

ARTICLE

# Sandwiching biregular random graphs

Tereza Klímošová<sup>1</sup> , Christian Reiher<sup>2</sup> , Andrzej Ruciński<sup>3</sup> and Matas Šileikis<sup>4\*</sup> 

<sup>1</sup>Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic, <sup>2</sup>Fachbereich Mathematik, Universität Hamburg, Hamburg, Germany, <sup>3</sup>Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań, Poland, and <sup>4</sup>Institute of Computer Science of the Czech Academy of Sciences, Prague, Czech Republic

\*Corresponding author. Email: [matas@cs.cas.cz](mailto:matas@cs.cas.cz)

(Received 30 January 2021; revised 1 October 2021; accepted 31 March 2022; first published online 6 June 2022)

## Abstract

Let  $\mathbb{G}(n_1, n_2, m)$  be a uniformly random  $m$ -edge subgraph of the complete bipartite graph  $K_{n_1, n_2}$  with bipartition  $(V_1, V_2)$ , where  $n_i = |V_i|$ ,  $i = 1, 2$ . Given a real number  $p \in [0, 1]$  such that  $d_1 := pn_2$  and  $d_2 := pn_1$  are integers, let  $\mathbb{R}(n_1, n_2, p)$  be a random subgraph of  $K_{n_1, n_2}$  with every vertex  $v \in V_i$  of degree  $d_i$ ,  $i = 1, 2$ . In this paper we determine sufficient conditions on  $n_1, n_2, p$  and  $m$  under which one can embed  $\mathbb{G}(n_1, n_2, m)$  into  $\mathbb{R}(n_1, n_2, p)$  and vice versa with probability tending to 1. In particular, in the balanced case  $n_1 = n_2$ , we show that if  $p \gg \log n/n$  and  $1 - p \gg (\log n/n)^{1/4}$ , then for some  $m \sim pn^2$ , asymptotically almost surely one can embed  $\mathbb{G}(n_1, n_2, m)$  into  $\mathbb{R}(n_1, n_2, p)$ , while for  $p \gg (\log^3 n/n)^{1/4}$  and  $1 - p \gg \log n/n$  the opposite embedding holds. As an extension, we confirm the Kim–Vu Sandwich Conjecture for degrees growing faster than  $(n \log n)^{3/4}$ .

**Keywords:** Random regular graph; Bipartite random graph; Embedding; Coupling; Monotone graph property

**2020 MSC Codes:** Primary: 05C80

## 1. Introduction

### 1.1. History and motivation

The Sandwich Conjecture of Kim and Vu [6] claims that if  $d \gg \log n$ , then for some sequences  $p_1 = p_1(n) \sim d/n$  and  $p_2 = p_2(n) \sim d/n$  there is a joint distribution of a random  $d$ -regular graph  $\mathbb{R}(n, d)$  and two binomial random graphs  $\mathbb{G}(n, p_1)$  and  $\mathbb{G}(n, p_2)$  such that with probability tending to 1

$$\mathbb{G}(n, p_1) \subseteq \mathbb{R}(n, d) \subseteq \mathbb{G}(n, p_2).$$

If true, the Sandwich Conjecture would essentially reduce the study of any monotone graph property of the random graph  $\mathbb{R}(n, d)$  in the regime  $d \gg \log n$  to the more manageable  $\mathbb{G}(n, p)$ .

For  $\log n \ll d \ll n^{1/3}(\log n)^{-2}$ , Kim and Vu proved the embedding  $\mathbb{G}(n, p_1) \subseteq \mathbb{R}(n, d)$  as well as an imperfect embedding  $\mathbb{R}(n, d) \setminus H \subseteq \mathbb{G}(n, p_2)$ , where  $H$  is some pretty sparse subgraph of  $\mathbb{R}(n, d)$ . In [2] the lower embedding was extended to  $d \ll n$  (and, in fact, to uniform hypergraph counterparts of the models  $\mathbb{G}(n, p)$  and  $\mathbb{R}(n, d)$ ). Recently, Gao, Isaev and McKay [4] came up with a result which confirms the conjecture for  $d \gg n/\sqrt{\log n}$  ([4] is the first paper that gives the

---

Research of TK was supported by the grant no. 19-04113Y of the Czech Science Foundation (GAČR) and the Center for Foundations of Modern Computer Science (Charles Univ. project UNCE/SCI/004).

Research of AR was supported by Narodowe Centrum Nauki, grant 2018/29/B/ST1/00426.

Research of MŠ was supported by the Czech Science Foundation, grant number 20-27757Y, with institutional support RVO:67985807.

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.



(perfect) embedding  $\mathbb{R}(n, d) \subseteq \mathbb{G}(n, p_2)$  for some range of  $d$ ) and, subsequently, Gao [3] widely extended this range to  $d = \Omega(\log^7 n)$ .

Initially motivated by a paper of Perarnau and Petridis [12] (see Section 9), we consider sandwiching for bipartite graphs, in which the natural counterparts of  $\mathbb{G}(n, p)$  and  $\mathbb{R}(n, d)$  are random subgraphs of the complete bipartite graph  $K_{n_1, n_2}$  rather than of  $K_n$ .

**1.2. New results**

We consider three models of random subgraphs of  $K_{n_1, n_2}$ , the complete bipartite graph with bipartition  $(V_1, V_2)$ , where  $|V_1| = n_1, |V_2| = n_2$ . Given an integer  $m \in [0, n_1 n_2]$ , let  $\mathbb{G}(n_1, n_2, m)$  be an  $m$ -edge subgraph of  $K_{n_1, n_2}$  chosen uniformly at random (the bipartite Erdős–Rényi model). Given a number  $p \in [0, 1]$ , let  $\mathbb{G}(n_1, n_2, p)$  be the binomial bipartite random graph where each edge of  $K_{n_1, n_2}$  is included independently with probability  $p$ . Note that in the latter model,  $pn_{3-i}$  is the expected degree of each vertex in  $V_i, i = 1, 2$ . If, in addition,

$$d_1 := pn_2 \quad \text{and} \quad d_2 := pn_1$$

are integers (we shall always make this implicit assumption), we let  $\mathcal{R}(n_1, n_2, p)$  be the class of subgraphs of  $K_{n_1, n_2}$  such that every  $v \in V_i$  has degree  $d_i$ , for  $i = 1, 2$  (it is an easy exercise to show that  $\mathcal{R}(n_1, n_2, p)$  is non-empty). We call such graphs *p-biregular*. Let  $\mathbb{R}(n_1, n_2, p)$  be a random graph chosen uniformly from  $\mathcal{R}(n_1, n_2, p)$ .

In this paper we establish an embedding of  $\mathbb{G}(n_1, n_2, m)$  into  $\mathbb{R}(n_1, n_2, p)$ . This easily implies an embedding of  $\mathbb{G}(n_1, n_2, p)$  into  $\mathbb{R}(n_1, n_2, p)$ . Moreover, by taking complements, our result translates immediately to the opposite embedding of  $\mathbb{R}(n_1, n_2, p)$  into  $\mathbb{G}(n_1, n_2, m)$  (and thus into  $\mathbb{G}(n_1, n_2, p)$ ). This idea was first used in [4] to prove  $\mathbb{R}(n, d) \subseteq \mathbb{G}(n, p)$  for  $p \gg \frac{1}{\sqrt{\log n}}$ . In particular, in the balanced case ( $n_1 = n_2 := n$ ), we prove this opposite embedding for  $p \gg (\log^3 n/n)^{1/4}$ .

The proof is far from a straightforward adaptation of the proof in [2]. The common aspect shared by the proofs is that the edges of  $\mathbb{R}(n_1, n_2, p)$  are revealed in a random order, giving a graph process which turns out to be, for most of the time, similar to the basic Erdős–Rényi process that generates  $\mathbb{G}(n_1, n_2, m)$ . The rest of the current proof is different in that it avoids using the configuration model. Instead, we focus on showing that both  $\mathbb{R}(n_1, n_2, p)$  and its random  $t$ -edge subgraphs are pseudorandom. We achieve this by applying the switching method (when  $\min\{p, 1 - p\}$  is small) and otherwise via asymptotic enumeration of bipartite graphs with a given degree sequence proved in [1] (see Theorem 5).

In addition, for  $p > 0.49$ , we rely on a non-probabilistic result about the existence of alternating cycles in 2-edge-coloured pseudorandom graphs (Lemma 20), which might be of separate interest.

Throughout the paper we assume that the underlying complete bipartite graph  $K_{n_1, n_2}$  grows on both sides, that is,  $\min\{n_1, n_2\} \rightarrow \infty$ , and any parameters (e.g.,  $p, m$ ), events, random variables, etc., are allowed to depend on  $(n_1, n_2)$ . In most cases we will make the dependence on  $(n_1, n_2)$  implicit, with all limits and asymptotic notation like  $O, \Omega, \sim$  considered with respect to  $\min\{n_1, n_2\} \rightarrow \infty$ . We say that an event  $\mathcal{E} = \mathcal{E}(n_1, n_2)$  holds *asymptotically almost surely (a.a.s.)* if  $\mathbb{P}(\mathcal{E}) \rightarrow 1$ .

Our main results are Theorem 2 below and its immediate Corollary 3. For a gentle start we first state an abridged version of both in the balanced case  $n_1 = n_2 = n$ . Note that  $pn^2$  is the number of edges in  $\mathbb{R}(n, n, p)$ .

**Theorem 1.** *If  $p \gg \frac{\log n}{n}$  and  $1 - p \gg \left(\frac{\log n}{n}\right)^{1/4}$ , then for some  $m \sim pn^2$  there is a joint distribution of random graphs  $\mathbb{G}(n, n, m)$  and  $\mathbb{R}(n, n, p)$  such that*

$$\mathbb{G}(n, n, m) \subseteq \mathbb{R}(n, n, p) \quad \text{a.a.s.} \tag{1}$$

If  $p \gg \left(\frac{\log^3 n}{n}\right)^{1/4}$ , then for some  $m \sim pn^2$  there is a joint distribution of random graphs  $\mathbb{G}(n, n, m)$  and  $\mathbb{R}(n, n, p)$  such that

$$\mathbb{R}(n, n, p) \subseteq \mathbb{G}(n, n, m) \quad \text{a.a.s.} \tag{2}$$

Moreover, in (1) and (2) one can replace  $\mathbb{G}(n, n, m)$  by the binomial random graph  $\mathbb{G}(n, n, p')$ , for some  $p' \sim p$ .

The condition  $p \gg (\log n)/n$  is necessary for (1) to hold with  $m \sim pn^2$  (see Remark 26), since otherwise the maximum degree of  $\mathbb{G}(n, n, m)$  is no longer  $pn(1 + o(1))$  a.a.s. We guess that the maximum degree is, vaguely speaking, the only obstacle for embedding  $\mathbb{G}(n, n, m)$  (or  $\mathbb{G}(n, n, p')$ ) into  $\mathbb{R}(n_1, n_2, p)$ , see Conjecture 1 in Section 10.

We further write  $N := n_1 n_2$ ,  $q := 1 - p$ ,  $\hat{n} := \min\{n_1, n_2\}$ ,  $\hat{p} := \min\{p, q\}$ , and let

$$\mathbb{I} := \mathbb{I}(n_1, n_2, p) = \begin{cases} 1, & \hat{p} < 2 \frac{n_1 n_2^{-1} + n_1^{-1} n_2}{\log N} \\ 0, & \hat{p} \geq 2 \frac{n_1 n_2^{-1} + n_1^{-1} n_2}{\log N}. \end{cases} \tag{3}$$

Note that  $\mathbb{I} = 0$  entails that the vertex classes are rather balanced: the ratio of their sizes cannot exceed  $\frac{1}{4} \log N$ . Moreover,  $\mathbb{I} = 0$  implies that  $\hat{p} \geq 4/\log N$ .

**Theorem 2.** For every constant  $C > 0$ , there is a constant  $C^*$  such that whenever the parameter  $p \in [0, 1]$  satisfies

$$q \geq 680 \left(\frac{3(C + 4) \log N}{\hat{n}}\right)^{1/4} \tag{4}$$

and

$$1 \geq \gamma := \begin{cases} C^* \left(p^2 \mathbb{I} + \sqrt{\frac{\log N}{p \hat{n}}}\right), & p \leq 0.49 \\ C^* \left(q^{3/2} \mathbb{I} + \left(\frac{\log N}{\hat{n}}\right)^{1/4} + \frac{1}{q} \sqrt{\frac{\log N}{\hat{n}}} \log \frac{\hat{n}}{\log N}\right), & p > 0.49, \end{cases} \tag{5}$$

there is, for  $m := \lceil (1 - \gamma)pN \rceil$ , a joint distribution of random graphs  $\mathbb{G}(n_1, n_2, m)$  and  $\mathbb{R}(n_1, n_2, p)$  such that

$$\mathbb{P}(\mathbb{G}(n_1, n_2, m) \subseteq \mathbb{R}(n_1, n_2, p)) = 1 - O(N^{-C}). \tag{6}$$

If, in addition  $\gamma \leq 1/2$ , then for  $p' := (1 - 2\gamma)p$ , there is a joint distribution of  $\mathbb{G}(n_1, n_2, p')$  and  $\mathbb{R}(n_1, n_2, p)$  such that

$$\mathbb{P}(\mathbb{G}(n_1, n_2, p') \subseteq \mathbb{R}(n_1, n_2, p)) = 1 - O(N^{-C}). \tag{7}$$

By inflating  $C^*$ , the constant 0.49 in (5) can be replaced by any constant smaller than 1/2. It can be shown (see Remark 27 in Section 10) that if  $p \leq 1/4$  and  $\gamma = \Theta\left(\sqrt{\frac{\log N}{p \hat{n}}}\right)$ , then  $\gamma$  has optimal order of magnitude. For more remarks about the conditions of Theorem 2 and the role of the indicator  $\mathbb{I}$ , see Section 10.

By taking the complements of  $\mathbb{R}(n_1, n_2, p)$  and  $\mathbb{G}(n_1, n_2, m)$  and swapping  $p$  and  $q$ , we immediately obtain the following consequence of Theorem 2 which provides the opposite embedding.

