &= \sum_{k=2}^{r_n} \sum_{i=1}^n a_k \Gamma_k(\lambda_i) - \sum_{k=2}^{r_n} \frac{a_k}{[(d_1-1)(d_2-1)]^{k/2}} \mu_k(d_1, d_2) \\ &= \sum_{k=2}^{r_n} \sum_{i=1}^n \left( 2a_k T_k(\lambda_i/2) + \frac{a_k(d_1-2)}{(d_1-1)^{k/2}} \mathbf{1}_{\{k \text{ is even}\}} \right) - \sum_{k=2}^{r_n} \frac{a_k}{[(d_1-1)(d_2-1)]^{k/2}} \mu_k(d_1, d_2) \\ &= \sum_{i=1}^n f_{r_n}(\lambda_i) - n a_0 + \sum_{k=2}^{r_n} \frac{n a_k(d_1-2)}{(d_1-1)^{k/2}} \mathbf{1}_{\{k \text{ is even}\}} - \sum_{k=2}^{r_n} \frac{a_k}{[(d_1-1)(d_2-1)]^{k/2}} \mu_k(d_1, d_2) \\ &= \sum_{i=1}^n f_{r_n}(\lambda_i) - m_f^{(n)}, \end{aligned}$$

where in the third line we use the fact given in (2.53) that

$$\sum_{i=1}^n 2a_1 T_1(\lambda_i/2) = \sum_{i=1}^n a_1 \Gamma_1(\lambda_i) = a_1 [(d_1-1)(d_2-1)]^{-1/2} \text{CNBW}_1^{(n)} = 0.$$

From the definition of  $N_k^{(n)}$  in (4.14),

$$\tilde{X}_f^{(n)} = \sum_{k=2}^{r_n} a_k N_k^{(n)}.$$

By Lemma 4.6 and Proposition 4.1 (2),  $\tilde{X}_f^{(n)}$  converges in distribution to a centered Gaussian random variable with variance  $\sigma_f = \sum_{k=2}^{\infty} 2ka_k^2$ .

From Corollary 2.15, the total variation distance between  $X_f^{(n)}$  and  $\tilde{X}_f^{(n)}$  satisfies

$$\begin{aligned} d_{\text{TV}}(X_f^{(n)}, \tilde{X}_f^{(n)}) &\leq d_{\text{TV}}\left((\text{CNBW}_k^{(n)}, 2 \leq k \leq r_n), (\text{CNBW}_k^{(\infty)}, 2 \leq k \leq r_n)\right) \\ &\leq \frac{c_8 \sqrt{r_n} [(d_1 - 1)(d_2 - 1)]^{3r_n/2}}{nd_1}, \end{aligned}$$

which converges to 0 as  $n \rightarrow \infty$  from the assumption (4.21). Therefore  $X_f^{(n)}$  and  $\tilde{X}_f^{(n)}$  converge to the same limit.

It remains to show  $Y_f^{(n)}$  and  $X_f^{(n)}$  converge in distribution to the same limit. We have  $f_{r_n}(\lambda_1) \rightarrow f(\lambda_1)$  as  $n \rightarrow \infty$  from (4.22). Then for any  $\delta > 0$ ,  $|f(\lambda_1) - f_{r_n}(\lambda_1)| \leq \delta/2$  for sufficiently large  $n$ .

Suppose that all the non-trivial eigenvalues are contained in  $[-K_1, K_1]$ . From Condition (4.20), we have for sufficiently large  $n$ ,

$$\left| Y_f^{(n)} - X_f^{(n)} \right| \leq \sum_{i=1}^n |f(\lambda_i) - f_{r_n}(\lambda_i)| \leq \frac{\delta}{2} + (n-1)M \exp(-\alpha r_n h(r_n)) \leq \frac{\delta}{2} + Mn^{1-\alpha\beta} < \delta.$$

Therefore

$$(4.26) \quad \mathbb{P}\left(\left| Y_f^{(n)} - X_f^{(n)} \right| \geq \delta\right) \leq \mathbb{P}\left(\max_{2 \leq i \leq n} |\lambda_i| \geq K_1\right) = o(1),$$

where the last inequality is from parts (2) and (3) in Theorem 3.1. Hence  $Y_f^{(n)}$  and  $X_f^{(n)}$  converge in distribution to the same limit. This proves the CLT for (4.24).

We now extend the results to a random vector  $(Y_{g_1}, \dots, Y_{g_t})$ . By Lemma 4.6 and part (2) in Proposition 4.1, the random vector  $(\tilde{X}_{g_1}^{(n)}, \dots, \tilde{X}_{g_t}^{(n)})$  converges in distribution to the Gaussian random vector  $(Z_{g_1}, \dots, Z_{g_t})$  with covariance given in (4.25).

Note that each entry in the vector  $(X_{g_1}^{(n)}, \dots, X_{g_t}^{(n)})$  is a measurable function of  $(\text{CNBW}_k^{(n)})_{2 \leq k \leq r_n}$ , and we can find a measurable map  $\psi$  such that

$$\psi((\text{CNBW}_k^{(n)})_{2 \leq k \leq r_n}) = (X_{g_1}^{(n)}, \dots, X_{g_t}^{(n)}), \quad \psi((\text{CNBW}_k^{(\infty)})_{2 \leq k \leq r_n}) = (\tilde{X}_{g_1}^{(n)}, \dots, \tilde{X}_{g_t}^{(n)}).$$

Since any measurable map reduces the total variation distance between two random variables, we obtain from (2.41),

$$\begin{aligned} &d_{\text{TV}}\left((X_{g_1}^{(n)}, \dots, X_{g_t}^{(n)}), (\tilde{X}_{g_1}^{(n)}, \dots, \tilde{X}_{g_t}^{(n)})\right) \\ &\leq d_{\text{TV}}\left((\text{CNBW}_k^{(n)}, 2 \leq k \leq r_n), (\text{CNBW}_k^{(\infty)}, 2 \leq k \leq r_n)\right) \\ &\leq \frac{c_8 \sqrt{r_n} [(d_1 - 1)(d_2 - 1)]^{3r_n/2}}{nd_1} = o(1). \end{aligned}$$

Therefore  $(X_{g_1}^{(n)}, \dots, X_{g_t}^{(n)})$  converges in distribution to  $(Z_{g_1}, \dots, Z_{g_t})$ . Finally, according to (4.26),  $(X_{g_1}^{(n)}, \dots, X_{g_t}^{(n)})$  and  $(Y_{g_1}^{(n)}, \dots, Y_{g_t}^{(n)})$  converge in distribution to the same limit. This finishes the proof.  $\square$

*Remark 4.8.* In [24], the authors proved a CLT for linear spectral statistics for normalized sample covariance matrices  $A = \frac{1}{\sqrt{np}}(XX^\top - pI)$ , where  $p/n \rightarrow \infty$  and  $X = (X_{ij})_{n \times p}$  has i.i.d. entries with mean 0 variance 1. It is shown in Theorem 1 of [24] that the fluctuations of linear statistics for two analytic functions  $g_1, g_2$  converge in distribution to a centered Gaussian vector with covariance given by  $(\nu_4 - 3)a_1(g_1)a_1(g_2) + 2\sum_{k=1}^{\infty} ka_k(g_1)a_k(g_2)$ , where  $\nu_4 = \mathbb{E}X_{11}^4$ . The covariance given in (4.25) is the same, except for the fact that the coefficient in front of  $a_1(f_1)a_1(f_2)$  is 0. This can be explained by the fact that the number of 2-cycles is 0 in RBBGs, whereas in the model used in [24] it is not. The same phenomenon was also observed in uniform random regular graphs [40], where the limiting variance is the same as the eigenvalue fluctuations for the GOE except for the first two terms, see Remark 22 in [40].