**Corollary 3.** For every constant  $C > 0$  there is a constant  $C^*$  such that whenever the parameter  $p \in [0, 1]$  satisfies

$$p \geq 680 (3(C + 4)\hat{n}^{-1} \log N)^{1/4} \tag{8}$$

and

$$1 \geq \bar{\gamma} := \begin{cases} C^* \left( p^{3/2} \mathbb{I} + \left( \frac{\log N}{\hat{n}} \right)^{1/4} + \frac{1}{p} \sqrt{\frac{\log N}{\hat{n}}} \log \frac{\hat{n}}{\log N} \right), & p < 0.51, \\ C^* \left( q^2 \mathbb{I} + \sqrt{\frac{\log N}{q\hat{n}}} \right), & p \geq 0.51, \end{cases}$$

there is, for  $\bar{m} = \lfloor (p + \bar{\gamma}q)N \rfloor$ , a joint distribution of random graphs  $\mathbb{G}(n_1, n_2, \bar{m})$  and  $\mathbb{R}(n_1, n_2, p)$  such that

$$\mathbb{P}(\mathbb{R}(n_1, n_2, p) \subseteq \mathbb{G}(n_1, n_2, \bar{m})) = 1 - O(N^{-C}). \tag{9}$$

If, in addition,  $\bar{\gamma} \leq 1/2$ , then for  $p'' := (p + 2\bar{\gamma}q)N$  there is a joint distribution of  $\mathbb{G}(n_1, n_2, p'')$  and  $\mathbb{R}(n_1, n_2, p)$  such that

$$\mathbb{P}(\mathbb{R}(n_1, n_2, p) \subseteq \mathbb{G}(n_1, n_2, p'')) = 1 - O(N^{-C}). \tag{10}$$

**Proof.** The assumptions of Corollary 3 yield the assumptions of Theorem 2 with  $\gamma = \bar{\gamma}$  and with  $p$  and  $q$  swapped. Note also that

$$m = \lceil (1 - \bar{\gamma})qN \rceil = \lceil N - pN - \bar{\gamma}qN \rceil = N - \lfloor (p + \bar{\gamma}q)N \rfloor = N - \bar{m}.$$

Thus, by Theorem 2, with probability  $1 - O(N^{-C})$  we have  $\mathbb{G}(n_1, n_2, N - \bar{m}) \subseteq \mathbb{R}(n_1, n_2, q)$ , which, by taking complements, translates into (9). Similarly

$$p' = (1 - 2\bar{\gamma})qN = N - (p + 2\bar{\gamma}q)N = N - p''N.$$

Thus, by Theorem 2, with probability  $1 - O(N^{-C})$  we have  $\mathbb{G}(n_1, n_2, N - p''N) \subseteq \mathbb{R}(n_1, n_2, q)$ , which, by taking complements yields embedding (10).  $\square$

**Proof of Theorem 1.** We apply Theorem 2 and Corollary 3 with  $C = 1$  and the corresponding  $C^*$ . Note that for  $n_1 = n_2 = n$ , the ratio in (3) equals  $4/\log N = 2/\log n$ , so

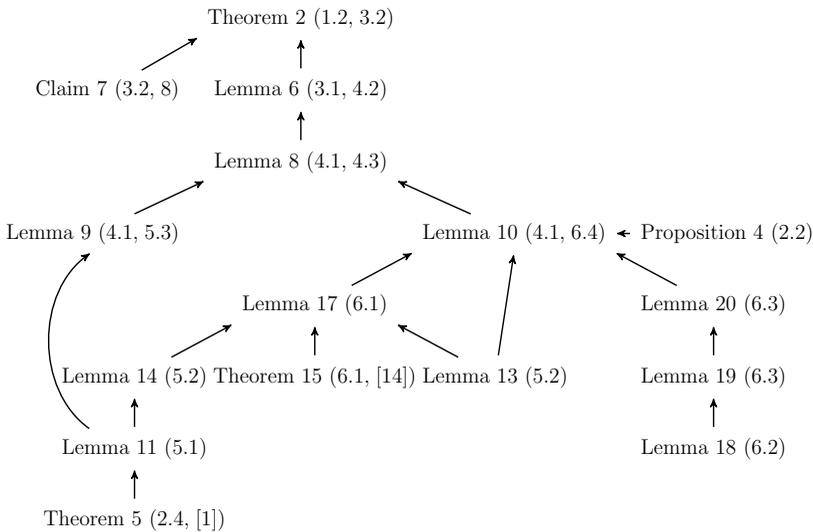
$$\mathbb{I} = 1 \quad \text{if and only if} \quad \hat{p} < 2/\log n. \tag{11}$$

To prove (1), assume  $p \gg \log n/n$  and  $q \gg (\log n/n)^{1/4}$  and apply Theorem 2. Note that condition (4) holds and it is straightforward to check that, regardless of whether  $p \leq 0.49$  or  $p > 0.49$ ,  $\gamma \rightarrow 0$ . In particular,  $\gamma \leq 1/2$ . We conclude that, indeed, embedding (1) holds with  $m = \lceil (1 - \gamma)pN \rceil \sim pn^2$  and (1) still holds if we replace  $\mathbb{G}(n, n, m)$  by  $\mathbb{G}(n, n, p')$  with  $p' = (1 - 2\gamma)p \sim p$ .

For (2) first note that when  $p \rightarrow 1$ , embedding (2) holds trivially with  $m = n^2$  (even though a nontrivial embedding follows in this case under an additional assumption  $q \gg \log n/n$ . Hence, we further assume  $q = \Omega(1)$ ,  $p \gg (\log^3 n/n)^{1/4}$  and apply Corollary 1.2. Note that condition (8) holds. It is routine to check, taking into account (11), that  $\bar{\gamma} \leq 1/2$  and, moreover,  $\bar{\gamma}q = o(p)$ . We conclude that (2) holds with  $m = \lfloor (p + \bar{\gamma}q)N \rfloor \sim Np = n^2p$  and (2) still holds if  $\mathbb{G}(n, n, m)$  is replaced by  $\mathbb{G}(n, n, p')$  with  $p' = p + 2\bar{\gamma}q \sim p$ .  $\square$

**1.3. A note on the second version of the manuscript.**

This project was initially aimed at extending the result in [2] to bipartite graphs and, thus, limited to the lower embedding  $\mathbb{G}(n, n, p_1) \subseteq \mathbb{R}(n, n, p)$  only. While it was in progress, Gao, Isaev and McKay [4] made an improvement on the Sandwich Conjecture by using a surprisingly fruitful idea of taking complements to obtain the upper embedding  $\mathbb{R}(n, d) \subseteq \mathbb{G}(n, p_2)$  directly from the lower embedding  $\mathbb{G}(n, p_1) \subseteq \mathbb{R}(n, d)$ . We then decided to borrow this idea (but nothing else) and strengthen some of our lemmas to get a significantly broader range of  $p$  for which the upper embedding (i.e. Corollary 1.2) holds. It turned out that our approach works for *non-bipartite* regular graphs, too. Therefore, prompted by the recent substantial progress of Gao [3] on the Sandwich Conjecture (which appeared on arXiv after the first version of this manuscript), in the current



**Figure 1.** The structure of the proof of Theorem 2. An arrow from statement A to statement B means that A is used in the proof of B. The numbers in the brackets point to the section where a statement is formulated and where it is proved (unless the proof follows the statement immediately); external results have instead an article reference in square brackets.

version of the manuscript we added Section 7, which outlines how to modify our proofs to get a corresponding sandwiching for non-bipartite graphs. This improves upon the results in [4] (for regular graphs), but is now superseded by [3].

**1.4. Organisation**

In Section 2 we introduce the notation and tools used throughout the paper: the switching technique, probabilistic inequalities and an enumeration result for bipartite graphs with a given degree sequence. In Section 3 we state a crucial Lemma 6 and show how it implies Theorem 2.

In Section 4 we give a proof of Lemma 6 based on two technical lemmas, one about the concentration of a degree-related parameter (Lemma 9), the other (Lemma 10) facilitating the switching technique used in the proof of Lemma 6. Lemma 9 is proved in Section 5, after giving some auxiliary results establishing the concentration of degrees and co-degrees in  $\mathbb{R}(n_1, n_2, p)$  as well as in its conditional versions. In Section 6, we present a proof of Lemma 10, preceded by a purely deterministic result about alternating cycles in 2-edge-coloured pseudorandom graphs. We defer some technical but straightforward results and their proofs (e.g., the proof of Claim 7) to Section 8. A flowchart of the results ultimately leading to the proof of Theorem 2 is presented in Figure 1.

The contents of Section 7 were already described (see Subsection 1.3 above). Section 9 contains an application of our main Theorem 2, which was part of the motivation for our research. In Section 10 we present some concluding remarks and our version of the (bipartite) sandwiching conjecture.

**2. Preliminaries**

**2.1. Notation**

Recall that

$$d_1 = pn_2 \quad \text{and} \quad d_2 = pn_1$$

are the degrees of vertices in a  $p$ -biregular graph, and thus, the number of edges in any  $p$ -biregular graph  $H \in \mathcal{R}(n_1, n_2, p)$  is

$$M := pN \quad \text{where} \quad N = n_1 n_2.$$

Throughout the proofs we also use shorthand notation

$$q = 1 - p, \quad \hat{p} = \min \{p, q\}, \quad \hat{n} = \min \{n_1, n_2\}, \tag{12}$$

and  $[n] := \{1, \dots, n\}$ . All logarithms appearing in this paper are natural.

By  $\Gamma_G(v)$  we denote the set of neighbours of a vertex  $v$  in a graph  $G$ .

### 2.2. Switchings

The switching technique is used to compare the size of two classes of graphs, say  $\mathcal{R}$  and  $\mathcal{R}'$ , by defining an auxiliary bipartite graph  $B := B(\mathcal{R}, \mathcal{R}')$ , in which two graphs  $H \in \mathcal{R}$ ,  $H' \in \mathcal{R}'$  are connected by an edge whenever  $H$  can be transformed into  $H'$  by some operation (a *forward switching*) that deletes and/or creates some edges of  $H$ . By counting the number of edges of  $B(\mathcal{R}, \mathcal{R}')$  in two ways, we see that

$$\sum_{H \in \mathcal{R}} \deg_B(H) = \sum_{H' \in \mathcal{R}'} \deg_B(H'), \tag{13}$$

which easily implies that

$$\frac{\min_{H' \in \mathcal{R}'} \deg_B(H')}{\max_{H \in \mathcal{R}} \deg_B(H)} \leq \frac{|\mathcal{R}|}{|\mathcal{R}'|} \leq \frac{\max_{H' \in \mathcal{R}'} \deg_B(H')}{\min_{H \in \mathcal{R}} \deg_B(H)}. \tag{14}$$

The reverse operation mapping  $H' \in \mathcal{R}'$  to its neighbours in graph  $B$ , is called a *backward switching*. Usually, one defines the forward switching in such a way that the backward switching can be easily described.

All switchings used in this paper follow the same pattern. For a fixed graph  $G \subseteq K$  (possibly empty), where  $K := K_{n_1, n_2}$ , the families  $\mathcal{R}, \mathcal{R}'$  will be subsets of

$$\mathcal{R}_G := \{H \in \mathcal{R}(n_1, n_2, p) : G \subseteq H\}. \tag{15}$$

Every  $H \in \mathcal{R}_G$  will be interpreted as a blue-red colouring of the edges of  $K \setminus G$ : those in  $H \setminus G$  will be coloured *blue* and those in  $K \setminus H - G$  *red*. Given  $H \in \mathcal{R}$ , consider a subset  $S$  of the edges of  $K \setminus G$  in which for every vertex  $v \in V_1 \cup V_2$  the blue degree equals the red degree, i.e.,  $\deg_{(H \setminus G) \cap S}(v) = \deg_{(K \setminus H) \cap S}(v)$ . Then switching the colours within  $S$  produces another graph  $H' \in \mathcal{R}_G$ . Formally,  $E(H')$  is the symmetric difference  $E(H) \Delta S$ . In each application of the switching technique, we will restrict the choices of  $S$  to make sure that  $H' \in \mathcal{R}'$ .

As an elementary illustration of this technique, which nevertheless turns out to be useful in Section 6, we prove here the following result. A cycle in  $K \setminus G$  is *alternating* if it is a union of a blue matching and a red matching. Note that the definition depends on  $H$ . We will omit mentioning this dependence, as  $H$  will always be clear from the context. Given  $e \in K \setminus G$ , set

$$\mathcal{R}_{G,e} := \{H \in \mathcal{R}_G : e \in H\} \quad \text{and} \quad \mathcal{R}_{G,-e} := \{H \in \mathcal{R}_G : e \notin H\}. \tag{16}$$

**Proposition 4.** *Let a graph  $G \subseteq K$  be such that  $\mathcal{R}_G \neq \emptyset$  and let  $e \in K \setminus G$ . Assume that for some number  $D > 0$  and every  $H \in \mathcal{R}_G$  the edge  $e$  is contained in an alternating cycle of length at most  $2D$ . Then  $\mathcal{R}_{G,-e} \neq \emptyset$ ,  $\mathcal{R}_{G,e} \neq \emptyset$ , and*

$$\frac{1}{N^D - 1} \leq \frac{|\mathcal{R}_{G,-e}|}{|\mathcal{R}_{G,e}|} \leq N^D - 1.$$

**Proof.** Let  $B = B(\mathcal{R}_{G,e}, \mathcal{R}_{G,-e})$  be the switching graph corresponding to the following forward switching: choose an alternating cycle  $S$  of length at most  $2D$  containing  $e$  and switch the colours

of edges within  $S$ . Note that the backward switching does precisely the same. Note that by the assumption, the minimum degree  $\delta(B)$  is at least 1.

Since  $\mathcal{R}_G = \mathcal{R}_{G,e} \cup \mathcal{R}_{G,-e} \neq \emptyset$ , one of the classes  $\mathcal{R}_{G,e}$  and  $\mathcal{R}_{G,-e}$  is non-empty and, in view of  $\delta(B) \geq 1$ , the other one is non-empty as well. For  $\ell = 2, \dots, D$ , the number of cycles of length  $2\ell$  containing  $e$  is, crudely, at most  $n_1^{\ell-1} n_2^{\ell-1} = N^{\ell-1}$ , hence the maximum degree is

$$\Delta(B) \leq N + N^2 + \dots + N^{D-1} \leq N^D - 1.$$

Thus, by (14), we obtain the claimed bounds. □

**2.3. Probabilistic inequalities**

We first state a few basic concentration inequalities that we apply in our proofs. For a sum  $X$  of independent Bernoulli (not necessarily identically distributed) random variables, writing  $\mu = \mathbb{E} X$ , we have (see Theorem 2.8, (2.5), (2.6), and (2.11) in [5]) that

$$\mathbb{P}(X \geq \mu + t) \leq \exp \left\{ -\frac{t^2}{2(\mu + t/3)} \right\}, \quad t \geq 0, \tag{17}$$

$$\mathbb{P}(X \leq \mu - t) \leq \exp \left\{ -\frac{t^2}{2\mu} \right\}, \quad t \geq 0. \tag{18}$$

Let  $\Gamma$  be a set of size  $|\Gamma| = g$  and let  $A \subseteq \Gamma$ ,  $|A| = a \geq 1$ . For an integer  $r \in [0, g]$ , choose uniformly at random a subset  $R \subseteq \Gamma$  of size  $|R| = r$ . The random variable  $Y = |A \cap R|$  has then the hypergeometric distribution  $\text{Hyp}(g, a, r)$  with expectation  $\mu := \mathbb{E} Y = ar/g$ . By Theorem 2.10 in [5], inequalities (17) and (18) hold for  $Y$ , too.

Moreover, by Remark 2.6 in [5], inequalities (17) and (18) also hold for a random variable  $Z$  which has Poisson distribution  $\text{Po}(\mu)$  with expectation  $\mu$ . In this case, we also have the following simple fact. For  $k \geq 0$ , set  $q_k = \mathbb{P}(Z = k)$ . Then  $q_k/q_{k-1} = \mu/k$ , and hence  $k = \lfloor \mu \rfloor$  maximises  $q_k$  (we say that such  $k$  is a *mode* of  $Z$ ). Since  $\text{Var} Z = \mu$ , by Chebyshev’s inequality,  $\mathbb{P}(|Z - \mu| < \sqrt{2\mu}) \geq 1/2$ . Moreover, the interval  $(\mu - \sqrt{2\mu}, \mu + \sqrt{2\mu})$  contains at most  $\lceil \sqrt{8\mu} \rceil$  integers, hence it follows that

$$q_{\lfloor \mu \rfloor} \geq \frac{1/2}{\lceil \sqrt{8\mu} \rceil}. \tag{19}$$

**2.4. Asymptotic enumeration of dense bipartite graphs**

To estimate co-degrees of  $\mathbb{R}(n, n, p)$  we will use the following asymptotic formula by Canfield, Greenhill and McKay [1]. We reformulate it slightly for our convenience.

Given two vectors  $\mathbf{d}_1 = (d_{1,v}, v \in V_1)$  and  $\mathbf{d}_2 = (d_{2,v}, v \in V_2)$  of positive integers such that  $\sum_{v \in V_1} d_{1,v} = \sum_{v \in V_2} d_{2,v}$ , let  $\mathcal{R}(\mathbf{d}_1, \mathbf{d}_2)$  be the class of bipartite graphs on  $(V_1, V_2)$  with vertex degrees  $\deg(v) = d_{i,v}, v \in V_i, i = 1, 2$ . Let  $|V_i| = n_i$  and write  $\bar{d}_i = n_i^{-1} \sum_{v \in V_i} d_{i,v}$ ,  $D_i = \sum_{v \in V_i} (d_{i,v} - \bar{d}_i)^2$ ,  $p = \bar{d}_1/n_2 = \bar{d}_2/n_1$ , and  $q = 1 - p$ .

**Theorem 5.** ([1]) *Given any positive constants  $a, b, C$  such that  $a + b < 1/2$ , there exists a constant  $\epsilon > 0$  so that the following holds. Consider the set of degree sequences  $\mathbf{d}_1, \mathbf{d}_2$  satisfying*

- i.  $\max_{v \in V_1} |d_{1,v} - \bar{d}_1| \leq Cn_2^{1/2+\epsilon}, \max_{v \in V_2} |d_{2,v} - \bar{d}_2| \leq Cn_1^{1/2+\epsilon}$
- ii.  $\max\{n_1, n_2\} \leq C(pq)^2(\min\{n_1, n_2\})^{1+\epsilon}$
- iii.  $\frac{(1-2p)^2}{4pq} \left(1 + \frac{5n_1}{6n_2} + \frac{5n_2}{6n_1}\right) \leq a \log \max\{n_1, n_2\}$ .

If  $\min\{n_1, n_2\} \rightarrow \infty$ , then uniformly for all such  $\mathbf{d}_1, \mathbf{d}_2$

$$|\mathcal{R}(\mathbf{d}_1, \mathbf{d}_2)| = \binom{n_1 n_2}{pn_1 n_2}^{-1} \prod_{v \in V_1} \binom{n_2}{d_{1,v}} \prod_{v \in V_2} \binom{n_1}{d_{2,v}} \times \exp \left[ -\frac{1}{2} \left( 1 - \frac{D_1}{pq n_1 n_2} \right) \left( 1 - \frac{D_2}{pq n_1 n_2} \right) + O \left( (\max\{n_1, n_2\})^{-b} \right) \right], \quad (20)$$

where the constants implicit in the error term may depend on  $a, b, C$ .

Note that condition (ii) of Theorem 5 implies the corresponding condition in [1] after adjusting  $\epsilon$ . Also, the uniformity of the bound is not explicitly stated in [1], but, given  $n_1, n_2$ , one should take  $\mathbf{d}_1, \mathbf{d}_2$  with the worst error and apply the result in [1].

### 3. A crucial lemma

#### 3.1. The set-up

Recall that  $K = K_{n_1, n_2}$  has  $N = n_1 n_2$  edges. Consider a sequence of graphs  $\mathbb{G}(t) \subseteq K$  for  $t = 0, \dots, N$ , where  $\mathbb{G}(0)$  is empty and, for  $t < N$ ,  $\mathbb{G}(t + 1)$  is obtained from  $\mathbb{G}(t)$  by adding an edge  $\epsilon_{t+1}$  chosen from  $K \setminus \mathbb{G}(t)$  uniformly at random, that is, for every graph  $G \subseteq K$  of size  $t$  and every edge  $e \in K \setminus G$

$$\mathbb{P}(\epsilon_{t+1} = e \mid \mathbb{G}(t) = G) = \frac{1}{N - t}. \quad (21)$$

Of course,  $(\epsilon_1, \dots, \epsilon_N)$  is just a uniformly random ordering of the edges of  $K$ .

Our approach to proving Theorem 2 is to represent the random regular graph  $\mathbb{R}(n_1, n_2, p)$  as the outcome of a random process which behaves similarly to  $(\mathbb{G}(t))_t$ . Recalling that a  $p$ -biregular graph has  $M = pN$  edges, let

$$(\eta_1, \dots, \eta_M)$$

be a uniformly random ordering of the edges of  $\mathbb{R}(n_1, n_2, p)$ . By taking the initial segments, we obtain a sequence of random graphs

$$\mathbb{R}(t) = \{\eta_1, \dots, \eta_t\}, \quad t = 0, \dots, M.$$

For convenience, we shorten

$$\mathbb{R} := \mathbb{R}(M) = \mathbb{R}(n_1, n_2, p).$$

Let us mention here that for a fixed  $H \in \mathcal{R}(n_1, n_2, p)$ , conditioning on  $\mathbb{R} = H$ , the edge set of the random subgraph  $\mathbb{R}(t)$  is a uniformly random  $t$ -element subset of the edge set of  $H$ . This observation often leads to a hypergeometric distribution and will be utilised several times in our proofs.

We say that a graph  $G$  with  $t$  edges is *admissible*, if the family  $\mathcal{R}_G$  (see definition (15)) is non-empty, or, equivalently,

$$\mathbb{P}(\mathbb{R}(t) = G) > 0.$$

For an admissible graph  $G$  with  $t$  edges and any edge  $e \in K \setminus G$ , let

$$p_{t+1}(e, G) := \mathbb{P}(\eta_{t+1} = e \mid \mathbb{R}(t) = G). \quad (22)$$

The conditional space underlying (22) can be described as first extending  $G$  uniformly at random to an element of  $\mathcal{R}_G$  and then randomly permuting the new  $M - t$  edges.

The main idea behind the proof of Theorem 2 is that the conditional probabilities in (22) behave similarly to those in (21). Observe that  $p_{t+1}(e, \mathbb{R}(t)) = \mathbb{P}(\eta_{t+1} = e \mid \mathbb{R}(t))$ . Given a real number  $\chi \geq 0$  and  $t \in \{0, \dots, M - 1\}$ , we define an  $\mathbb{R}(t)$ -measurable event

$$\mathcal{A}(t, \chi) := \left\{ p_{t+1}(e, \mathbb{R}(t)) \geq \frac{1 - \chi}{N - t} \text{ for every } e \in K \setminus \mathbb{R}(t) \right\}.$$

In the crucial lemma below, we are going to show that for suitably chosen  $\gamma_0, \gamma_1, \dots$ , a.a.s. the events  $\mathcal{A}(t, \gamma_t)$  occur simultaneously for all  $t = 0, \dots, t_0 - 1$  where  $t_0$  is quite close to  $M$ . Postponing the choice of  $\gamma_t$  and  $t_0$ , we define an event

$$\mathcal{A} := \bigcap_{t=0}^{t_0-1} \mathcal{A}(t, \gamma_t). \tag{23}$$

Intuitively, the event  $\mathcal{A}$  asserts that up to time  $t_0$  the process  $\mathbb{R}(t)$  stays ‘almost uniform’, which will enable us to embed  $\mathbb{G}(n_1, n_2, m)$  into  $\mathbb{R}(t_0)$ .

To define time  $t_0$ , it is convenient to parametrise the time by the proportion of edges of  $\mathbb{R}$  ‘not yet revealed’ after  $t$  steps. For this, we define by

$$\tau = \tau(t) := 1 - \frac{t}{M} \in [0, 1] \quad \text{and so} \quad t = (1 - \tau)M. \tag{24}$$

Given a constant  $C > 0$  and, we define (recalling the notation in (12)) the ‘final’ value  $\tau_0$  of  $\tau$  as

$$\tau_0 := \begin{cases} \frac{3 \cdot 3240^2(C + 4) \log N}{p\hat{n}}, & p \leq 0.49, \\ 700(3(C + 4))^{1/4} \left( q^{3/2}\mathbb{I} + \left( \frac{\log N}{\hat{n}} \right)^{1/4} \right), & p > 0.49. \end{cases} \tag{25}$$

(Some of the constants appearing here and below are sharp or almost sharp, but others have room to spare as we round them up to the nearest ‘nice’ number.)

Consider the following assumptions on  $p$  (which we will later show to follow from the assumptions of Theorem 2):

$$\hat{p} \geq \frac{3 \cdot 3240^2(C + 4) \log N}{\hat{n}}, \tag{26}$$

$$\hat{p} \cdot \mathbb{I} \leq \frac{49}{51} \cdot \frac{1}{340^2(C + 4)^{1/6}}, \tag{27}$$

and

$$q \geq 680 \left( \frac{3(C + 4) \log N}{\hat{n}} \right)^{1/4}. \tag{28}$$

At the end of this subsection we show that these three assumptions imply

$$\tau_0 \leq 1, \tag{29}$$

so that

$$t_0 := \lfloor (1 - \tau_0)M \rfloor$$

is a non-negative integer. Further, for  $t = 0, \dots, M - 1$ , define

$$\gamma_t := 1080\hat{p}^2\mathbb{I} + \begin{cases} 3240 \sqrt{\frac{2(C + 3) \log N}{\tau p \hat{n}}}, & p \leq 0.49, \\ 25,000 \sqrt{\frac{(C + 3) \log N}{\tau^2 q^2 \hat{n}}}, & p > 0.49. \end{cases} \tag{30}$$

Taking (29) for granted, we now state our crucial lemma, which is proved in Section 4.

**Lemma 6.** For every constant  $C > 0$ , if assumptions (26), (27) and (28) hold, then

$$\mathbb{P}(\mathcal{A}) = 1 - O(N^{-C}), \tag{31}$$

where the constant implicit in the  $O$ -term in (31) may also depend on  $C$ .

It remains to show (29). When  $p \leq 0.49$ , inequality (29) is equivalent to (26). For  $p > 0.49$ , we have  $q \leq 51\hat{p}/49$ , which together with assumptions (27) and (28) implies that

$$\tau_0 \leq 700 (3(C + 4))^{1/4} \left(\frac{51}{49}\hat{p} \cdot \mathbb{I}\right)^{3/2} + 700 \cdot \frac{q}{680} \leq \frac{700 \cdot 3^{1/4}}{340^3} + \frac{700 \cdot 0.51}{680} < 1.$$

**3.2. Proof of Theorem 2**

From Lemma 6 we are going to deduce Theorem 2 using a coupling argument similar to the one which was employed by Dudek, Frieze, Ruciński and Šileikis [2], but with an extra tweak (inspired by Kim and Vu [6]) of letting the probabilities  $\gamma_t$  of Bernoulli random variables depend on  $t$ , which reduces the error  $\gamma$  in (5).

It is easy to check that the assumptions of Lemma 6 follow from the assumptions of Theorem 2. Indeed, (28) coincides with assumption (4), while (26) and (27) follow from the assumption  $\gamma \leq 1$  (see (5)) for sufficiently large  $C^*$ .

Our aim is to couple  $(\mathbb{G}(t))_t$  and  $(\mathbb{R}(t))_t$  on the event  $\mathcal{A}$  defined in (23). For this we will define a graph process  $\mathbb{R}'(t) := \{\eta'_1, \dots, \eta'_t\}$ ,  $t = 0, \dots, t_0$  so that for every admissible graph  $G$  with  $t \in [0, M - 1]$  edges and every  $e \in K \setminus G$

$$\mathbb{P}(\eta'_{t+1} = e \mid \mathbb{R}'(t) = G) = p_{t+1}(e, G), \tag{32}$$

where  $p_{t+1}(e, G)$  was defined in (22). Note that  $\mathbb{R}'(0)$  is an empty graph. Since the distribution of the process  $(\mathbb{R}(t))_t$  is determined by the conditional probabilities (22), in view of (32), the distribution of  $\mathbb{R}'(t_0)$  is the same as that of  $\mathbb{R}(t_0)$  and therefore we will identify  $\mathbb{R}'(t_0)$  with  $\mathbb{R}(t_0)$ . As the second step, we will show that a.a.s.  $\mathbb{G}(n_1, n_2, m)$  can be sampled from  $\mathbb{R}'(t_0) = \mathbb{R}(t_0)$ .

Proceeding with the definition, set  $\mathbb{R}'(0)$  to be the empty graph and define graphs  $\mathbb{R}'(t)$ ,  $t = 1, \dots, t_0$ , inductively, as follows. Hence, let us further fix  $t \in [0, t_0 - 1]$  and suppose that

$$\mathbb{R}'(t) = R_t \quad \text{and} \quad \mathbb{G}(t) = G_t$$

have been already chosen. Our immediate goal is to select a random pair of edges  $\varepsilon_{t+1}$  and  $\eta'_{t+1}$ , according to, resp., (21) and (32), in such a way that the event  $\varepsilon_{t+1} \in \mathbb{R}'(t + 1)$  is quite likely.

To this end, draw  $\varepsilon_{t+1}$  uniformly at random from  $K \setminus G_t$  and, independently, generate a Bernoulli random variable  $\xi_{t+1}$  with the probability of success  $1 - \gamma_t$  (which is in  $[0, 1]$  by (146)). If event  $\mathcal{A}(t, \gamma_t)$  has occurred, that is, if

$$p_{t+1}(e, R_t) \geq \frac{1 - \gamma_t}{N - t} \quad \text{for every} \quad e \in K \setminus R_t, \tag{33}$$

then draw a random edge  $\zeta_{t+1} \in K \setminus R_t$  according to the distribution

$$\mathbb{P}(\zeta_{t+1} = e \mid \mathbb{R}'(t) = R_t) := \frac{p_{t+1}(e, R_t) - (1 - \gamma_t)/(N - t)}{\gamma_t} \geq 0,$$

where the inequality holds by (33). Observe also that

$$\sum_{e \in K \setminus R_t} \mathbb{P}(\zeta_{t+1} = e \mid \mathbb{R}'(t) = R_t) = 1,$$

so  $\zeta_{t+1}$  has a properly defined distribution. Finally, fix an arbitrary bijection

$$f_{R_t, G_t} : R_t \setminus G_t \rightarrow G_t \setminus R_t$$

between the sets of edges and define

$$\eta'_{t+1} = \begin{cases} \varepsilon_{t+1}, & \text{if } \xi_{t+1} = 1, \varepsilon_{t+1} \in K \setminus R_t, \\ f_{R_t, G_t}(\varepsilon_{t+1}), & \text{if } \xi_{t+1} = 1, \varepsilon_{t+1} \in R_t, \\ \zeta_{t+1}, & \text{if } \xi_{t+1} = 0. \end{cases}$$

On the other hand, if event  $\mathcal{A}(t, \gamma_t)$  has failed, then  $\eta'_{t+1}$  is sampled directly (without defining  $\zeta_{t+1}$ ) according to the distribution (32). With this definition of  $(\mathbb{R}'(t))_{t=0}^{t_0}$ , it is easy to check that for  $\eta'_{t+1}$  defined above, (32) indeed holds, so from now on we drop the prime ' and identify  $\mathbb{R}'(t)$  with  $\mathbb{R}(t)$ , which is a subset of  $\mathbb{R}(n_1, n_2, p)$ .