## 5. GLOBAL SEMICIRCLE LAW

Consider a random  $(n, m, d_1, d_2)$ -biregular bipartite graph with  $d_1 \geq d_2$ . We assume  $d_1, d_2$  satisfy the following:

$$(5.1) \quad \lim_{n \rightarrow \infty} d_1 = \infty,$$

$$(5.2) \quad d_1 = o(n^\varepsilon), \quad \forall \varepsilon > 0,$$

$$(5.3) \quad \frac{d_1}{d_2} \rightarrow \infty.$$

Here  $d_2$  can be fixed or a parameter depending on  $n$ . In this section, we prove a semicircle law for the matrix  $\frac{XX^\top - d_1 I}{\sqrt{(d_1-1)(d_2-1)}}$  under the assumptions (5.1)-(5.3).

For RBBGs in this regime, we have the locally tree-like structure in the following sense. Let  $R$  be fixed and  $\tau_1$  be the set of vertices in  $V_1$  without any cycles in the  $R$ -neighborhood. The following lemma holds.

**Lemma 5.1.** *Then under Condition (5.2),*

$$\mathbb{P}\left(\frac{n - |\tau_1|}{n} > n^{-1/4}\right) = o(n^{-5/4}).$$

To prove Lemma 5.1, the following estimates on the expectation and variance of the cycle counts of RBBGs given in [29] are needed.

**Lemma 5.2** (Proposition 4 in [29]). *Let  $C_k$  be the number of cycles of length  $2k$  in a random  $(d_1, d_2)$ -biregular bipartite graph. Denote  $\mu_k = \frac{[(d_1-1)(d_2-1)]^k}{2k}$ . If  $d_1 = o(n)$ ,  $k = O(\log n)$  and  $kd_1 = o(n)$ , then*

$$(5.4) \quad \mathbb{E}C_k = \mu_k \left(1 + O\left(\frac{k(k+d_1)}{n}\right)\right),$$

$$(5.5) \quad \text{Var}[C_k] = \mu_k \left(1 + O\left(\frac{d_1^{2k}(k(d_1/d_2)^{2k-1} + (d_1/d_2)^{-k}d_2)}{n}\right)\right).$$

*Proof of Lemma 5.1.* From Lemma 5.2, for each fixed  $k$ , under Condition (5.2),

$$(5.6) \quad \mathbb{E}C_k = (1 + o(1))\mu_k, \quad \text{Var}[C_k] = (1 + o(1))\mu_k.$$

If a vertex  $v_1 \in V_1$  is not in  $\tau_1$ , then for some  $s$  with  $2 \leq s \leq R$ , there exists a  $2s$ -cycle within  $(R-s)$ -neighborhood of  $v_1$ . Hence the size of all  $(R-s)$ -neighborhoods of  $2s$ -cycles from  $V_1$  gives an upper bound on  $(n - |\tau_1|)$ .

For any  $2s$ -cycle, the size of its  $(R-s)$ -neighborhood from  $V_1$  is bounded by

$$c_1 s [(d_1 - 1)(d_2 - 1)]^{(R-s)/2+1}$$

with an absolute constant  $c_1$ . Define

$$N_R := c_1 \sum_{s=2}^R s[(d_1 - 1)(d_2 - 1)]^{(R-s)/2+1} C_s.$$

We then have  $n - \tau_1 \leq N_R$ . From (5.6),  $\mathbb{E}N_R = O([(d_1 - 1)(d_2 - 1)]^{R+1})$ .

Recall  $R$  is fixed. By Cauchy inequality,

$$\text{Var}[N_R] \leq c_1^2 R \sum_{s=2}^R s^2 [(d_1 - 1)(d_2 - 1)]^{R-s+2} \text{Var}[C_s] = O([(d_1 - 1)(d_2 - 1)]^{R+2}).$$

Then from Markov's inequality, together with our assumptions (5.1)-(5.3),

$$\begin{aligned} \mathbb{P}\left(\frac{n - |\tau_1|}{n} > n^{-1/4}\right) &= \mathbb{P}(n - |\tau_1| > n^{3/4}) \leq \mathbb{P}(N_R \geq n^{3/4}) \\ &\leq \frac{\mathbb{E}[N_R^2]}{n^{3/2}} = O([(d_1 - 1)(d_2 - 1)]^{2R+2} n^{-3/2}) = o(n^{-5/4}). \end{aligned}$$

□

We now state our main result in this section. The proof is based on the moment method and the tree approximation of local neighborhoods, which were previously applied to random regular graphs in [31].

**Theorem 5.3.** *Let  $G_n$  be a sequence of random  $(d_1, d_2)$ -biregular bipartite graph. Under assumptions (5.1)-(5.3), the empirical spectral distribution of  $\frac{XX^\top - d_1 I}{\sqrt{(d_1-1)(d_2-1)}}$  converges weakly to the semicircle law almost surely.*

*Remark 5.4.* Recall in [29], when the ratio  $d_1/d_2 \geq 1$  converges to a positive constant, the ESD of  $\frac{XX^\top}{d_1}$  converges to Marčenko-Pastur law. With different scaling parameters, we obtain a different semicircle law when  $d_1/d_2 \rightarrow \infty$ . This can be seen as an analog of the semicircle law for sample covariance matrices proved in [7] when the aspect ratio is unbounded.

*Proof of Theorem 5.3.* Note that for all  $i \in V_1$ , by the degree constraint,

$$(5.7) \quad (XX^\top)_{ii} = \sum_j X_{ij} X_{ji} = \sum_j X_{ij} = \deg(i) = d_1.$$

Denote  $M = \frac{XX^\top - d_1 I}{\sqrt{(d_1-1)(d_2-1)}}$ . We start with the trace expansion of  $M$ .

$$\begin{aligned} (5.8) \quad \frac{1}{n} \text{tr} M^k &= \frac{1}{n((d_1-1)(d_2-1))^{k/2}} \text{tr}(XX^\top - d_1 I)^k \\ &= \frac{1}{n((d_1-1)(d_2-1))^{k/2}} \sum_{\substack{i_1, \dots, i_k \in [n] \\ i_1 \neq i_2, \dots, i_k \neq i_1 \\ j_1, \dots, j_k \in [m]}} X_{i_1 j_1} X_{i_2 j_1} \cdots X_{i_k j_k} X_{i_1 j_k}. \end{aligned}$$

From (5.7), the diagonal entries of  $XX^\top - d_1 I$  are 0, therefore we have the constraint that  $i_1 \neq i_2, \dots, i_k \neq i_1$  in (5.8).

Let  $A_k^{r,c}(v, v)$  be the number of all closed walks of length  $2k$  in  $G$  starting from  $v \in V_1$  that use  $r$  distinct vertices from  $V_1$ ,  $c$  distinct vertices from  $V_2$ , with the restriction that  $i_1 \neq i_2, \dots, i_k \neq i_1$ . We have  $r \leq k+1$  and  $c \leq k$ , since there are at most  $k+1$  vertices in  $V_1$  and  $k$  vertices in  $V_2$  that

are visited in one closed walk of length  $2k$ . From (5.8), the  $k$ -th moment of the empirical spectral distribution  $\mu_n$  satisfies

$$(5.9) \quad \int x^k d\mu_n(x) = \frac{1}{n} \text{tr} M^k = \frac{1}{n((d_1-1)(d_2-1))^{k/2}} \sum_{v \in V_1} \sum_{r=1}^{k+1} \sum_{c=1}^k A_k^{r,c}(v, v).$$

Since  $d_1 \geq d_2$ , for any fixed  $v \in V_1$ , we have

$$\sum_{r \leq k+1, c \leq k} A_k^{r,c}(v, v) \leq d_1^{2k}.$$

For ease of notation, in the following equations, we often omit the range of  $r, c$  in the summation.