Most importantly, we conclude that, for  $t = 0, \dots, t_0 - 1$

$$\mathcal{A}(t, \gamma_t) \cap \{\xi_{t+1} = 1\} \implies \varepsilon_{t+1} \in \mathbb{R}(t + 1). \tag{34}$$

In view of this, define

$$S := \{t \in [t_0] : \xi_t = 1\}$$

and recall that  $m = \lceil (1 - \gamma)M \rceil$ . If  $|S| \geq m$ , define  $\mathbb{G}(n_1, n_2, m)$  as, say, the edges indexed by the smallest  $m$  elements of  $S$  (note that since the vectors  $(\xi_i)$  and  $(\varepsilon_i)$  are independent, after conditioning on  $S$ , these  $m$  edges are uniformly distributed), and if  $|S| < m$ , then define  $\mathbb{G}(n_1, n_2, m)$  as, say, the graph with edges  $\{\varepsilon_1, \dots, \varepsilon_m\}$ . Recalling the definition (23) of the event  $\mathcal{A}$ , by (34) we observe that  $\mathcal{A}$  implies the inclusion  $\{\varepsilon_t : t \in S\} \subseteq \mathbb{R}(t_0) \subset \mathbb{R}(M) = \mathbb{R}(n_1, n_2, p)$ . On the other hand  $|S| \geq m$  implies  $\mathbb{G}(n_1, n_2, m) \subseteq \{\varepsilon_t : t \in S\}$ , so

$$\mathbb{P}(\mathbb{G}(n_1, n_2, m) \subseteq \mathbb{R}) \geq \mathbb{P}(\{|S| \geq m\} \cap \mathcal{A}).$$

Since, by Lemma 6, event  $\mathcal{A}$  holds with probability  $1 - O(N^{-C})$ , to complete the proof of (6) it suffices to show that also

$$\mathbb{P}(|S| \geq m) = 1 - O(N^{-C}).$$

For this we need the following claim whose technical proof is deferred to Section 8.

**Claim 7.** *We have*

$$\mathbb{E} |S| \geq t_0 - \theta M, \tag{35}$$

where

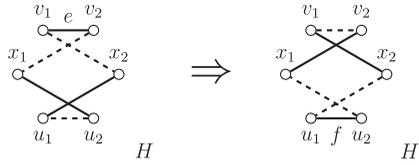
$$\theta := 1080\hat{p}^2\mathbb{I} + \begin{cases} 6480\sqrt{\frac{2(C+3)\log N}{p\hat{n}}}, & p \leq 0.49, \\ 6250\sqrt{\frac{(C+3)\log N}{q^2\hat{n}}} \log \frac{\hat{n}}{\log N}, & p > 0.49, \end{cases}$$

and, with  $\gamma$  as in (5),

$$\gamma \geq \tau_0 + \theta + 2/M + \sqrt{\frac{2C \log N}{M}}. \tag{36}$$

Recalling that  $t_0 = \lfloor (1 - \tau_0)M \rfloor$  and  $m = \lceil (1 - \gamma)M \rceil$ , we have

$$\begin{aligned} t_0 - \theta M - m &\geq (1 - \tau_0)M - 1 - \theta M - (1 - \gamma)M - 1 \\ &= (\gamma - \tau_0 - \theta)M - 2 \stackrel{(36)}{\geq} \sqrt{2CM \log N}. \end{aligned} \tag{37}$$



**Figure 2.** Switching between  $H$  and  $H'$  when  $e$  and  $f$  are disjoint: solid edges are in  $H \setminus G$  (or  $H' \setminus G$ ) and the dashed ones in  $K \setminus H$  (or  $K \setminus H'$ ).

Since  $|S|$  is a sum of independent Bernoulli random variables,

$$\begin{aligned} \mathbb{P}(|S| < m) &= \mathbb{P}(|S| < \mathbb{E}|S| - (\mathbb{E}|S| - m)) \stackrel{(35)}{\leq} \mathbb{P}(|S| < \mathbb{E}|S| - (t_0 - \theta M - m)) \\ \stackrel{(18), \mathbb{E}|S| \leq M}{\leq} \exp\left(-\frac{(t_0 - \theta M - m)^2}{2M}\right) &\stackrel{(38)}{\leq} \exp(-C \log N) = N^{-C}, \end{aligned}$$

which as mentioned above, implies (6).

Finally, we prove (7) by coupling  $\mathbb{G}(n_1, n_2, p')$  with  $\mathbb{G}(n_1, n_2, m)$  so that the former is a subset of the latter with probability  $1 - O(N^{-C})$ . Denote by  $X := e(\mathbb{G}(n_1, n_2, p'))$  the number of edges of  $\mathbb{G}(n_1, n_2, p')$ . Whenever  $X \leq m$ , choose  $X$  edges of  $\mathbb{G}(n_1, n_2, m)$  at random and declare them to be a copy of  $\mathbb{G}(n_1, n_2, p')$ . Otherwise sample  $\mathbb{G}(n_1, n_2, p')$  independently from  $\mathbb{G}(n_1, n_2, m)$ . Hence it remains to show that  $\mathbb{P}(X > m) = O(N^{-C})$ .

Recalling that  $m = \lceil(1 - \gamma)\rceil pN$  and  $p' = (1 - 2\gamma)p$ , we have  $m/N \geq (1 - \gamma)p = p' + \gamma p$ . Since  $X \sim \text{Bin}(N, p')$ , Chernoff's bound (17) implies that

$$\mathbb{P}(X > m) \leq \mathbb{P}(X > p'N + \gamma pN) \leq \exp\left\{-\frac{(\gamma pN)^2}{2(p'N + \gamma pN/3)}\right\} \leq e^{-\gamma^2 pN/2}. \tag{38}$$

Choosing  $C^*$  sufficiently large, the definition (5) implies  $\gamma \geq \sqrt{\frac{2C \log N}{pN}}$  with a lot of room (to see this easier in the case  $p > 0.49$ , note that (4) implies  $\log(\hat{n}/\log N) \geq 1$ , say). So (38) implies  $\mathbb{P}(X > m) \leq N^{-C}$ .

**4. Proof of Lemma 6**

**4.1. Preparations**

Recall our notation  $K = K_{n_1, n_2}$ , and  $\mathcal{R}_G$  defined in (15). Fix an integer  $t \in [0, t_0)$ . For a graph  $G$  with  $t$  edges and  $e, f \in K \setminus G$ , define

$$\mathcal{R}_{G, e, \neg f} := \{H \in \mathcal{R}_G : e \in H, f \notin H\}.$$

By skipping a few technical steps, we will see that the essence of the proof of Lemma 6 is to show that the ratio  $|\mathcal{R}_{G, e, \neg f}|/|\mathcal{R}_{G, f, \neg e}|$  is approximately 1 for any pair of edges  $e, f$ , where  $G$  is a ‘typical’ instance of  $\mathbb{R}(t)$ .

Recalling our generic definition of switchings from Subsection 2.2, let us treat the graphs in  $\mathcal{R} := \mathcal{R}_{G, e, \neg f}$  and  $\mathcal{R}' := \mathcal{R}_{G, f, \neg e}$  as blue-red edge colourings of the graph  $K \setminus G$ . Recall that a path or a cycle in  $K \setminus G$  is *alternating* if no two consecutive edges have the same colour.

When edges  $e, f$  are disjoint, we define the switching graph  $B = B(\mathcal{R}, \mathcal{R}')$  by putting an edge between  $H \in \mathcal{R}$  and  $H' \in \mathcal{R}'$  whenever there is an alternating 6-cycle containing  $e$  and  $f$  in  $H$  such that switching the colours in the cycle gives  $H'$  (see Figure 2). If  $e$  and  $f$  share a vertex, instead of 6-cycles we use alternating 4-cycles containing  $e$  and  $f$ .

It is easy to describe the vertex degrees in  $B$ . For distinct  $u, v \in V_i$ , let  $\theta_{G, H}(u, v)$  be the number of (alternating) paths  $uxv$  such that  $ux$  is blue and  $xv$  is red. Note that  $\theta_{G, H}(u, v) = |\Gamma_{H \setminus G}(u) \cap$

$\Gamma_{K \setminus H}(v)$ . Then, setting  $f = u_1 u_2$  and  $e = v_1 v_2$ , where  $u_i, v_i \in V_i, i = 1, 2$ ,

$$\deg_B(H) = \prod_{i:u_i \neq v_i} \theta_{G,H}(u_i, v_i), \quad H \in \mathcal{R} \tag{39}$$

and

$$\deg_B(H') = \prod_{i:u_i \neq v_i} \theta_{G,H'}(v_i, u_i), \quad H' \in \mathcal{R}'. \tag{40}$$

Note that equation (13) is equivalent to

$$\frac{|\mathcal{R}|}{|\mathcal{R}'|} = \frac{\frac{1}{|\mathcal{R}'|} \sum_{H' \in \mathcal{R}'} \deg_B(H')}{\frac{1}{|\mathcal{R}|} \sum_{H \in \mathcal{R}} \deg_B(H)}. \tag{41}$$

In view of (39) and (40), the denominator of the RHS above is the (conditional) expectation of the random variable  $\prod_{i:u_i \neq v_i} \theta_{G,\mathbb{R}}(u_i, v_i)$ , given that  $\mathbb{R}$  contains  $G$  and  $e$ , but not  $f$  (and similarly for the numerator).

To get an idea of how large that expectation could be, let us focus on one factor, say,  $\theta_{G,\mathbb{R}}(u_1, v_1)$ , assuming  $u_1 \neq v_1$ . Clearly, the red degree of  $v_1$  equals  $|\Gamma_{K \setminus \mathbb{R}}(v_1)| = n_2 - d_2 = qn_2$ . Since  $|\Gamma_{\mathbb{R}}(v_1)| = pn_2$ , viewing  $\mathbb{R} \setminus G$  as a  $\tau$ -dense subgraph of  $\mathbb{R}$  (see (24)) we expect that the blue neighbourhood  $|\Gamma_{\mathbb{R} \setminus G}(u_1)|$  is approximately  $\tau pn_2$ . It is reasonable to expect that for typical  $G$  and  $\mathbb{R}$  the red and blue neighbourhoods intersect proportionally, that is, on a set of size about  $q \cdot \tau p \cdot n_2 = \tau qd_1$ .

Inspired by this heuristic, we say that, for  $\delta > 0$ , an admissible graph  $G$  with  $t$  edges is  $\delta$ -typical if

$$\max_{u_1 u_2, v_1 v_2 \in K \setminus G} \mathbb{P} \left( \max_{i \in [2]: u_i \neq v_i} \left| \frac{\theta_{G,\mathbb{R}}(u_i, v_i)}{\tau qd_i} - 1 \right| > \delta \mid G \subseteq \mathbb{R}, v_1 v_2 \in \mathbb{R}, u_1 u_2 \notin \mathbb{R} \right) \leq \tau^2 \delta, \tag{42}$$

where the outer maximum is taken over distinct pairs of edges. (The bound  $\tau^2 \delta$  has hardly any intuition, but it is simple and sufficient for our purposes.)

Lemma 6 is a relatively easy consequence of the upcoming Lemma 8, which states that for a suitably chosen function  $\delta(t)$  it is very likely that the initial segments of  $\mathbb{R}$  are  $\delta(t)$ -typical.

For each  $t = 0, \dots, t_0 - 1$ , define

$$\delta(t) := 120\hat{p}^2\mathbb{I} + 360\sqrt{\frac{(C+3)\lambda(t)}{6\tau p q \hat{n}}}, \tag{43}$$

where

$$\lambda(t) := \begin{cases} 6 \log N, & p \leq 0.49, \\ 6 \log N + \frac{64 \log N}{\tau p q}, & p > 0.49. \end{cases} \tag{44}$$

For future reference, note that

$$\delta(t) \leq \gamma_t/9. \tag{45}$$

Indeed, recalling (30), for  $p \leq 0.49$ ,

$$9\delta(t) \leq 1080\hat{p}^2\mathbb{I} + 3240\sqrt{\frac{2(C+3)\log N}{\tau p \hat{n}}} = \gamma_t,$$

while, for  $p > 0.49$ , noting that  $\lambda(t) \leq 70 \log N / (\tau pq)$ , we have

$$9\delta(t) \leq 1080\hat{p}^2\mathbb{I} + 3240\sqrt{\frac{70(C+3)\log N}{0.49^2 \cdot 6\tau^2q^2\hat{n}}} \leq \gamma_t.$$

**Lemma 8.** *For every constant  $C > 0$ , under the conditions of Lemma 6,*

$$\mathbb{P}(\mathbb{R}(t) \text{ is } \delta(t)\text{-typical for all } t < t_0) = 1 - O(N^{-C}).$$

Note that the LHS of (42) is a function of graph  $G$ , say  $f(G)$ . Hence, Lemma 8 asserts that, with high probability,  $f(\mathbb{R}(t))$  is small for all  $t < t_0$ . The main idea of the proof of Lemma 8, which we defer to Subsection 4.3, is to bound  $f(\mathbb{R}(t))$  by a ratio of two simpler functions (again conditional probabilities) of  $\mathbb{R}(t)$ . Lemmas 9 and 10 below bound each of these conditional probabilities separately.

For any  $t = 0, \dots, M$  and distinct  $u_i, v_i \in V_i, i = 1, 2$ , set

$$\theta_t(u_i, v_i) := \theta_{\mathbb{R}(t), \mathbb{R}}(u_i, v_i), \tag{46}$$

where  $\theta_{G,H}(u_i, v_i) = |\Gamma_{H \setminus G}(u_i) \cap \Gamma_{K \setminus H}(v_i)|$  was introduced earlier in this subsection.

**Lemma 9.** *For every constant  $C > 0$ , under the conditions of Lemma 6, for  $i \in [2]$  we have that with probability  $1 - O(N^{-C})$ ,*

$$\mathbb{P}\left(\max_{u,v \in V_i, u \neq v} \left| \frac{\theta_t(u, v)}{\tau q d_i} - 1 \right| > \delta(t) \mid \mathbb{R}(t)\right) \leq e^{-2\lambda(t)}, \quad \text{for all } t < t_0. \tag{47}$$

**Lemma 10.** *For every constant  $C > 0$ , under the conditions of Lemma 6, with probability  $1 - O(N^{-C})$*

$$\min_{e,f \in K \setminus \mathbb{R}(t), e \neq f} \mathbb{P}(e \in \mathbb{R}, f \notin \mathbb{R} \mid \mathbb{R}(t)) \geq e^{-\lambda(t)}, \quad \text{for all } t < t_0. \tag{48}$$

**4.2. Proof of Lemma 6**

In view of Lemma 8, it suffices to show that if  $t < t_0$ , and a  $t$ -edge graph  $G$  is  $\delta(t)$ -typical, then

$$\min_{e \in K \setminus G} p_{t+1}(e, G) \geq \frac{1 - \gamma_t}{N - t}.$$

Fix  $f \in K \setminus G$  which maximises  $p_{t+1}(f, G)$ . Since the average of  $p_{t+1}(e, G)$  over  $e \in K \setminus G$  is exactly  $\frac{1}{N-t}$ , we have  $p_{t+1}(f, G) \geq \frac{1}{N-t}$  and therefore it is enough to prove that, for every  $e \in K \setminus (G \cup \{f\})$ ,

$$\frac{p_{t+1}(e, G)}{p_{t+1}(f, G)} \geq 1 - \gamma_t.$$

Recall the definitions of  $\mathcal{R}_G$  in (15),  $\mathcal{R}_{G,e}, \mathcal{R}_{G,-e}$  in (16), and  $\mathcal{R} = \mathcal{R}_{G,e,-f}$  and  $\mathcal{R}' = \mathcal{R}_{G,f,-e}$  from Subsection 4.1 and observe that  $\mathcal{R} = \mathcal{R}_{G \cup \{e\}, -f}$  too. In view of the remark immediately following (22),

$$p_{t+1}(e, G) = \mathbb{P}(e \in \mathbb{R} \mid \mathbb{R}(t) = G) \cdot \mathbb{P}(\eta_{t+1} = e \mid \mathbb{R}(t) = G, e \in \mathbb{R}) = \frac{|\mathcal{R}_{G,e}|}{|\mathcal{R}_G|} \cdot \frac{1}{M - t}.$$

Therefore,

$$1 \geq \frac{p_{t+1}(e, G)}{p_{t+1}(f, G)} = \frac{|\mathcal{R}_{G,e}|}{|\mathcal{R}_{G,f}|} = \frac{|\mathcal{R}_{G \cup \{e,f\}}| + |\mathcal{R}_{G,e,-f}|}{|\mathcal{R}_{G \cup \{e,f\}}| + |\mathcal{R}_{G,f,-e}|} \geq \frac{|\mathcal{R}|}{|\mathcal{R}'|}.$$

Write  $e = v_1 v_2$  and  $f = u_1 u_2$  and for simplicity assume that both  $u_1 \neq v_1$  and  $u_2 \neq v_2$  (otherwise the proof goes mutatis mutandis and is, in fact, a bit simpler). By (39), (40), and (41),

$$\frac{|\mathcal{R}|}{|\mathcal{R}'|} = \frac{\mathbb{E}_{\text{top}}}{\mathbb{E}_{\text{bottom}}},$$

where

$$\mathbb{E}_{\text{top}} := \frac{1}{|\mathcal{R}'|} \sum_{H' \in \mathcal{R}} \text{deg}_B(H') = \mathbb{E} [\theta_{G, \mathbb{R}}(v_1, u_1) \theta_{G, \mathbb{R}}(v_2, u_2) \mid G \subseteq \mathbb{R}, f \in \mathbb{R}, e \notin \mathbb{R}]$$

and

$$\mathbb{E}_{\text{bottom}} := \frac{1}{|\mathcal{R}|} \sum_{H \in \mathcal{R}} \text{deg}_B(H) = \mathbb{E} [\theta_{G, \mathbb{R}}(u_1, v_1) \theta_{G, \mathbb{R}}(u_2, v_2) \mid G \subseteq \mathbb{R}, e \in \mathbb{R}, f \notin \mathbb{R}].$$

Since  $G$  is  $\delta$ -typical with  $\delta = \delta(t)$ , denoting the LHS of (42) as  $p_*$ , we have  $p_* \leq \tau^2 \delta$ , so

$$\begin{aligned} \mathbb{E}_{\text{top}} &\geq (1 - \delta) \tau q d_1 \cdot (1 - \delta) \tau q d_2 \cdot (1 - p_*) + 0 \cdot p_* \\ &\geq (1 - \delta)^2 \tau^2 p^2 q^2 n_1 n_2 (1 - \tau^2 \delta) \geq (1 - \delta)^3 \tau^2 p^2 q^2 N. \end{aligned} \tag{49}$$

Moreover, since, deterministically,

$$\theta_{G, \mathbb{R}}(u_1, v_1) \theta_{G, \mathbb{R}}(u_2, v_2) \leq \min\{p, q\} n_2 \cdot \min\{p, q\} n_1 \leq 4p^2 q^2 N,$$

using (42) again, we infer that

$$\begin{aligned} \mathbb{E}_{\text{bottom}} &\leq (1 + \delta) \tau q d_1 \cdot (1 + \delta) \tau q d_2 \cdot (1 - p_*) + 4p^2 q^2 N \cdot p_* \\ &\leq (1 + \delta)^2 \tau^2 p^2 q^2 N + 4p^2 q^2 N \cdot \tau^2 \delta \leq (1 + 6\delta + \delta^2) \tau^2 p^2 q^2 N. \end{aligned} \tag{50}$$

Finally, combining the bounds on  $\mathbb{E}_{\text{top}}$  and  $\mathbb{E}_{\text{bottom}}$  and using (45), we conclude that

$$\frac{p_{t+1}(e, G)}{p_{t+1}(f, G)} \geq \frac{(1 - \delta)^3}{(1 + 6\delta + \delta^2)} \geq 1 - 9\delta \geq 1 - \gamma_t, \tag{51}$$

and the proof of Lemma 6 is complete. □

### 4.3. Proof of Lemma 8

By Lemmas 9 and 10 the events (47) (for each  $i \in [2]$ ) and (48) hold simultaneously with probability  $1 - O(N^{-C})$ . Hence it suffices to fix an arbitrary integer  $t < t_0$  and a realisation of  $\mathbb{R}(t)$ , say  $\mathbb{R}(t) = G$  satisfying inequalities (47) (for each  $i \in [2]$ ) and (48) and to prove that  $G$  is  $\delta(t)$ -typical. Fix any such  $G$  and define events

$$\mathcal{E} := \bigcup_{i \in [2]} \left\{ \max_{u, v \in V_i, u \neq v} \left| \frac{\theta_{G, \mathbb{R}}(u, v)}{\tau q d_i} - 1 \right| > \delta(t) \right\} \quad \text{and} \quad \mathcal{F}_{e,f} := \{e \in \mathbb{R}, f \notin \mathbb{R}\}.$$

Since, conditioning on the event  $\mathbb{R}(t) = G$ , we have  $\theta_{G, \mathbb{R}}(u_i, v_i) = \theta_t(u_i, v_i)$ , by the choice of  $G$  we have

$$\mathbb{P}(\mathcal{E} \mid \mathbb{R}(t) = G) \leq 2e^{-2\lambda(t)} \tag{52}$$

and

$$\min_{e, f \in K \setminus G} \mathbb{P}(\mathcal{F}_{e,f} \mid \mathbb{R}(t) = G) \geq e^{-\lambda(t)}. \tag{53}$$

Note that the probability on the LHS of (42) does not change if we replace  $G \subseteq \mathbb{R}$  by  $\mathbb{R}(t) = G$ , since conditioning on either event makes  $\mathbb{R}$  uniformly distributed over  $\mathcal{R}_G$  (i.e., biregular graphs containing  $G$ ) and the random variables  $\theta_{G, \mathbb{R}}(u_i, v_i)$  do not depend on the random ordering of the

edges of  $\mathbb{R}$ . Hence it remains to prove that (52) and (53) imply, for any distinct edges  $e = v_1v_2$ ,  $f = u_1u_2 \in K \setminus G$ , that

$$\mathbb{P} \left( \max_{i \in [2]: u_i \neq v_i} \left| \frac{\theta_{G, \mathbb{R}}(u_i, v_i)}{\tau q d_i} - 1 \right| > \delta(t) \mid \mathbb{R}(t) = G, e \in \mathbb{R}, f \notin \mathbb{R} \right) \leq \tau^2 \delta(t). \tag{54}$$

The probability on the LHS of (54) is at most  $\mathbb{P}(\mathcal{E} \mid \mathbb{R}(t) = G, \mathcal{F}_{e,f})$ . Inequalities (52) and (53) imply that

$$\mathbb{P}(\mathcal{E} \mid \mathbb{R}(t) = G, \mathcal{F}_{e,f}) = \frac{\mathbb{P}(\mathcal{E} \cap \mathcal{F}_{e,f} \mid \mathbb{R}(t) = G)}{\mathbb{P}(\mathcal{F}_{e,f} \mid \mathbb{R}(t) = G)} \leq \frac{\mathbb{P}(\mathcal{E} \mid \mathbb{R}(t) = G)}{\mathbb{P}(\mathcal{F}_{e,f} \mid \mathbb{R}(t) = G)} \leq 2e^{-\lambda(t)}.$$

Finally, even a quick glance at the definitions of  $\delta(t)$  and  $\tau_0$  (see (43) and (25)) reveals that  $\min\{\tau_0, \delta(t)\} \geq 1/\hat{n}$ . Thus, we get, with a huge margin,

$$2e^{-\lambda(t)} \leq 2/N^6 \leq 1/N^2 \leq 1/\hat{n}^4 \leq \tau_0^2 \delta(t) \leq \tau^2 \delta(t). \quad \square$$

### 5. Degrees and co-degrees

In this section we prove facts about the neighbourhood structure of  $\mathbb{R}(t)$ , one of which will be enough to deduce Lemma 9, while the other two will be used in the proof of Lemma 10 in Section 6. We start, in Subsection 5.1, with a tail bound for the co-degrees in  $\mathbb{R} = \mathbb{R}(n_1, n_2, p)$  (Lemma 11). In Subsection 5.2 we analyse the process  $(\mathbb{R}(t))_t$ . First, we show that the vertex degrees in the process grow proportionally until almost the very end (Lemma 13). Then, conditioning on  $\mathbb{R}$  having concentrated co-degrees, we prove that the co-degrees in  $\mathbb{R}(t)$  do not exceed their expectation too much (Lemma 14). Finally, in Subsection 5.3, we present a proof of Lemma 9 based on Lemma 11.

#### 5.1. Co-degrees in the random biregular graph

Recall that  $\Gamma_F(v)$  is the set of neighbours of a vertex  $v$  in a graph  $F$ . We define the *co-degree* of two distinct vertices  $u, v$  as

$$\text{cod}_F(u, v) := |\Gamma_F(u) \cap \Gamma_F(v)|. \tag{55}$$

A few times we will use the following simple observation: for  $F \subseteq K$  and distinct  $u, v \in V_1$ ,

$$\begin{aligned} \text{cod}_F(u, v) &= |\Gamma_F(u)| + |\Gamma_F(v)| - |\Gamma_F(u) \cup \Gamma_F(v)| \\ &= \text{deg}_F(u) + \text{deg}_F(v) - (n_2 - \text{cod}_{K \setminus F}(u, v)), \end{aligned} \tag{56}$$

where, recall,  $K = K_{n_1, n_2}$  is the complete bipartite graph.

Due to symmetry, we prove the following concentration result for pairs of vertices on one side of the bipartition only.

**Lemma 11.** *Suppose that  $\hat{p}\hat{n} \rightarrow \infty$  and let  $\lambda = \lambda(n_1, n_2)$  be such that  $\lambda \rightarrow \infty$ . Then, for any distinct  $u_1, v_1 \in V_1$ ,*

$$\mathbb{P} \left( \left| \text{cod}_{\mathbb{R}}(u_1, v_1) - p^2 n_2 \right| > 20 \left( \hat{p}^3 n_2 \mathbb{I} + \frac{\hat{p} n_2}{\hat{n}} + \sqrt{\lambda \hat{p}^2 n_2} \right) + \lambda \right) = O \left( \sqrt{N} e^{-\lambda} \right),$$

where  $\mathbb{I}$  is defined in (3).

**Proof.** We first claim that it is sufficient to assume  $p \leq 1/2$  and prove that for any distinct  $u_1, v_1 \in V_1$

$$\mathbb{P} \left( \left| \text{cod}_{\mathbb{R}}(u_1, v_1) - p^2 n_2 \right| > 20 \left( p^3 n_2 \mathbb{I} + \frac{p n_2}{\hat{n}} + \sqrt{\lambda p^2 n_2} \right) + \lambda \right) = O \left( \sqrt{N} e^{-\lambda} \right). \tag{57}$$

To see this, note that by (56), recalling  $q = 1 - p$ ,

$$\text{cod}_{\mathbb{R}}(u_1, v_1) - p^2 n_2 = 2pn_2 - (n_2 - \text{cod}_{K \setminus \mathbb{R}}(u_1, v_1)) - p^2 n_2 = \text{cod}_{K \setminus \mathbb{R}}(u_1, v_1) - q^2 n_2.$$

Further,  $K \setminus \mathbb{R} = K \setminus \mathbb{R}(n_1, n_2, p)$  has the same distribution as  $\mathbb{R}(n_1, n_2, q)$ . So, if  $p > 1/2$ , the lemma follows by applying (57) with  $q$  instead of  $p$ .  $\square$

The crude proof idea comes from our anticipation that  $\text{cod}_{\mathbb{R}}(u_1, v_1)$  behaves similarly to  $\text{cod}_{\mathbb{G}(n_1, n_2, p)}(u_1, v_1)$ , which is distributed as  $\text{Bin}(n_2, p^2)$  or, approximately, as  $\text{Po}(p^2 n_2)$ . We will show that each tail of  $\text{cod}_{\mathbb{R}}(u_1, v_1)$  is comparable to the tail of  $\text{Po}(\mu)$  with expectation  $\mu$  fairly close to  $p^2 n_2$  and then apply the Chernoff bounds (this is packaged in Claim 12 below). Further we consider the cases  $\mathbb{I} = 1$  and  $\mathbb{I} = 0$  and apply Claim 12 analogously to the proof of Theorem 2.1 in [7]: in the case  $\mathbb{I} = 1$  we use switchings and in the case  $\mathbb{I} = 0$  we use asymptotic enumeration (Theorem 5).