We may decompose the sum in (5.9) into two parts depending on whether  $v \in \tau_1$  or not. For any  $v \in \tau_1$ , we write  $A_k^{r,c} =: A_k^{r,c}(v, v)$  since all neighborhoods of  $v \in \tau_1$  of radius  $k$  looks the same and the number of such closed walks is independent of  $v$ . Now we have the following upper bound on (5.9):

$$\begin{aligned} \int x^k d\mu_n(x) &\leq \frac{1}{n((d_1-1)(d_2-1))^{k/2}} \sum_{v \in \tau_1} \sum_{r,c} A_k^{r,c}(v, v) + \frac{(n - |\tau_1|)d_1^{2k}}{n((d_1-1)(d_2-1))^{k/2}} \\ &= \frac{|\tau_1|}{n((d_1-1)(d_2-1))^{k/2}} \sum_{r,c} A_k^{r,c} + \frac{(n - |\tau_1|)d_1^{2k}}{n((d_1-1)(d_2-1))^{k/2}} \\ &\leq \frac{1}{((d_1-1)(d_2-1))^{k/2}} \sum_{r,c} A_k^{r,c} + \frac{(n - |\tau_1|)d_1^{2k}}{n((d_1-1)(d_2-1))^{k/2}}. \end{aligned}$$

Similarly, a lower bound holds by only counting closed walks starting with vertices in  $\tau_1$ :

$$\int x^k d\mu_n(x) \geq \frac{1}{n((d_1-1)(d_2-1))^{k/2}} \sum_{v \in \tau_1} \sum_r \sum_c A_k^{r,c}(v, v) = \frac{|\tau_1|}{n((d_1-1)(d_2-1))^{k/2}} \sum_r \sum_c A_k^{r,c}.$$

From Lemma 5.1 and assumption (5.2), with probability at least  $1 - o(n^{-5/4})$ , for any fixed  $k \geq 0$ ,

$$\frac{(n - |\tau_1|)}{n((d_1-1)(d_2-1))^{k/2}} d_1^{2k} = o(1), \quad \text{and} \quad \frac{|\tau_1|}{n} = 1 - o(n^{-1/4}).$$

To show the almost sure convergence of the empirical measure to semicircle law, by the upper and lower bounds above, it suffices to show

$$(5.10) \quad \lim_{n \rightarrow \infty} \frac{1}{((d_1-1)(d_2-1))^{k/2}} \sum_{r,c} A_k^{r,c} = \begin{cases} 0 & \text{if } k \text{ is odd,} \\ C_{k/2} & \text{if } k \text{ is even,} \end{cases}$$

where  $C_k := \frac{1}{k+1} \binom{2k}{k}$  is the  $k$ -th Catalan number.

Recall  $A_k^{r,c}$  counts the closed walks of length  $2k$  on a rooted  $(d_1, d_2)$ -biregular tree starting from a root with degree  $d_1$ , ending at the same root. Now we consider the quantity

$$\frac{1}{((d_1-1)(d_2-1))^{k/2}} \sum_{r,c} A_k^{r,c}$$

more carefully. We first consider possible ranges of  $r$  and  $c$  in the expression above.

The walk  $(i_1, j_1, i_2, j_2, \dots, i_k, j_k, i_1)$  in the summation satisfies  $i_1 \neq i_2, \dots, i_{k-1} \neq i_k, i_k \neq i_1$ . This implies when a walk goes from  $i_t$  to  $j_t$  for some  $t$ , it cannot backtrack immediately to  $i_t$ . Namely, any such walk is not allowed to backtrack at even depths (here, we define the depth of the root in a tree as 1). To have a closed walk of length  $2k$  on a tree, each edge is repeated at

least twice, so the number of distinct edges is at most  $k$ . Therefore the number of distinct vertices satisfies

$$(5.11) \quad r + c \leq k + 1.$$

For fixed  $r$  and  $c$ , the number of such unlabeled rooted trees with  $r + c - 1$  distinct edges is  $C_{r+c-1}$ . Let  $I$  be the set of vertices in the odd depths of the biregular tree and  $J$  be the set of vertices in the even depths. Since the first vertex of the walk is fixed (we always start from the fixed root), for any closed walk, there are at most  $d_1^c$  many ways to choose distinct vertices from  $J$  and  $d_2^{r-1}$  many ways to choose distinct vertices from  $I$ . Therefore we have

$$(5.12) \quad A_k^{r,c} \leq d_1^c d_2^{r-1} C_{r+c-1} \leq d_1^c d_2^{r-1} C_k,$$

where the last inequality is from (5.11). We also know that  $r - 1 \geq c$ , because whenever a new vertex in  $J$  is reached by the walk, the walk cannot backtrack, so it must reach a new vertex in  $I$ . Therefore we have

$$c \leq r - 1 \quad \text{and} \quad r + c \leq k + 1,$$

which implies the following conditions on  $c$  and  $r$ :

$$(5.13) \quad c \leq k/2 \quad \text{and} \quad r - 1 \leq k - c.$$

From (5.12), for any  $(r, c)$  satisfying (5.13), the following holds:

$$(5.14) \quad \frac{A_k^{r,c}}{((d_1 - 1)(d_2 - 1))^{k/2}} \leq \frac{d_1^c}{(d_1 - 1)^{k/2}} \frac{d_2^{r-1}}{(d_2 - 1)^{k/2}} C_k \leq \frac{d_1^c}{(d_1 - 1)^{k/2}} \frac{d_2^{k-c}}{(d_2 - 1)^{k/2}} C_k.$$

Now we discuss two cases depending on the parity of  $k$ . When  $k$  is odd, from (5.13),  $c \leq \frac{k-1}{2}$ . Since  $d_1/d_2 \rightarrow \infty$ , we obtain

$$(5.15) \quad \frac{1}{((d_1 - 1)(d_2 - 1))^{k/2}} \sum_{r,c} A_k^{r,c} \leq \left(\frac{d_1}{d_2}\right)^c \frac{d_2^k C_k}{[(d_1 - 1)(d_2 - 1)]^{k/2}} \leq \left(\frac{d_1}{d_2}\right)^{c-k/2} 2^k C_k = o(1).$$

When  $k$  is even, to have a non-vanishing term in the limit for  $A_k^{r,c}$ , we must have  $c = k/2$  and  $r = k/2 + 1$ . Then we have

$$(5.16) \quad \frac{1}{((d_1 - 1)(d_2 - 1))^{k/2}} \sum_{r,c} A_k^{r,c} = \frac{1}{((d_1 - 1)(d_2 - 1))^{k/2}} A_k^{k/2+1,k/2} + o(1).$$

We continue our proof with a more refined estimate on  $A_k^{k/2+1,k/2}$ . Since every edge is repeated exactly twice in the closed walk, it's a depth-first search on the biregular tree.

If the root is at level 1, and subsequent vertices are at a level  $i + 1$  where  $i$  is the distance from the root, then all leaves must be at odd levels, since we can never backtrack at an even level. This implies that every vertex at an even level has at least one child, which means  $r \geq c + 1$ , with equality if and only if every vertex at an even level has *exactly* one child. Thus, one can see the tree as a subdivision of a smaller tree, where a vertex has been introduced on each edge (the “new” vertices being the vertices on an even level in the bigger tree). This is a bijection between the kind of planar rooted tree on  $k + 1$  vertices we are trying to count and the set of all planar rooted trees on  $k/2 + 1$  vertices. There are  $C_{k/2}$  of the latter. See Figure 3 for an example of a valid closed walk and an illustration of the aforementioned bijection.