Fix distinct  $u_1, v_1 \in V_1$  and set  $X(u_1, v_1) := \text{cod}_{\mathbb{R}}(u_1, v_1)$ . Further, recall that  $\mathcal{R}(n_1, n_2, p)$  is the class of subgraphs of  $K_{n_1, n_2}$  such that every  $v \in V_i$  has degree  $d_i$ , for  $i = 1, 2$ , and let

$$\mathcal{R}_k(u_1, v_1) = \{H \in \mathcal{R}(n_1, n_2, p) : \text{cod}_H(u_1, v_1) = k\}, \quad k = 0, \dots, d_1.$$

**Claim 12.** Fix two vertices  $u_1, v_1 \in V_1$ . Suppose that positive numbers  $r_k, k = 0, \dots, d_1$ , are such that

$$|\mathcal{R}_k(u_1, v_1)| \sim r_k, \text{ uniformly in } k \text{ as } (n_1, n_2) \rightarrow \infty,$$

and there exist numbers  $0 \leq \mu_- \leq \mu_+ \leq N$  such that

$$\frac{r_k}{r_{k-1}} \leq \frac{\mu_+}{k}, \text{ for } k \in [\mu_+, d_1] \quad \text{and} \quad \frac{r_{k-1}}{r_k} \leq \frac{k}{\mu_-}, \text{ for } k \in [1, \mu_-].$$

If  $x \geq \sqrt{2\mu_+ \lambda} + \lambda$ , then

$$\mathbb{P}(X(u_1, v_1) \geq \mu_+ + x) + \mathbb{P}(X(u_1, v_1) \leq \mu_- - x) = O(\sqrt{N}e^{-\lambda}). \tag{58}$$

**Proof.** Note that if  $Z_+ \sim \text{Po}(\mu_+)$ , then  $\mu_+/k = \mathbb{P}(Z_+ = k) / \mathbb{P}(Z_+ = k - 1)$ . Setting  $m = \lfloor \mu_+ \rfloor$ , for any integer  $i \in \{m, \dots, d_1\}$ , and abbreviating  $X := X(u_1, v_1)$  and  $\mathcal{R}_j := \mathcal{R}_j(u_1, v_1)$ , we have

$$\begin{aligned} \mathbb{P}(X \geq i) &= \frac{\sum_{j \geq i} |\mathcal{R}_j|}{|\mathcal{R}(n_1, n_2, p)|} \leq \sum_{j \geq i} \frac{|\mathcal{R}_j|}{|\mathcal{R}_m|} \sim \sum_{j \geq i} \frac{r_j}{r_m} = \sum_{j \geq i} \prod_{k=m+1}^j \frac{r_k}{r_{k-1}} \leq \sum_{j \geq i} \prod_{k=m+1}^j \frac{\mu_+}{k} \\ &= \sum_{j \geq i} \prod_{k=m+1}^j \frac{\mathbb{P}(Z_+ = k)}{\mathbb{P}(Z_+ = k - 1)} = \sum_{j \geq i} \frac{\mathbb{P}(Z_+ = j)}{\mathbb{P}(Z_+ = m)} = \frac{\mathbb{P}(Z_+ \geq i)}{\mathbb{P}(Z_+ = m)}. \end{aligned}$$

Since  $\mu_+ \leq N$ , by (19) we have  $\mathbb{P}(Z_+ = m) = \Omega(1/\sqrt{\mu_+}) = \Omega(1/\sqrt{N})$ , and conclude that

$$\mathbb{P}(X \geq i) = O\left(\sqrt{N} \cdot \mathbb{P}(Z_+ \geq i)\right), \quad i \geq \mu_+. \tag{59}$$

Similarly, if  $Z_- \sim \text{Po}(\mu_-)$ , then for  $i \leq m := \lfloor \mu_- \rfloor$ , using  $k/\mu_- = \mathbb{P}(Z_- = k - 1) / \mathbb{P}(Z_- = k)$ ,

$$\mathbb{P}(X \leq i) \leq \sum_{j \leq i} \frac{|\mathcal{R}_j|}{|\mathcal{R}_m|} \sim \sum_{j \leq i} \prod_{k=j+1}^m \frac{r_{k-1}}{r_k} \leq \sum_{j \leq i} \prod_{k=j+1}^m \frac{\mathbb{P}(Z_- = k - 1)}{\mathbb{P}(Z_- = k)} = \frac{\mathbb{P}(Z_- \leq i)}{\mathbb{P}(Z_- = m)},$$

and therefore, again using (19),

$$\mathbb{P}(X \leq i) = O\left(\sqrt{N} \cdot \mathbb{P}(Z_- \leq i)\right), \quad i \leq \mu_-. \tag{60}$$

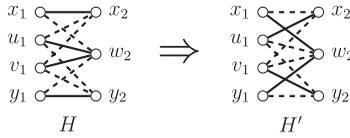


Figure 3. Switching between  $H \in \mathcal{R}_k(u_1, v_1)$  and  $H' \in \mathcal{R}_{k-1}(u_1, v_1)$ : solid edges are in  $H$  and  $H'$  and the dashed ones in  $K \setminus H$  and  $K \setminus H'$ , respectively.

Since the RHS of (58) does not depend on  $x$ , we can assume the equality:  $x = \sqrt{2\mu_+ \lambda} + \lambda$ . Inequalities (17) and (18) imply

$$\begin{aligned} \mathbb{P}(Z_+ \geq \mu_+ + x) + \mathbb{P}(Z_- \leq \mu_- - x) &\leq \exp\left(-\frac{x^2}{2(\mu_+ + x/3)}\right) + \exp\left(-\frac{x^2}{2\mu_-}\right) \\ \boxed{\mu_- \leq \mu_+, x > 0} &\leq 2 \exp\left(-\frac{x^2}{2(\mu_+ + x/3)}\right) \\ \boxed{x = \sqrt{2\mu_+ \lambda}} &= 2 \exp\left(-\frac{\lambda(2\mu_+ + 2\sqrt{2\mu_+ \lambda} + \lambda)}{2\mu_+ + \sqrt{2\mu_+ \lambda}/3 + \lambda/3}\right) \leq 2e^{-\lambda}. \end{aligned}$$

Combining this with (59) and (60) yields (58). □

It remains to prove (57). We consider separately the cases  $\mathbb{I} = 1$  and  $\mathbb{I} = 0$ .

**Case  $\mathbb{I} = 1$ .** If  $p \geq 1/5$ , then deterministically  $\text{cod}_{\mathbb{R}}(u_1, u_2) \leq pn_2 \leq 5p^2n_2 \leq p^2n_2 + 20p^3n_2$  and  $\text{cod}_{\mathbb{R}}(u_1, u_2) - p^2n_2 \geq -p^2n_2 \geq -5p^3n_2$ , and hence (57) holds trivially. So we further assume  $p < 1/5$ .

Setting  $r_k = |\mathcal{R}_k|$ , we are going to prove bounds on  $r_k/r_{k-1}$ ,  $k \geq 1$ , using a switching between  $\mathcal{R} = \mathcal{R}_k$  and  $\mathcal{R}' = \mathcal{R}_{k-1}$ . To this end, recall our terminology from Subsection 2.2 (with  $G = \emptyset$ ): any graph  $H \subseteq K$  is interpreted as a colouring of the edges of  $K$ , with the edges in  $H$  blue and the rest red. We define a forward switching as follows: pick a common blue neighbour  $w_2 \in \Gamma_H(u_1) \cap \Gamma_H(v_1)$  and find two alternating 4-cycles  $u_1 w_2 x_1 x_2$  and  $v_1 w_2 y_1 y_2$  so that  $x_1 \neq y_1$ ,  $x_2 \neq y_2$ . Moreover, we restrict the choice of the cycles to those for which  $v_1 x_2$  and  $u_1 y_2$  are red; this is to make sure that swapping of the colours on each of the two cycles (but keeping the colour of  $v_1 x_2, u_1 y_2$  red) indeed decreases  $\text{cod}_{\mathbb{R}}(u_1, v_1)$  by one, thus mapping  $H \in \mathcal{R}_k$  to  $H' \in \mathcal{R}_{k-1}$  (see Figure 3).

Let  $B_k := B(\mathcal{R}_k, \mathcal{R}_{k-1})$  be the auxiliary graph corresponding to the described switching. Since the number of choices of  $w_2$  is  $k$  for any  $H \in \mathcal{R}_k$ , we will show upper and lower bounds on  $\text{deg}_{B_k}(H)/k$  by fixing  $w_2$  and proving upper and lower bounds on the possible choices of  $(x_1, x_2, y_1, y_2)$ .

For the upper bound, since there are  $n_2 - 2d_1 + k$  common red neighbours of  $u_1$  and  $v_1$  (one can use (56) for this), the number of choices of distinct  $x_2, y_2$  is exactly  $(n_2 - 2d_1 + k)_2$ , while the number of choices of  $x_1$  and  $y_1$ , as blue neighbours of, resp.,  $x_2$  and  $y_2$ , is at most  $d_2^2$ . Ignoring the requirements that  $x_1 \neq y_1$  and that both  $x_1 w_2$  and  $y_1 w_2$  should be red, we have shown

$$\text{deg}_{B_k}(H)/k \leq (n_2 - 2d_1 + k)_2 d_2^2. \tag{61}$$

For the lower bound, we subtract from the upper bound (61) the number of choices of  $(x_1, x_2, y_1, y_2)$  for which either  $x_1 = y_1$  or at least one of  $x_1 w_2$  and  $y_1 w_2$  is blue (the assumption  $x_2 \neq y_2$  was taken into account in the upper bound). The number of choices with  $x_1 = y_1$  is at most  $n_1 d_1^2$ , and those with, say,  $x_1 w_2$  blue, is at most  $d_2 d_1 \cdot (n_2 - 2d_1 + k)_2$ . Indeed, there

are  $\deg_H(w_2) \leq d_2$  candidates for  $x_1$  and, then, at most  $d_1$  candidates for  $x_2$ ,  $\text{cod}_{K \setminus H}(u_1, u_2) = n_2 - 2d_1 + k$  candidates for  $y_2$ , and then  $d_2$  choices of  $y_1$ . Therefore

$$\begin{aligned} \deg_{B_k}(H)/k &\geq (n_2 - 2d_1 + k)_2 d_2^2 - n_1 d_1^2 - 2(n_2 - 2d_1 + k) d_1 d_2^2 \\ &= (n_2 - 2d_1 + k)_2 d_2^2 \left( 1 - \frac{n_1 d_1^2}{(n_2 - 2d_1 + k)_2 d_2^2} - \frac{2d_1}{n_2 - 2d_1 + k - 1} \right) \\ \boxed{d_1 = n_2 p, d_2 = n_1 p, k \geq 1} &\geq (n_2 - 2d_1 + k)_2 d_2^2 \left( 1 - \frac{n_2^2}{(n_2 - 2pn_2)^2 n_1} - \frac{2pn_2}{n_2 - 2pn_2} \right) \\ &= (n_2 - 2d_1 + k)_2 d_2^2 \left( 1 - \frac{1}{(1 - 2p)^2 n_1} - \frac{2p}{1 - 2p} \right) \\ \boxed{p \leq 1/5} &\geq (n_2 - 2d_1 + k)_2 d_2^2 \left( 1 - \left( \frac{25}{9pn_1} + \frac{10}{3} \right) p \right) \\ \boxed{\hat{p}\hat{n} \rightarrow \infty} &\geq (n_2 - 2d_1 + k)_2 d_2^2 (1 - 4p). \end{aligned}$$

A moment of thought (and a glance at Figure 3) reveals that the backward switching corresponds to choosing a common red neighbour  $w_2$  of  $u_1$  and  $v_1$ , choosing alternating cycles  $u_1 w_2 x_1 x_2$  and  $v_1 w_2 y_1 y_2$  such that  $x_1 \neq y_1$  and the edges  $v_1 x_2$  and  $u_1 y_2$  are red (note that these assumptions imply  $x_2 \neq y_2$ ), and swapping the colours along the cycles.

Since for every  $H' \in \mathcal{R}_{k-1}$  the number of choices of  $x_2 \in \Gamma_{H'}(u_1) \setminus \Gamma_{H'}(v_1)$  and  $y_2 \in \Gamma_{H'}(v_1) \setminus \Gamma_{H'}(u_1)$  is exactly  $(d_1 - k + 1)^2$ , we will bound  $\deg_{B_k}(H')/(d_1 - k + 1)^2$  from above and below by fixing  $x_2, y_2$  and estimating the number of possible triplets  $(w_2, x_1, y_1)$ . For the upper bound, we can choose  $w_2$  in exactly  $\text{cod}_{K \setminus H'}(u_1, u_2) = n_2 - 2d_1 + k - 1$  ways, and a pair  $x_1, y_1$  of distinct blue neighbours of  $w_2$  in at most  $(d_2)_2$  ways (we ignore the requirement that  $x_1 x_2$  and  $y_1 y_2$  are red). Thus,

$$\deg_{B_k}(H')/(d_1 - k + 1)^2 \leq (n_2 - 2d_1 + k - 1)(d_2)_2. \tag{62}$$

For the lower bound, given  $x_2, y_2$ , we need to subtract from the upper bound (62) the number of choices of  $w_2, x_1, y_1$  for which  $x_1 x_2$  or  $y_1 y_2$  is blue. In the former case, ignoring other constraints, there are at most  $d_2$  choices of a blue neighbour  $x_1$  of  $x_2$ , at most  $d_1$  choices of a blue neighbour  $w_2$  of  $x_1$ , and at most  $d_2$  choices of a blue neighbour  $y_1$  of  $w_2$ . By symmetry, the total number of bad choices is at most  $2d_1 d_2^2$ , whence

$$\begin{aligned} \deg_{B_k}(H')/(d_1 - k + 1)^2 &\geq (n_2 - 2d_1 + k - 1)(d_2)_2 - 2d_1 d_2^2 \\ \boxed{k \geq 1} &\geq (n_2 - 2d_1 + k - 1) d_2^2 \left( 1 - \frac{1}{d_2} - \frac{2d_1}{n_2 - 2d_1} \right) \\ \boxed{d_1 = n_2 p, p \leq 1/5} &\geq (n_2 - 2d_1 + k - 1) d_2^2 \left( 1 - d_2^{-1} - 4p \right). \end{aligned}$$

Since, by our assumptions,  $d_2 \geq \hat{p}\hat{n} \rightarrow \infty$  and  $p \leq 1/5$ , it follows that the graphs  $B_k$  for  $k = 1, \dots, d_1$  have minimum degrees at least 1. In particular, this means that starting with any  $p$ -biregular graph, by applying a certain number of switchings, we can obtain a  $p$ -biregular graph with arbitrary co-degree. Since we implicitly assume  $\mathcal{R}(n_1, n_2, p)$  is non-empty, this implies that all classes  $\mathcal{R}_k, k = 0, \dots, d_1$ , are non-empty, i.e., the numbers  $r_k = |\mathcal{R}_k|$  are all positive, satisfying one of the conditions of Claim 12.

Using (14) and bounds on the degrees in  $B_k$ , for  $k = 1, \dots, d_1$ , we have

$$\begin{aligned} \frac{r_k}{r_{k-1}} &\leq \frac{\max_{H' \in \mathcal{R}'} \deg_{B_k}(H')}{\min_{H \in \mathcal{R}} \deg_{B_k}(H)} \leq \frac{(d_1 - k + 1)^2}{k(n_2 - 2d_1 + k)} \cdot \frac{1}{1 - 4p} \\ &\leq \frac{(d_1 - k + 1)^2}{k(n_2 - 2d_1 + k)} \cdot (1 + 20p), \end{aligned} \tag{63}$$

(where the last inequality holds because  $p \leq 1/5$  implies  $\frac{1}{1-4p} = 1 + \frac{4p}{1-4p} \leq 1 + 20p$ ), and

$$\frac{r_k}{r_{k-1}} \geq \frac{\min_{H' \in \mathcal{R}'} \deg_{B_k}(H')}{\max_{H \in \mathcal{R}} \deg_{B_k}(H)} \geq \frac{(d_1 - k + 1)^2}{k(n_2 - 2d_1 + k)} \left(1 - d_2^{-1} - 4p\right). \tag{64}$$

Let  $\mu$  be a real number such that

$$\frac{(d_1 - \mu + 1)^2}{\mu(n_2 - 2d_1 + \mu)} = 1. \tag{65}$$

After solving for  $\mu$ , we have

$$\mu = \frac{(d_1 + 1)^2}{n_2 + 2} \in [p^2 n_2, p^2 n_2(1 + 3d_1^{-1})]. \tag{66}$$

Define

$$\mu_+ := \mu(1 + 20p), \quad \text{and} \quad \mu_- := \mu(1 - d_2^{-1} - 4p). \tag{67}$$

From (63), (65), and (67), for  $\mu \leq k \leq d_1$ , we have

$$\frac{r_k}{r_{k-1}} \leq \frac{(d_1 - \mu + 1)^2 (1 + 20p)}{\mu(n_2 - 2d_1 + \mu)} \cdot \frac{\mu}{k} = \frac{\mu_+}{k}.$$

On the other hand, from (64), (65), and (67) that, for  $1 \leq k \leq \mu$ , we have

$$\frac{r_{k-1}}{r_k} \leq \frac{\mu(n_2 - 2d_1 + \mu)}{(d_1 - \mu + 1)^2 (1 - 1/d_2 - 4p)} \cdot \frac{k}{\mu} = \frac{k}{\mu_-}.$$

Note that, since  $\min\{d_1, d_2\} = p\hat{n} \rightarrow \infty$  and  $p \leq 1/5$ , it follows by (66), and (67) that

$$p^2 n_2(1 - (p\hat{n})^{-1} - 4p) \leq \mu_- \leq \mu_+ \leq p^2 n_2(1 + 15(p\hat{n})^{-1} + 20p) \leq 6p^2 n_2. \tag{68}$$

Therefore,

$$\mu_+ \leq p^2 n_2 + 20 \left(p^3 n_2 + \frac{pn_2}{\hat{n}}\right) \quad \text{and} \quad \mu_- \geq p^2 n_2 - 20 \left(p^3 n_2 + \frac{pn_2}{\hat{n}}\right). \tag{69}$$

Consequently, setting

$$x := 20\sqrt{\lambda p^2 n_2} + \lambda, \tag{70}$$

we have, by (69)

$$\mathbb{P}\left(|X - p^2 n_2| \geq 20 \left(p^3 n_2 + \frac{pn_2}{\hat{n}}\right) + x\right) \leq \mathbb{P}(X \geq \mu_+ + x) + \mathbb{P}(X \leq \mu_- - x).$$

Noting that the last inequality in (68) implies  $x \geq \sqrt{\mu_+ n_2} + \lambda$ , using Claim 12 we obtain (57), completing the proof in the case  $\mathbb{I} = 1$ .

**Case  $\mathbb{I} = 0$ .** For the switching argument used in the previous case it was crucial that  $p$  was small, as otherwise several estimates would be negative and so meaningless. Therefore, it cannot be used now. Fortunately, in this case we are in a position to apply an asymptotic enumeration approach based on Theorem 5 and analogous to the proof of Theorem 2.1 in [7].

Recall the notation  $\mathcal{R}(\mathbf{d}_1, \mathbf{d}_2)$  defined before Theorem 5. Writing  $A = \Gamma_H(u_1)$  and  $B = \Gamma_H(v_1)$ , we note that every graph in  $\mathcal{R}_k$  induces an ordered partition of  $V_2$  into four sets  $A \cap B, A \setminus B, B \setminus A$  and  $V_2 \setminus (A \cup B)$ , of sizes, respectively,  $k, d_1 - k, d_1 - k$ , and  $n_2 - 2d_1 + k$ . There are exactly

$$\Pi(n_2, d_1, d_2, k) := \frac{n_2!}{k!(d_1 - k)!^2(n_2 - 2d_1 + k)!}$$

such partitions. After removing vertices  $u_1$  and  $v_1$ , we obtain a graph  $H^* \in \mathcal{R}(\mathbf{d}_1, \mathbf{d}_2)$  on  $(V_1 \setminus \{u_1, v_1\}, V_2)$ , with  $\mathbf{d}_1$  having all its  $n_1 - 2$  entries equal to  $d_1$ , and entries of  $\mathbf{d}_2$  determined by the partition: entries equal to  $d_2 - 2$  on  $A \cap B$ , to  $d_2 - 1$  on  $A \Delta B$  and the remaining ones equal to  $d_2$ . Since  $H^*$  together with the ordered partition uniquely determines  $H$ , we have

$$|\mathcal{R}_k| = \Pi(n_2, d_1, d_2, k) \times |\mathcal{R}(\mathbf{d}_1, \mathbf{d}_2)|.$$

Let us check that the three assumptions (i)–(iii) of Theorem 5 are satisfied. When doing so, we should remember that we have  $n_1 - 2$  instead of  $n_1$ , but that the parameter  $p = d_1/n_2$  remains intact. With foresight, choose  $a = 0.35$  and, say,  $b = 0.1$ , and, for convenience  $C = 1$ , which determine, via Theorem 5, an  $\epsilon > 0$ .

Note that  $\mathbb{I} = 0$  implies

$$\hat{n} \geq \frac{4 \max\{n_1, n_2\}}{\log N} \quad \text{and} \quad \frac{n_1}{n_2} + \frac{n_2}{n_1} \leq \frac{p \log N}{2} \leq p \log \max\{n_1, n_2\}. \tag{71}$$

As the degrees on one side are all equal and on the other side they differ from each other by at most 2, assumption (i) holds true with a big margin. Using  $q \geq 1/2$  and the first inequality in (71), we get

$$(pq)^2 \min\{n_1 - 2, n_2\}^{1+\epsilon} \geq \frac{p^2 \hat{n}^{1+\epsilon}}{4} (1 + o(1)) = \Omega\left(\frac{(\max\{n_1, n_2\})^{1+\epsilon}}{(\log N)^{3+\epsilon}}\right) \gg \max\{n_1, n_2\},$$

which implies assumption (ii) for large  $\hat{n}$ . Finally, using elementary inequalities  $(1 - 2p)^2 \leq q$  (since  $p \leq 1/2$ ),  $1 \leq (n_1/n_2 + n_2/n_1)/2$ , and the second inequality in (71), we obtain

$$\begin{aligned} \frac{(1 - 2p)^2}{4pq} \left(1 + \frac{5(n_1 - 2)}{6n_2} + \frac{5n_2}{6(n_1 - 1)}\right) &\leq \frac{1}{4p} \left(\frac{1}{2} + \frac{5}{6}\right) \left(\frac{n_1}{n_2} + \frac{n_2}{n_1}\right) (1 + o(1)) \\ &\leq \frac{1}{3} \log \max\{n_1, n_2\} (1 + o(1)) \ll 0.35 \cdot \log \max\{n_1 - 2, n_2\}, \end{aligned}$$

which for large  $\hat{n}$  implies assumption (iii) with  $a = 0.35$ .

Note that in (20) we have  $D_1 = 0$ , while  $D_2 \leq 4n_2 \ll pqn_1n_2$ , by the assumption  $\hat{p}\hat{n} \rightarrow \infty$ . Thus, uniformly over  $k$ , the exponent in (20) is  $-1/2 + o(1)$  and, by Theorem 5,

$$|\mathcal{R}_k| \sim \frac{e^{-1/2} n_2! \binom{n_2}{d_1}^{n_1-2} \binom{n_1-2}{d_2-2}^k \binom{n_1-2}{d_2-1}^{2d_1-2k} \binom{n_1-2}{d_2}^{n_1-2d_1+k}}{k!(d_1 - k)!^2(n_2 - 2d_1 + k)! \binom{(n_1-2)n_2}{(n_1-2)d_1}} =: r_k.$$

Straightforward calculations yield

$$\frac{r_k}{r_{k-1}} = \frac{(d_1 - k + 1)^2(n_1 - 2 - d_2 + 1)(d_2 - 1)}{k(n_2 - 2d_1 + k) \cdot d_2(n_1 - d_2)} \leq \frac{(d_1 - k + 1)^2}{k(n_2 - 2d_1 + k)}$$

and, because  $d_2/n_1 = p \leq 1/2$ ,

$$\frac{r_k}{r_{k-1}} \geq \frac{(d_1 - k + 1)^2}{k(n_2 - 2d_1 + k)} \left(1 - 2d_2^{-1}\right).$$

(Compare with (63) and (64) to note the absence of  $\Theta(p)$  error terms.) With  $\mu$  as in (66), we redefine

$$\mu_+ := \mu \quad \text{and} \quad \mu_- := \mu(1 - 2d_2^{-1}). \tag{72}$$

From (66) and (72) it follows that

$$\mu_+ = \mu \leq p^2 n_2(1 + 3d_1^{-1}) \leq p^2 n_2 + \frac{3pn_2}{n_1} \leq p^2 n_2 + \frac{20pn_2}{\hat{n}} \tag{73}$$

and

$$\mu_- = p^2 n_2(1 - 2d_2^{-1}) \geq p^2 n_2 - 2p \geq p^2 n_2 - \frac{20pn_2}{\hat{n}}. \tag{74}$$

From (73) it follows  $\mu_+ \leq 6p^2 n_2$ , which implies that  $x$  defined in (70) satisfies  $x \geq \sqrt{2\mu_+ \lambda} + \lambda$ . Consequently, using (73), (74) and Claim 12 we infer

$$\mathbb{P} \left( |X - p^2 n_2| \geq \frac{20pn_2}{\hat{n}} + x \right) \leq \mathbb{P}(X \geq \mu_+ + x) + \mathbb{P}(X \leq \mu_- - x) = O \left( \sqrt{N} e^{-\lambda} \right).$$

We have obtained (57) in the case  $\mathbb{I} = 0$ . □

### 5.2. Degrees and co-degrees in $\mathbb{R}(t)$

For convenience, having fixed  $H \in \mathcal{R}(n_1, n_2, p)$ , we will denote by  $\mathbb{P}_H$  and  $\mathbb{E}_H$  the conditional probability and expectation with respect to the event  $\mathbb{R} = H$ . In the conditional space defined by such an event,  $(\eta_1, \dots, \eta_M)$  is just a uniformly random permutation of the edges of  $H$  and therefore  $\mathbb{R}(t)$  is a uniformly random  $t$ -subset of edges of  $H$ .

We first show that the degrees of vertices in the process  $\mathbb{R}(t)$  grow proportionally almost until the end. Recall that  $d_1 = pn_2$ ,  $d_2 = pn_1$ , while  $\tau = \tau(t)$  and  $\tau_0$  are defined, respectively, in (24) and (25).

**Lemma 13.** *If  $\lambda = \lambda(n_1, n_2) \leq \tau_0 p \hat{n}$ , then for  $i = 1, 2$ , with probability  $1 - O(N^2 e^{-9\lambda/4})$*

$$\forall t < t_0 \quad \forall v \in V_i \quad |\deg_{\mathbb{R}(t)}(v) - (1 - \tau)d_i| \leq 3\sqrt{\lambda \tau d_i}. \tag{75}$$

**Proof.** Fix  $t < t_0$  and  $v \in V_i$  and set  $X_t(v) := \deg_{\mathbb{R}(t)}(v)$ . For any  $H \in \mathcal{R}(n_1, n_2, p)$ , conditioning on  $\mathbb{R} = H$ , the random variable  $X_t(v) \sim \text{Hyp}(M, d_i, M - t)$  has a hypergeometric distribution. Since the distribution does not depend on  $H$ ,  $X_t(v)$  has the same distribution unconditionally. In view of (24),  $\mathbb{E} X_t(v) = \tau d_i$ , therefore combining (17) and (18), for all  $x > 0$ ,

$$\mathbb{P}(|X_t(v) - \tau d_i| \geq x) \leq 2 \exp \left\{ -\frac{x^2}{2(\tau d_i + x/3)} \right\} = 2 \exp \left\{ -\frac{x^2}{2\tau d_i(1 + x/(3\tau d_i))} \right\}. \tag{76}$$

Let  $x := 3\sqrt{\lambda \tau d_i}$ . Since  $\tau \geq \tau_0$ , by the assumption  $\lambda \leq \tau_0 p \hat{n}$ ,

$$x/(3\tau d_i) = \sqrt{\lambda/(\tau d_i)} \leq \sqrt{\tau_0 \hat{n} p / (\tau_0 d_i)} \leq 1. \tag{77}$$

Consequently, taking the union bound over all  $(1 - \tau_0)Mn_i \leq N^2$  choices of  $t$  and  $v$  and using (76) and (77), we conclude that (75) fails with probability at most

$$2N^2 \exp \left\{ -\frac{x^2}{2\tau d_i(1 + x/(3\tau d_i))} \right\} \leq 2N^2 \exp \left\{ -\frac{9\tau d_i \lambda}{4\tau d_i} \right\} = 2N^2 e^{-9\lambda/4}. \tag{78}$$

□

In the proof of Lemma 10, to have a pseudorandom-like property of  $\mathbb{R} \setminus \mathbb{R}(t)$  for  $p > 0.49$ , we will need a bound on the upper tail of co-degrees in  $\mathbb{R}(t)$ . Recall the definition of co-degree from (55).

**Lemma 14.** *Assume that  $n_1 \geq n_2$  and  $p > 0.49$ . If  $\lambda = \lambda(n_1, n_2) \geq \log N$ , then, with probability  $1 - O(N^3 e^{-\lambda/3})$ ,*

$$\forall t \leq M \quad \forall u_1 \neq v_1 \quad \text{cod}_{\mathbb{R}(t)}(u_1, v_1) \leq (1 - \tau)^2 p^2 n_2 + 20q^3 n_2 \mathbb{I} + 15\sqrt{\lambda n_2}. \tag{78}$$

**Proof.** If  $\lambda \geq n_2$ , inequality (78) holds trivially, so let us assume  $\lambda < n_2$ . We can condition on  $\mathbb{R} = H$  satisfying

$$\begin{aligned} \forall u_1, v_1 \in V_1 \quad \text{cod}_H(u_1, v_1) &\leq p^2 n_2 + 20 \left( \hat{p}^3 n_2 \mathbb{I} + n_2 \hat{p} / \hat{n} + \sqrt{\hat{p}^2 \lambda n_2} \right) + \lambda \\ \boxed{\hat{p} \leq \min\{q, 0.5\}, n_2 = \hat{n}} &\leq p^2 n_2 + 20 \left( q^3 n_2 \mathbb{I} + 0.5 + 0.5\sqrt{\lambda n_2} \right) + \lambda \\ \boxed{1 \ll \lambda < n_2} &\leq p^2 n_2 + 20q^3 n_2 \mathbb{I} + 12\sqrt{\lambda n_2}. \end{aligned} \tag{79}$$

Indeed, the probability of the opposite event can be bounded, by Lemma 11 and the union bound, by  $O(n_1^2 \sqrt{N} e^{-\lambda}) \ll N^3 e^{-\lambda}$ . Taking this bound into account, it thus suffices to show that if  $H$  satisfies (79), then for any distinct  $u_1, v_1 \in V_1$  and  $t \leq M$

$$\mathbb{P}_H \left( \text{cod}_{\mathbb{R}(t)}(u_1, v_1) > (1 - \tau)^2 p^2 n_2 + 20q^3 n_2 \mathbb{I} + 15\sqrt{\lambda n_2} \right) \leq 2e^{-\lambda/3}, \tag{80}$$

since then the proof is completed by a union bound over  $O(N^3)$  choices of  $t, u_1, v_1$ .