Moreover, given a fixed root with a vertex label, the number of all possible ways to label the tree with vertices in a biregular bipartite graph is between  $d_1^{k/2} (d_2 - 1)^{k/2}$  and  $(d_1 - k/2)^{k/2} (d_2 - 1)^{k/2}$ , so the following inequality for  $A_k^{k/2+1,k/2}$  holds:

$$(5.17) \quad (d_1 - k/2 - 1)^{k/2} (d_2 - 1)^{k/2} C_{k/2} \leq A_k^{k/2+1,k/2} \leq d_1^{k/2} (d_2 - 1)^{k/2} C_{k/2}.$$

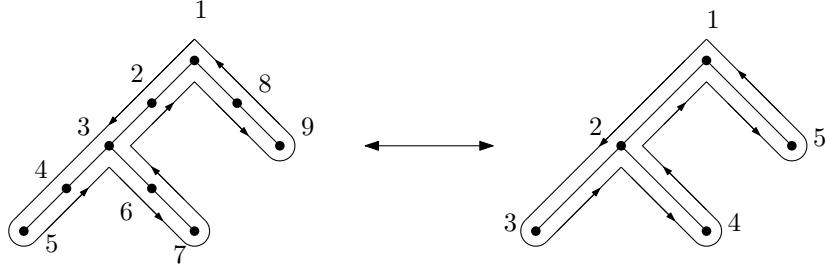


FIGURE 3. On the left we have a closed walk  $(1, 2, 3, 4, 5, 4, 3, 6, 7, 6, 3, 2, 1, 8, 9, 8, 1)$  on a rooted planar tree which only backtracks at odd depths, and the tree has no new branches at any even depth along the walk. Its correspondent under the bijection is the closed walk  $(1, 2, 3, 2, 4, 2, 1, 5, 1)$  on the smaller rooted planar tree induced by the depth-first search on the right.

From (5.16) and (5.17), we obtain for even  $k$ ,

$$(5.18) \quad \lim_{n \rightarrow \infty} \frac{1}{((d_1 - 1)(d_2 - 1))^{k/2}} \sum_{r,c} A_k^{r,c} = C_{k/2}.$$

With (5.15) and (5.18), the asymptotic behavior of moments given in (5.10) holds. This completes the proof of Theorem 5.3.  $\square$

## 6. RANDOM REGULAR HYPERGRAPHS

We first include some definitions for hypergraphs and describe a bijection between a subset of biregular bipartite graphs and the set of regular hypergraphs studied in [32]. We will use the map given in Definition 6.2 to apply some of our results for RBBGs to random regular hypergraphs, see [32] for more details.

**Definition 6.1** (hypergraph). A *hypergraph*  $H$  consists of a set  $V$  of vertices and a set  $E$  of hyperedges such that each hyperedge is a nonempty set of  $V$ . A hypergraph  $H$  is  $k$ -uniform for an integer  $k \geq 2$  if every hyperedge  $e \in E$  contains exactly  $k$  vertices. The *degree* of  $i$ , denoted  $\deg(i)$ , is the number of all hyperedges incident to  $i$ . A hypergraph is  $d$ -regular if all of its vertices have degree  $d$ . A hypergraph is  $(d, k)$ -regular if it is both  $d$ -regular and  $k$ -uniform.

**Definition 6.2** (incidence matrix and associated bipartite graph). A vertex  $i$  is *incident* to a hyperedge  $e$  if and only  $v$  is an element of  $e$ . We can define the *incidence matrix*  $X$  of a hypergraph  $H = (V, E)$  to be a  $|V| \times |E|$  matrix indexed by elements in  $V$  and  $E$  such that  $X_{i,e} = 1$  if  $i \in e$  and 0 otherwise. Moreover, if we regard  $X$  as the adjacency matrix of a graph, it defines a bipartite graph  $G$  with two vertex sets  $V$  and  $E$ . We call  $G$  the *bipartite graph associated to  $H$* , given by a map  $\Phi$  (so  $\Phi(H) = G$ ). See Figure 4 for an example.

**Definition 6.3** (adjacency matrix). For a hypergraph  $H$  with  $n$  vertices, we associate a  $n \times n$  symmetric matrix  $A$  called the *adjacency matrix* of  $H$ . For  $i \neq j$ , we define  $A_{ij}$  as the number of hyperedges containing both  $i$  and  $j$ ; we define  $A_{ii} = 0$  for all  $1 \leq i \leq n$ . When the hypergraph is 2-uniform (i.e., it is a graph), this is the usual definition for the adjacency matrix of a graph.

The following lemma connects the adjacency matrix of a regular hypergraph with its associated biregular bipartite graph. It formally appears in [45, 32], and it is also informally mentioned in [34].

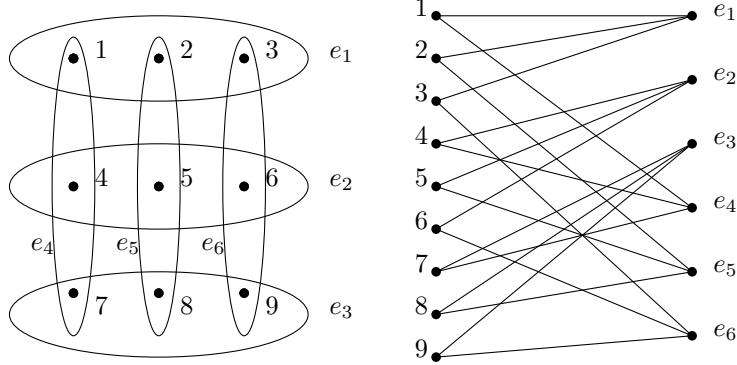


FIGURE 4. a  $(2, 3)$ -regular hypergraph and its associated biregular bipartite graph where all vertices in  $V_2$  have different neighborhoods in  $V_1$

**Lemma 6.4** (Lemma 4.5 in [32]). *Let  $H$  be a  $(d_1, d_2)$ -regular hypergraph, and let  $G$  be the corresponding  $(d_1, d_2)$ -biregular bipartite graph. Let  $A_H$  be the adjacency matrix of  $H$  and  $A_G$  be the adjacency matrix of  $G$  given by*

$$(6.1) \quad A_G = \begin{pmatrix} 0 & X \\ X^\top & 0 \end{pmatrix}.$$

*Then  $A_H = XX^\top - d_1 I$ .*

**Definition 6.5** (walks and cycles). A *walk* of length  $l$  on a hypergraph  $H$  is a vertex-hyperedge sequence  $(i_0, e_1, i_1, \dots, e_l, i_l)$  such that  $i_{j-1} \neq i_j$  and  $\{i_{j-1}, i_j\} \subset e_j$  for all  $1 \leq j \leq l$ . A walk is closed if  $i_0 = i_l$ . A *cycle* of length  $l$  in a hypergraph  $H$  is a closed walk  $(v_0, e_1, \dots, v_{l-1}, e_l, v_0)$  such that all edges are distinct and all vertices are distinct subject to  $v_{l+1} = v_0$ . In the associated bipartite graph  $G$ , a cycle of length  $2l$  corresponds to a cycle of length  $l$  in  $H$ .

Let  $\mathcal{G}(n, m, d_1, d_2)$  be the set of all simple biregular bipartite random graphs with vertex set  $V = V_1 \cup V_2$  such that  $|V_1| = n, |V_2| = m$ , and every vertex in  $V_i$  has degree  $d_i$  for  $i = 1, 2$ . Without loss of generality, we assume  $d_1 \geq d_2$ . Let  $\mathcal{H}(n, d_1, d_2)$  be the set of all simple (without multiple hyperedges)  $(d_1, d_2)$ -regular hypergraphs with labeled vertex set  $[n]$  and  $\frac{nd_1}{d_2}$  many labeled hyperedges denoted by  $\{e_1, \dots, e_{nd_1/d_2}\}$ .

*Remark 6.6.* We can also consider all  $(d_1, d_2)$ -regular hypergraphs with labeled vertices and unlabeled hyperedges. Since all hyperedges are distinct, any such regular hypergraph with unlabeled hyperedges corresponds to  $(nd_1/d_2)!$  regular hypergraphs with labeled hyperedges.

It is well known (see for example [34]) that the map  $\Phi$  defined can be extended to a bijection  $\tilde{\Phi}$  between labeled regular multi-hypergraphs and biregular bipartite graphs. See Figure 4 as an example of the bijection. For a given biregular bipartite graph, if there are two vertices in  $V_2$  that have the same set of neighbors in  $V_1$ , the corresponding regular hypergraph will have multiple hyperedges, see Figure 5. Let  $\mathcal{G}'(n, m, d_1, d_2)$  be a subset of  $\mathcal{G}(n, m, d_1, d_2)$  such that for any  $G \in \mathcal{G}'(n, m, d_1, d_2)$ , any two vertices in  $V_2$  have different neighborhoods in  $V_1$ . The following lemma holds.

**Lemma 6.7** (Lemma 4.2 in [32]).  *$\Phi$  is the restriction of the bijection  $\tilde{\Phi}$  to  $\mathcal{H}(n, d_1, d_2)$  and its image is  $\mathcal{G}'(n, m, d_1, d_2)$ . Hence  $|\mathcal{H}(n, d_1, d_2)| = |\mathcal{G}'(n, m, d_1, d_2)|$ .*

From Lemma 6.7, the uniform distribution on  $\mathcal{G}'(n, m, d_1, d_2)$  for biregular bipartite graphs induces the uniform distribution on  $\mathcal{H}(n, d_1, d_2)$  for regular hypergraphs. With this observation, we can translate some of the results for spectra of random biregular bipartite graphs into results

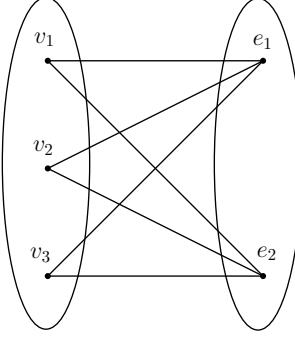


FIGURE 5. a subgraph in a biregular bipartite graph which gives multiple hyperedges  $e_1$  and  $e_2$  in the corresponding regular hypergraph

for spectra of random regular hypergraphs. A similar approach was applied in [12] to enumerate uniform hypergraphs with given degrees.

**Lemma 6.8** (Lemma 4.8 in [32]). *Let  $G$  be a random biregular bipartite graph sampled uniformly from  $\mathcal{G}(n, m, d_1, d_2)$  such that  $3 \leq d_2 \leq d_1 \leq \frac{n}{32}$ . Let  $\mathcal{G}'(n, m, d_1, d_2)$  be the set of biregular bipartite graphs corresponding to simple regular hypergraphs. Then*

$$(6.2) \quad \mathbb{P}(G \in \mathcal{G}'(n, m, d_1, d_2)) \geq 1 - \left(\frac{nd_1}{d_2}\right)^2 \left(\frac{4ed_2}{n}\right)^{d_2}.$$

In particular,

$$(6.3) \quad \mathbb{P}(G \in \mathcal{G}'(n, m, d_1, d_2)) = 1 - O\left(\frac{d_1^2}{nd_2^2}\right).$$

Lemma 6.8 implies the following total variation bound.

**Lemma 6.9** (total variation bound). *Let  $\mu_n$  be the probability measure of the random  $(d_1, d_2)$ -regular hypergraph with  $n$  vertices induced on the set of all  $(n, m, d_1, d_2)$ -biregular bipartite graphs, and let  $\mu'_n$  be the uniform measure on the set of all  $(n, m, d_1, d_2)$ -biregular bipartite graphs. We have*

$$(6.4) \quad d_{\text{TV}}(\mu_n, \mu'_n) \leq \left(\frac{nd_1}{d_2}\right)^2 \left(\frac{4ed_2}{n}\right)^{d_2}.$$

*Proof.* Since  $\mathcal{G}'(n, m, d_1, d_2)$  is the set of all biregular bipartite graphs that are bijective to regular hypergraphs. We have  $\mu_n(\mathcal{G}'(n, m, d_1, d_2)) = 1$  and  $\mu'_n(\mathcal{G}'(n, m, d_1, d_2)) = \frac{|\mathcal{G}'(n, m, d_1, d_2)|}{|\mathcal{G}(n, m, d_1, d_2)|}$ . Let  $\mathcal{F}$  be the power set of  $\mathcal{G}(n, m, d_1, d_2)$ . Taking into account the fact that both  $\mu_n$  and  $\mu'_n$  are uniform measures, we obtain that

$$\begin{aligned} d_{\text{TV}}(\mu_n, \mu'_n) &= \sup_{A \in \mathcal{F}} |\mu_n(A) - \mu'_n(A)| = |1 - \mu'_n(\mathcal{G}'(n, m, d_1, d_2))| \\ &= \mathbb{P}(G \notin \mathcal{G}'(n, m, d_1, d_2)) \leq \left(\frac{nd_1}{d_2}\right)^2 \left(\frac{4ed_2}{n}\right)^{d_2}, \end{aligned}$$

where the last inequality is from Lemma 6.8.  $\square$

Equipped with Lemma 6.9, we obtain several corollaries for random regular hypergraphs in the following subsections.

**6.1. Cycle counts.** Recall the definition of cycles in a hypergraph given in Definition 6.5. Let  $C_k$  be the number of cycles of length  $k$  in a  $(d_1, d_2)$ -regular hypergraph. The following result holds.

**Corollary 6.10.** *Let  $H$  be a  $(d_1, d_2)$ -random regular hypergraph with cycle counts  $(C_k, k \geq 2)$ . Let  $(Z_k, k \geq 2)$  be independent Poisson random variables with  $\mathbb{E} Z_k = \frac{(d_1 - 1)^k (d_2 - 1)^k}{2k}$ . For any  $n, m \geq 1$ ,  $r \geq 3$ , and  $3 \leq d_2 \leq d_1 \leq \frac{n}{32}$ ,*

$$d_{\text{TV}}((C_2, \dots, C_r), (Z_2, \dots, Z_r)) \leq \frac{c_6 \sqrt{r} (d_1 - 1)^{3r/2} (d_2 - 1)^{3r/2}}{nd_1} + \left( \frac{nd_1}{d_2} \right)^2 \left( \frac{4ed_2}{n} \right)^{d_2}.$$

*Proof.* Let  $\tilde{C}_k$  be the number of cycles with length  $2k$  in a uniform random  $(d_1, d_2)$ -biregular bipartite graph. From Lemma 6.9,

$$(6.5) \quad d_{\text{TV}}((C_2, \dots, C_r), (\tilde{C}_2, \dots, \tilde{C}_r)) \leq d_{\text{TV}}(\mu_n, \mu'_n) \leq \left( \frac{nd_1}{d_2} \right)^2 \left( \frac{4ed_2}{n} \right)^{d_2}.$$

Then the conclusion follows from Theorem 2.10 and the triangle inequality.  $\square$

**6.2. Global laws.** The limiting spectral distributions for the adjacency matrix of a random regular hypergraph can be summarized in the following corollary.

**Corollary 6.11.** *Let  $H$  be a random  $(d_1, d_2)$ -regular hypergraph.*

(1) *If  $d_1, d_2$  are fixed, the empirical spectral distribution of  $\frac{A - (d_2 - 2)}{\sqrt{(d_1 - 1)(d_2 - 1)}}$  converges in probability to a measure  $\mu$  with density function given by*

$$(6.6) \quad f(x) := \frac{1 + \frac{d_2 - 1}{q}}{(1 + \frac{1}{q} - \frac{x}{\sqrt{q}})(1 + \frac{(d_2 - 1)^2}{q} + \frac{(d_2 - 1)x}{\sqrt{q}})} \frac{1}{\pi} \sqrt{1 - \frac{x^2}{4}} dx,$$

where  $q = (d_1 - 1)(d_2 - 1)$ .

(2) *For  $d_1, d_2 \rightarrow \infty$  with  $\frac{d_1}{d_2} \rightarrow \alpha \geq 1$  and  $d_1 \leq \frac{n}{32}$ , the empirical spectral distribution of  $\frac{A - (d_2 - 2)}{\sqrt{(d_1 - 1)(d_2 - 1)}}$  converges in probability to a measure supported on  $[-2, 2]$  with a density function given by*

$$(6.7) \quad g(x) = \frac{\alpha}{1 + \alpha + \sqrt{\alpha}x} \frac{1}{\pi} \sqrt{1 - \frac{x^2}{4}}.$$

(3) *If  $d_1 \rightarrow \infty$ ,  $d_1 = o(n^\varepsilon)$  for any  $\varepsilon > 0$  and  $\frac{d_1}{d_2} \rightarrow \infty$ , the ESD of  $\frac{A}{\sqrt{(d_1 - 1)(d_2 - 1)}}$  converges to the semicircle law in probability.*

*Proof of Corollary 6.11.* Claim (1) is proved in Theorem 6.4 of [32] based on a result for deterministic regular hypergraphs in Theorem 5 of [34].

Claim (2) is a combination of several results. When  $d_1 = o(n^{1/2})$ , it is proved in Theorem 6.6 of [32] based on the global law for random biregular bipartite graphs in [29] and [63]. When  $d_1 = \omega(\log^4 n)$ , the optimal local law for RBBGs was recently proved in [65], which also implies the global law for RBBGs. When  $d_1 \leq \frac{n}{32}$  and  $\frac{d_1}{d_2} \rightarrow \alpha$ , from Lemma 6.8,

$$\mathbb{P}(G \in \mathcal{G}'(n, m, d_1, d_2)) \rightarrow 1.$$

Therefore by the same proof of Theorem 6.6 in [32], the ESD of  $\frac{A - (d_2 - 2)}{\sqrt{(d_1 - 1)(d_2 - 1)}}$  for random regular hypergraphs converges in probability.

Under the assumptions  $d_1 \rightarrow \infty$ ,  $d_1 = o(n^\varepsilon)$  for any  $\varepsilon > 0$  and  $\frac{d_1}{d_2} \rightarrow \infty$ , from Lemma 6.8, we have again  $\mathbb{P}(G \in \mathcal{G}'(n, m, d_1, d_2)) \rightarrow 1$ . Then Claim (3) follows from Theorem 5.3 and Lemma 6.9.  $\square$

*Remark 6.12.* The ESD in Corollary 6.11 (2) is a shifted and scaled Marčenko-Pastur law. Taking  $\alpha \rightarrow \infty$ ,  $g(x)$  converges to the density function of the semicircle law. The transition from Marčenko-Pastur law to the semicircle law was also proved for sample covariance matrices in [7] when the aspect ratio goes to infinity.

*Remark 6.13.* A semicircle law for the adjacency matrix of  $d_2$ -uniform Erdős-Rényi random hypergraphs with growing expected degrees was proved in Theorem 5 of [46] when  $d_2$  is a constant. Part (3) of Corollary 6.11 proves a corresponding semicircle law for random  $d_2$ -uniform  $d_1$ -regular hypergraphs where  $d_2$  can be a parameter depending on  $n$ .

**6.3. Spectral gaps.** The spectral gap for random regular hypergraphs with fixed  $d_1, d_2$  was studied in [32]. Here we include the results for the case when  $d_1, d_2$  are growing with  $n$ .

**Corollary 6.14.** *Let  $H$  be a random  $(d_1, d_2)$ -regular hypergraph with  $d_1 \geq d_2$ . Let  $\lambda_1 \geq \dots \geq \lambda_n$  be the eigenvalues of  $A$ . Let  $\lambda = \max_{2 \leq i \leq n} |\lambda_i|$ .*

(1) *Suppose  $d_1 \geq d_2 \geq 3$  is fixed. There exists a sequence  $\varepsilon_n \rightarrow 0$  such that*

$$\mathbb{P}(|\lambda - (d_2 - 2)| \geq 2\sqrt{(d_1 - 1)(d_2 - 1)} + \varepsilon_n) \rightarrow 0$$

*as  $n \rightarrow \infty$ .*

(2) *Suppose  $3 \leq d_2 \leq \frac{1}{2}n^{2/3}$ ,  $d_1 \geq d_2 \geq cd_1$  for some  $c \in (0, 1)$ . Then for some constant  $K > 0$  depending on  $c$ , for all  $n \geq 1$ ,*

$$\mathbb{P}\left(\lambda \geq K\sqrt{(d_1 - 1)(d_2 - 1)}\right) = O\left(\frac{1}{n}\right).$$

(3) *Suppose  $3 \leq d_2 \leq C_1$  for a constant  $C_1$ , and  $d_1 = o(n^{1/2})$ . There exists a constant  $C$  depending on  $C_1$  such that*

$$\mathbb{P}\left(\lambda \geq C\sqrt{(d_1 - 1)(d_2 - 1)}\right) = O\left(\frac{d_1^2}{n^2}\right).$$

*Proof.* Claim (1) is proved in Theorem 4.3 in [32]. Claim (2) and (3) follow from part (2) and (3) in Theorem 3.1 with Lemma 6.8.  $\square$

*Remark 6.15.* Results in [66] that Claim (2) and (3) are based on have stronger probability estimates. However, Lemma 6.8 we used here yields a weaker failure probability.

**6.4. Eigenvalue fluctuations.** The following eigenvalue fluctuation results for random regular hypergraphs can be derived from Lemma 6.4, Lemma 6.9, and the eigenvalue fluctuations results for random biregular bipartite graphs in Section 4.

**Corollary 6.16.** *For fixed  $d_1 \geq d_2 \geq 3$ , let  $H$  be a random  $(d_1, d_2)$ -regular hypergraph with adjacency matrix  $A$ . Let  $\lambda_1 \geq \dots \geq \lambda_n$  be the eigenvalues of  $\frac{A}{\sqrt{(d_1 - 1)(d_2 - 1)}}$ . Suppose  $f$  is a function satisfying the same conditions in Theorem 4.4. Then  $Y_f^{(n)} := \sum_{i=1}^n f(\lambda_i) - na_0$  converges in distribution as  $n \rightarrow \infty$  to the infinitely divisible random variable*

$$Y_f := \sum_{k=2}^{\infty} \frac{a_k}{[(d_1 - 1)(d_2 - 1)]^{k/2}} \text{CNBW}_k^{(\infty)},$$

where  $\text{CNBW}_k^{(\infty)}$  is defined in (2.37).

*Proof.* Let  $\tilde{Y}_f^{(n)}$  be the corresponding random variable of  $Y_f^{(n)}$  for the uniform random biregular bipartite graphs considered in Theorem 4.4. From the total variation distance bound in Lemma

6.9, we have

$$d_{\text{TV}}(Y_f^{(n)}, \tilde{Y}_f^{(n)}) \leq d_{\text{TV}}(\mu_n, \mu'_n) \leq \left(\frac{nd_1}{d_2}\right)^2 \left(\frac{4ed_2}{n}\right)^{d_2} = o(1).$$

Therefore  $\tilde{Y}_f^{(n)}$  and  $Y_f^{(n)}$  converge in distribution to the same law.  $\square$

**Corollary 6.17.** *Let  $H$  be a random  $(d_1, d_2)$ -regular hypergraph with  $d_1 d_2 \rightarrow \infty$  as  $n \rightarrow \infty$  and  $d_1 d_2 = n^{o(1)}$ . Let  $\lambda_1 \geq \dots \geq \lambda_n$  be the eigenvalues of  $\frac{A}{\sqrt{(d_1-1)(d_2-1)}}$ . Let  $f$  be a function satisfying (4.20) and (4.22). Suppose one of the two assumptions holds:*

- (1) *there exists a constant  $c \geq 1$  such that  $1 \leq \frac{d_1}{d_2} \leq c$ ,*
- (2)  *$3 \leq d_2 \leq c_1$  for a constant  $c_1 \geq 3$ .*

*Then the random variable*

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

*converges in law to a Gaussian random variable with mean zero and variance  $\sigma_f = \sum_{k=2}^{\infty} 2ka_k^2$ . Moreover, for any fixed  $t$ , consider the entire functions  $g_1, \dots, g_t$  satisfying (4.20) and (4.22). The corresponding random vector  $(Y_{g_1}^{(n)}, \dots, Y_{g_t}^{(n)})$  converges in distribution to a centered Gaussian random vector  $(Z_{g_1}, \dots, Z_{g_t})$  with covariance*

$$\text{Cov}(Z_{g_i}, Z_{g_j}) = 2 \sum_{k=2}^{\infty} k a_k(g_i) a_k(g_j)$$

*for  $1 \leq i, j \leq t$ , where  $a_k(g_i), a_k(g_j)$  are the  $k$ -th coefficients in the expansion (4.19) for  $g_i, g_j$ , respectively.*

*Proof.* Recall Lemma 6.9 and our assumption  $d_1 d_2 = n^{o(1)}$ . Under Case (1), we have  $d_2 \rightarrow \infty$  and

$$d_{\text{TV}}(\mu_n, \mu'_n) \leq \left(\frac{nd_1}{d_2}\right)^2 \left(\frac{4ed_2}{n}\right)^{d_2} = O(n^2)(n^{(-1+o(1))d_2}) = o(1).$$

Under Case (2), we have

$$d_{\text{TV}}(\mu_n, \mu'_n) \leq \left(\frac{nd_1}{d_2}\right)^2 \left(\frac{4ed_2}{n}\right)^{d_2} = O(n^2 d_1^2) \left(\frac{4ec_1}{n}\right)^3 = o(1).$$

Then with Lemma 6.9, in both cases  $Y_f^{(n)}$  converges in distribution to the same limiting random variable defined in Theorem 4.7. The proof of the covariance part follows in the same way.  $\square$

## REFERENCES

- [1] Kartick Adhikari, Indrajit Jana, and Koushik Saha. Linear eigenvalue statistics of random matrices with a variance profile. *Random Matrices: Theory and Applications*, 10(03):2250004, 2021.
- [2] Greg W Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*, volume 118. Cambridge university press, 2010.
- [3] Greg W Anderson and Ofer Zeitouni. A CLT for a band matrix model. *Probability Theory and Related Fields*, 134(2):283–338, 2006.
- [4] Zhidong Bai and Jack W Silverstein. CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Annals of Probability*, pages 553–605, 2004.
- [5] Zhidong Bai, Xiaoying Wang, and Wang Zhou. Functional CLT for sample covariance matrices. *Bernoulli*, 16(4):1086–1113, 2010.
- [6] Zhidong Bai and Jianfeng Yao. On the convergence of the spectral empirical process of Wigner matrices. *Bernoulli*, 11(6):1059–1092, 2005.
- [7] Zhidong Bai and Yong Q Yin. Convergence to the semicircle law. *The Annals of Probability*, pages 863–875, 1988.

[8] Roland Bauerschmidt, Jiaoyang Huang, Antti Knowles, and Horng-Tzer Yau. Edge rigidity and universality of random regular graphs of intermediate degree. *Geometric and Functional Analysis*, pages 1–77, 2020.

[9] Gérard Ben Arous and Kim Dang. On fluctuations of eigenvalues of random permutation matrices. *Annales de l'IHP Probabilités et statistiques*, 51(2):620–647, 2015.

[10] Florent Benaych-Georges, Alice Guionnet, and Camille Male. Central limit theorems for linear statistics of heavy tailed random matrices. *Communications in Mathematical Physics*, 329(2):641–686, 2014.

[11] Ian F Blake and Shu Lin. On short cycle enumeration in biregular bipartite graphs. *IEEE Transactions on Information Theory*, 64(10):6526–6535, 2017.

[12] Vladimir Blinovsky and Catherine Greenhill. Asymptotic enumeration of sparse uniform linear hypergraphs with given degrees. *The Electronic Journal of Combinatorics*, 23(3):P3–17, 2016.

[13] Charles Bordenave. A new proof of friedman’s second eigenvalue theorem and its extension to random lifts. *Annales Scientifiques de l’École Normale Supérieure*, 4(6):1393–1439, 2020.

[14] Charles Bordenave and Marc Lelarge. Resolvent of large random graphs. *Random Structures & Algorithms*, 37(3):332–352, 2010.

[15] Charles Bordenave, Arnab Sen, and Bálint Virág. Mean quantum percolation. *Journal of the European Mathematical Society*, 19(12):3679–3707, 2017.

[16] Gerandy Brito, Ioana Dumitriu, and Kameron Decker Harris. Spectral gap in random bipartite biregular graphs and applications. *Combinatorics, Probability and Computing*, 31(2):229–267, 2022.

[17] Andrei Z Broder, Alan M Frieze, Stephen Suen, and Eli Upfal. Optimal construction of edge-disjoint paths in random graphs. *SIAM Journal on Computing*, 28(2):541–573, 1998.

[18] Shantanu Prasad Burnwal and Mathukumalli Vidyasagar. Deterministic completion of rectangular matrices using asymmetric ramanujan graphs: Exact and stable recovery. *IEEE Transactions on Signal Processing*, 68:3834–3848, 2020.

[19] Thierry Cabanal-Duvillard. Fluctuations de la loi empirique de grandes matrices aléatoires. *Annales de l’Institut Henri Poincaré (B) Probability and Statistics*, 37(3):373–402, 2001.

[20] E Rodney Canfield, Catherine Greenhill, and Brendan D McKay. Asymptotic enumeration of dense 0–1 matrices with specified line sums. *Journal of Combinatorial Theory, Series A*, 115(1):32–66, 2008.

[21] E Rodney Canfield and Brendan D McKay. Asymptotic enumeration of dense 0–1 matrices with equal row sums and equal column sums. *the Electronic Journal of Combinatorics*, 12(1):R29, 2005.

[22] Sourav Chatterjee. Fluctuations of eigenvalues and second order Poincaré inequalities. *Probability Theory and Related Fields*, 143(1–2):1–40, 2009.

[23] Sourav Chatterjee, Persi Diaconis, and Elizabeth Meckes. Exchangeable pairs and poisson approximation. *Probability Surveys*, 2:64–106, 2005.

[24] Binbin Chen and Guangming Pan. CLT for linear spectral statistics of normalized sample covariance matrices with the dimension much larger than the sample size. *Bernoulli*, 21(2):1089–1133, 2015.

[25] Nicholas Cook, Larry Goldstein, and Tobias Johnson. Size biased couplings and the spectral gap for random regular graphs. *The Annals of Probability*, 46(1):72–125, 2018.

[26] Ali Dehghan and Amir H Banihashemi. On the tanner graph cycle distribution of random LDPC, random protograph-based LDPC, and random quasi-cyclic LDPC code ensembles. *IEEE Transactions on Information Theory*, 64(6):4438–4451, 2018.

[27] Ali Dehghan and Amir H Banihashemi. On computing the multiplicity of cycles in bipartite graphs using the degree distribution and the spectrum of the graph. *IEEE Transactions on Information Theory*, 65(6):3778–3789, 2019.

[28] Amir Dembo, Eyal Lubetzky, and Yumeng Zhang. Empirical spectral distributions of sparse random graphs. In *In and Out of Equilibrium 3: Celebrating Vladas Sidoravicius*, pages 319–345. Springer, 2021.

[29] Ioana Dumitriu and Tobias Johnson. The Marčenko-Pastur law for sparse random bipartite biregular graphs. *Random Structures & Algorithms*, 48(2):313–340, 2016.

[30] Ioana Dumitriu, Tobias Johnson, Soumik Pal, and Elliot Paquette. Functional limit theorems for random regular graphs. *Probability Theory and Related Fields*, 156(3–4):921–975, 2013.

[31] Ioana Dumitriu and Soumik Pal. Sparse regular random graphs: spectral density and eigenvectors. *The Annals of Probability*, 40(5):2197–2235, 2012.

[32] Ioana Dumitriu and Yizhe Zhu. Spectra of random regular hypergraphs. *The Electronic Journal of Combinatorics*, pages P3–36, 2021.

[33] Nathanaël Enriquez and Laurent Ménard. Spectra of large diluted but bushy random graphs. *Random Structures & Algorithms*, 49(1):160–184, 2016.

[34] Keqin Feng and Wen-Ching Winnie Li. Spectra of hypergraphs and applications. *Journal of number theory*, 60(1):1–22, 1996.

[35] Joel Friedman. *A Proof of Alon's Second Eigenvalue Conjecture and Related Problems*. Memoirs of the American Mathematical Society. American Mathematical Society, 2008.

[36] David Gamarnik, Quan Li, and Hongyi Zhang. Matrix completion from  $O(n)$  samples in linear time. In *Conference on Learning Theory*, pages 940–947, 2017.

[37] Chris D Godsil and Bojan Mohar. Walk generating functions and spectral measures of infinite graphs. *Linear Algebra and its Applications*, 107:191–206, 1988.

[38] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.

[39] Kurt Johansson. On fluctuations of eigenvalues of random Hermitian matrices. *Duke mathematical journal*, 91(1):151–204, 1998.

[40] Tobias Johnson. Exchangeable pairs, switchings, and random regular graphs. *The Electronic Journal of Combinatorics*, 22(1):P1–33, 2015.

[41] Dag Jonsson. Some limit theorems for the eigenvalues of a sample covariance matrix. *Journal of Multivariate Analysis*, 12(1):1–38, 1982.

[42] Paul Jung and Jaehun Lee. Delocalization and limiting spectral distribution of Erdős-Rényi graphs with constant expected degree. *Electronic Communications in Probability*, 23, 2018.

[43] Alexei M Khorunzhy, Boris A Khoruzhenko, and Leonid A Pastur. Asymptotic properties of large random matrices with independent entries. *Journal of Mathematical Physics*, 37(10):5033–5060, 1996.

[44] Timothy Kusalik, James A Mingo, and Roland Speicher. Orthogonal polynomials and fluctuations of random matrices. *Journal für die reine und angewandte Mathematik*, 2007(604):1–46, 2007.

[45] Wen-Ching Winnie Li and Patrick Solé. Spectra of regular graphs and hypergraphs and orthogonal polynomials. *European Journal of Combinatorics*, 17(5):461–477, 1996.

[46] Linyuan Lu and Xing Peng. Loose Laplacian spectra of random hypergraphs. *Random Structures & Algorithms*, 41(4):521–545, 2012.

[47] Brendan D McKay. The expected eigenvalue distribution of a large regular graph. *Linear Algebra and its Applications*, 40:203–216, 1981.

[48] Brendan D McKay. Subgraphs of random graphs with specified degrees. In *Congressus Numerantium*, volume 33, pages 213–223, 1981.

[49] Brendan D McKay, Nicholas C Wormald, and Beata Wysocka. Short cycles in random regular graphs. *the electronic journal of combinatorics*, pages R66–R66, 2004.

[50] Fernando L Metz, Giorgio Parisi, and Luca Leuzzi. Finite-size corrections to the spectrum of regular random graphs: An analytical solution. *Physical Review E*, 90(5):052109, 2014.

[51] Fernando L Metz and Jeferson D Silva. Spectral density of dense random networks and the breakdown of the Wigner semicircle law. *Physical Review Research*, 2(4):043116, 2020.

[52] Hirobumi Mizuno and Iwao Sato. The semicircle law for semiregular bipartite graphs. *Journal of Combinatorial Theory, Series A*, 101(2):174–190, 2003.

[53] Nathan Noiry. Spectral asymptotic expansion of Wishart matrices with exploding moments. *ALEA*, 15(2):897–911, 2018.

[54] Guillem Perarnau and Giorgis Petridis. Matchings in random biregular bipartite graphs. *The Electronic Journal of Combinatorics*, pages P60–P60, 2013.

[55] Mariya Shcherbina and Brunello Tirozzi. Central limit theorem for fluctuations of linear eigenvalue statistics of large random graphs. *Journal of Mathematical Physics*, 51(2):023523, 2010.

[56] Mariya Shcherbina and Brunello Tirozzi. Central limit theorem for fluctuations of linear eigenvalue statistics of large random graphs: Diluted regime. *Journal of Mathematical Physics*, 53(4):043501, 2012.

[57] Galen R Shorack. *Probability for Statisticians*. Springer Texts in Statistics. Springer International Publishing, 2017.

[58] Ya Sinai and Alexander Soshnikov. Central limit theorem for traces of large random symmetric matrices with independent matrix elements. *Boletim da Sociedade Brasileira de Matemática-Bulletin/Brazilian Mathematical Society*, 29(1):1–24, 1998.

[59] Michael Sipser and Daniel A Spielman. Expander codes. *IEEE transactions on Information Theory*, 42(6):1710–1722, 1996.

[60] Philippe Sosoe and Percy Wong. Regularity conditions in the CLT for linear eigenvalue statistics of Wigner matrices. *Advances in Mathematics*, 249:37–87, 2013.

[61] R Tanner. A recursive approach to low complexity codes. *IEEE Transactions on information theory*, 27(5):533–547, 1981.

[62] Konstantin Tikhomirov and Pierre Youssef. The spectral gap of dense random regular graphs. *The Annals of Probability*, 47(1):362–419, 2019.

- [63] Linh V Tran. Local law for eigenvalues of random regular bipartite graphs. *Bulletin of the Malaysian Mathematical Sciences Society*, 43(2):1517–1526, 2020.
- [64] Lloyd N Trefethen. *Approximation Theory and Approximation Practice*, volume 128. SIAM, 2013.
- [65] Kevin Yang. Local Marchenko-Pastur law for random bipartite graphs. *arXiv preprint arXiv:1704.08672*, 2017.
- [66] Yizhe Zhu. On the second eigenvalue of random bipartite biregular graphs. *Journal of Theoretical Probability*, pages 1–35, 2022.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, SAN DIEGO, LA JOLLA, CA 92093  
*Email address:* `idumitriu@ucsd.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA IRVINE, IRVINE, CA 92697  
*Email address:* `yizhe.zhu@uci.edu`