Fix  $t \leq M, u_1 \neq v_1$  and let  $Y := |\Gamma_{\mathbb{R}(t)}(u_1) \cap \Gamma_H(v_1)|$  and  $\text{cod}_H := \text{cod}_H(u_1, v_1)$ . The distribution of  $Y$  conditioned on  $\mathbb{R} = H$  is hypergeometric  $\text{Hyp}(M, \text{cod}_H, t)$ , and hence, by (79),

$$\mu_Y := \mathbb{E}_H Y = \frac{t \text{cod}_H}{M} = (1 - \tau) \text{cod}_H \leq (1 - \tau) p^2 n_2 + 20q^3 n_2 \mathbb{I} + 12\sqrt{\lambda n_2}.$$

Using (17), a trivial bound  $\mu_Y \leq n_2$ , and our assumption  $\lambda < n_2$ ,

$$\mathbb{P}_H \left( Y \geq \mu_Y + \sqrt{\lambda n_2} \right) \leq \exp \left\{ -\frac{\lambda n_2}{2(\mu_Y + \sqrt{\lambda n_2}/3)} \right\} \leq e^{-3\lambda/8} \leq e^{-\lambda/3}.$$

Since also trivially  $Y \leq \min\{d_1, t\}$ , we have shown that given  $\mathbb{R} = H$ , with probability at least  $1 - e^{-\lambda/3}$ ,

$$Y \leq y_0 := \min\{\mu_Y + \sqrt{\lambda n_2}, d_1, t\} \leq (1 - \tau) p^2 n_2 + 20q^3 n_2 \mathbb{I} + 13\sqrt{\lambda n_2}. \tag{81}$$

Let  $\mathcal{E}_{(80)}$  be the event that the inequality in (80) holds. By the law of total probability,

$$\begin{aligned} \mathbb{P}_H(\mathcal{E}_{(80)}) &= \sum_y \mathbb{P}_H(\mathcal{E}_{(80)} \mid Y = y) \mathbb{P}_H(Y = y) \\ \boxed{(81)} &\leq \sum_{y \leq y_0} \mathbb{P}_H(\mathcal{E}_{(80)} \mid Y = y) \mathbb{P}_H(Y = y) + e^{-\lambda/3}. \end{aligned}$$

Hence, to prove (80), it suffices to show that, for any  $y = 0, \dots, y_0$ ,

$$\mathbb{P}_H(\mathcal{E}_{(80)} \mid Y = y) \leq e^{-\lambda/3}. \tag{82}$$

Fix an integer  $y \in [0, y_0]$  and a set  $S \subseteq V_2$  of size  $|S| = y \leq y_0$ . Under an additional conditioning  $\Gamma_{\mathbb{R}(t)}(u_1) \cap \Gamma_H(v_1) = S$ , set  $\mathbb{R}(t)$  is the union of a fixed  $y$ -element set  $\{u_1 w : w \in S\}$  and a random  $(t - y)$ -element subset of  $E(H) \setminus \{u_1 w : w \in \Gamma_H(u_1) \cap \Gamma_H(v_1)\}$ . Thus, in this conditional space,  $X := \text{cod}_{\mathbb{R}(t)}(u_1, v_1)$  counts how many of these  $t - y$  random edges fall into the set  $\{v_1 w : w \in S\}$

and therefore  $X \sim \text{Hyp}(M - \text{cod}_H, y, t - y)$ . Moreover, the distribution of  $X$  is the same for all  $S$  of size  $y$ , so  $X$  has the same distribution when conditioned on  $Y = y$ . In particular,

$$\begin{aligned} \mu_X &:= \mathbb{E}_H(X \mid Y = y) = \frac{y(t - y)}{M - \text{cod}_H} \\ \boxed{y \geq 0} &\leq \frac{yt}{M} \left( 1 + \frac{\text{cod}_H}{M - \text{cod}_H} \right) \\ \boxed{M = pn_1n_2, \text{cod}_H \leq pn_2} &\leq (1 - \tau)y \left( 1 + \frac{1}{n_1 - 1} \right) \\ \boxed{y \leq y_0 \text{ and (81)}} &\leq (1 - \tau)^2 p^2 n_2 + 20q^3 n_2 \mathbb{I} + 13\sqrt{\lambda n_2} + \frac{n_2}{n_1 - 1}. \\ \boxed{n_1 \geq n_2} &\leq (1 - \tau)^2 p^2 n_2 + 20q^3 n_2 \mathbb{I} + 14\sqrt{\lambda n_2}. \end{aligned}$$

Using again (17) as well as inequalities  $\mu_X \leq n_2$  and  $\lambda \leq n_2$ , we infer that

$$\mathbb{P}_H \left( X \geq \mu_X + \sqrt{\lambda n_2} \mid Y = y \right) \leq \exp \left\{ -\frac{\lambda n_2}{2(\mu_X + \sqrt{\lambda n_2}/3)} \right\} \leq e^{-\lambda/3},$$

which, together with the above upper bound on  $\mu_X$ , implies (82). This completes the proof of Lemma 14.  $\square$

**5.3. Proof of Lemma 9**

Recall the definitions of  $\tau = \tau(t)$  in (24),  $\delta(t)$  in (43), and  $\lambda(t)$  in (44). In this proof we utilise some technical bounds on  $\tau := \tau(t)$  and  $\gamma_t$  proved in Section 8 (see Proposition 23). In particular, the bound (146) on  $\gamma_t$ , together with (45), implies that

$$\delta(t) \leq 1/9 \leq 1. \tag{83}$$

(We do not even need to remember what  $\gamma_t$  is to see that.)

We now derive an upper bound on  $\lambda(t)$ . By (147), with a huge margin,

$$6 \log N \leq \frac{\tau \hat{p} \hat{n}}{16(C + 3)} \leq \frac{\tau p q \hat{n}}{8(C + 3)}.$$

On the other hand, for  $p > 0.49$ , squaring and rearranging the inequality (145) implies

$$\frac{64 \log N}{\tau p q} \leq \frac{\tau p q \hat{n}}{8(C + 3)}.$$

Summing up, for any  $p$ ,

$$\lambda(t) \leq \frac{\tau p q \hat{n}}{4(C + 3)}. \tag{84}$$

Without loss of generality, we assume that  $i = 1$ . Let  $\mathcal{F}_t$  be the family of graphs  $H$  satisfying, for any distinct  $u_1, v_1 \in V_1$ ,

$$|d_1 - \text{cod}_H(u_1, v_1) - p q n_2| \leq 20 \hat{p} n_2 \left( \hat{p}^2 \mathbb{I} + \frac{1}{\hat{n}} + \sqrt{\frac{(C + 3)\lambda(t)}{n_2} + \frac{(C + 3)\lambda(t)}{20 \hat{p} n_2}} \right). \tag{85}$$

Lemma 11 with  $\lambda = (C + 3)\lambda(t)$  and a union bound over the  $O(N^2)$  choices of  $u_1, v_1$  imply

$$\mathbb{P}(\mathbb{R} \in \mathcal{F}_t) = 1 - O(N^{2.5} e^{-(C+3)\lambda(t)}) \stackrel{\lambda(t) \geq 6 \log n}{\geq} 1 - O(N^{-(C+1)} e^{-2\lambda(t)}). \tag{86}$$

Writing  $\delta_* = \sqrt{\frac{(C+3)\lambda(t)}{\tau pq \hat{n}}}$  and noting that (84) implies  $\delta_* \leq 1/2$ , the sum of the last three terms in the parentheses in (85) is at most

$$2 \cdot \sqrt{\frac{(C+3)\lambda(t)}{\hat{n}}} + \frac{(C+3)\lambda(t)}{20\hat{p}\hat{n}} \leq 2\sqrt{pq}\delta_* + \frac{pq\delta_*^2}{20\hat{p}} \leq (1 + 1/40)\delta_* \leq \frac{3\delta_*}{\sqrt{6}}.$$

In addition, the factor in front of the brackets is at most  $40pq n_2$ . Thus, (85) implies

$$|d_1 - \text{cod}_H(u_1, v_1) - pq n_2| \leq 40pq n_2 \left( \hat{p}^2 \mathbb{I} + \frac{3\delta_*}{\sqrt{6}} \right) = pq n_2 \cdot \delta(t)/3. \tag{87}$$

We claim that for  $t = 0, \dots, t_0 - 1$

$$\max_{H \in \mathcal{F}_t} \mathbb{P}_H \left( \max_{u_1 \neq v_1} \left| \frac{\theta_t(u_1, v_1)}{\tau pq n_2} - 1 \right| \geq \delta(t) \right) \leq 2N^2 e^{-(C+3)\lambda(t)}, \tag{88}$$

and deferring its proof to the end we first show how (86) and (88) imply the lemma.

Inequalities (86) and (88) imply

$$\begin{aligned} \mathbb{P} \left( \max_{u_1 \neq v_1} \left| \frac{\theta_t(u_1, v_1)}{\tau pq n_2} - 1 \right| \geq \delta(t) \right) &\leq 2N^2 e^{-(C+3)\lambda(t)} + \mathbb{P}(\mathbb{R} \notin \mathcal{F}_t) \\ &= O(N^{-(C+1)} e^{-2\lambda(t)}). \end{aligned} \tag{89}$$

Consider random variables, for  $t = 0, \dots, t_0 - 1$ ,

$$Y_t := \mathbb{P} \left( \max_{u_1 \neq v_1} \left| \frac{\theta_t(u_1, v_1)}{\tau pq n_2} - 1 \right| > \delta(t) \mid \mathbb{R}(t) \right).$$

Using Markov’s inequality and (89), we infer that

$$\mathbb{P} \left( Y_t > e^{-2\lambda(t)} \right) \leq e^{2\lambda(t)} \mathbb{E} Y_t = e^{2\lambda(t)} \mathbb{P} \left( \max_{u_1 \neq v_1} \left| \frac{\theta_t(u_1, v_1)}{\tau pq n_2} - 1 \right| \geq \delta(t) \right) = O(N^{-(C+1)}),$$

which, taking the union bound over the  $O(N)$  choices of  $t$ , implies that (47) holds with the desired probability, completing the proof of lemma.

Returning to the proof of (88), fix  $t < t_0, H \in \mathcal{F}_t$  and two distinct vertices  $u_1, v_1 \in V_1$ . Conditioning on  $\mathbb{R} = H$ , note that, recalling (46), random variable  $X := \theta_t(u_1, v_1) = |\Gamma_{H \setminus \mathbb{R}(t)}(u_i) \cap \Gamma_{K \setminus H}(v_i)|$  counts elements in the intersection of two subsets of  $E(H)$ : a fixed set  $\{(u_i, w) : w \in \Gamma_H(u_i) \cap \Gamma_{K \setminus H}(v_i)\}$  of size  $d_1 - \text{cod}_H(u_1, v_1)$  and a random set  $H \setminus \mathbb{R}(t)$  of size  $M - t$ . Hence  $X \sim \text{Hyp}(M, d_1 - \text{cod}_H(u_1, v_1), M - t)$  has the hypergeometric distribution with expectation

$$\mu_X := \mathbb{E} X = (d_1 - \text{cod}_H(u_1, v_1))(M - t)/M = \tau(d_1 - \text{cod}_H(u_1, v_1)).$$

Note that by (87),

$$|\mu_X - \tau pq n_2| \leq \tau pq n_2 \cdot \delta(t)/3. \tag{90}$$

Let  $\lambda^* := (3C + 9)\lambda(t)$ . From (90), (83), and (84) it follows that

$$\mu_X \geq \frac{26}{27} \tau pq n_2 \geq \lambda^*. \tag{91}$$

Note that (43) implies

$$\delta(t)^2 \geq 7200 \frac{\lambda^*}{\tau pq n_2} \geq 10 \frac{\lambda^*}{\tau pq n_2}. \tag{92}$$

By (17), (18), and (91),

$$\begin{aligned} \mathbb{P}_H \left( |X - \mu_X| \geq \sqrt{\mu_X \lambda^*} \right) &\leq 2 \exp \left\{ -\frac{\lambda^*}{2 \left( 1 + \frac{1}{3} \sqrt{\lambda^* / \mu_X} \right)} \right\} \\ &\leq 2e^{-3\lambda^*/8} \leq 2e^{-\lambda^*/3} = 2e^{-(C+3)\lambda(t)}. \end{aligned}$$

Thus, with probability  $1 - 2e^{-(C+3)\lambda(t)}$

$$\begin{aligned} |\theta_t(u_1, v_1) - \tau pqn_2| &\leq |X - \mu_X| + |\mu_X - \tau pqn_2| \\ &\leq \sqrt{\mu_X \lambda^*} + |\mu_X - \tau pqn_2| \\ \tag{90, 92} &\leq \sqrt{(1 + \delta(t)/3) \tau pqn_2 \cdot \delta^2(t) \tau pqn_2 / 10} + (\delta(t)/3) \tau pqn_2 \\ \tag{83} &\leq \delta(t) \tau pqn_2 (\sqrt{(1 + 1/27)/10} + 1/3) \\ &\leq \delta(t) \tau pqn_2. \end{aligned}$$

Hence, applying also the union bound over all  $n_1^2 \leq N^2$  choices of  $u_1, v_1$ , we infer (88). □

### 6. Alternating cycles in regularly 2-edge-coloured jumbled graphs

The ultimate goal of this section is to prove Lemma 10. While for  $p \leq 0.49$  the proof follows relatively easily from Lemma 13 by a standard switching technique, the case  $p > 0.49$  is much more involved. To cope with it, we first study the existence of alternating walks and cycles in a class of 2-edge-coloured pseudorandom graphs.

In Subsection 6.1, we define an appropriate notion of pseudorandom bipartite graphs (*jumbledness*), inspired by a similar notion introduced implicitly by Thomason in [14]. We show that for  $p > 0.49$  and suitably chosen parameters, the random graph  $K \setminus \mathbb{R}(t)$  is jumbled with high probability (Lemma 17).

The next two subsections are devoted to 2-edge-coloured jumbled graphs which are almost regular in each colour. After proving a technical Lemma 18 in Subsection 6.2, in Subsection 6.3 we show the existence of alternating short walks between any two vertices of almost regular 2-edge-coloured jumbled graphs (Lemma 19).

An immediate consequence of Lemma 19 is Lemma 20, which states that every edge belongs to an alternating short cycle. The latter result together with a standard switching argument (Proposition 4) will be used in the proof of Lemma 10 for  $p > 0.49$ . That proof, for both cases  $p \leq 0.49$  and  $p > 0.49$ , is presented in Subsection 6.4.

#### 6.1. Jumbled graphs

Let  $K := K_{n_1, n_2}$  be the complete bipartite graph with partition classes  $V_1$  and  $V_2$ , where  $|V_i| = n_i$ . Given a bipartite graph  $F \subseteq K$  and two subsets  $A \subseteq V_1, B \subseteq V_2$ , denote by  $e_F(A, B)$  the number of edges of  $F$  between  $A$  and  $B$ . Recall that  $N = n_1 n_2$  and  $M = pN$ .

Given real numbers  $\pi, \delta \in (0, 1)$ , we say that a graph  $F \subseteq K$  is  $(\pi, \delta)$ -jumbled if for every  $A \subseteq V_1$  and  $B \subseteq V_2$

$$|e_F(A, B) - \pi |A| |B|| \leq \delta \sqrt{N |A| |B|}.$$

The following result of Thomason [14, Theorem 2], which quantifies a variant of jumbledness in terms of the degrees and co-degrees of a graph, will turn out to be crucial for us.

**Theorem 15** ([14]). Let  $F \subseteq K$  be a bipartite graph and  $\rho \in (0, 1)$  and  $\mu \geq 0$  be given. If

$$\min_{v \in V_1} \deg_F(v) \geq \rho n_2 \quad \text{and} \quad \max_{u, v \in V_1: u \neq v} \text{cod}_F(u, v) \leq \rho^2 n_2 + \mu, \tag{93}$$

then, for all  $A \subseteq V_1$  and  $B \subseteq V_2$ ,

$$|e_F(A, B) - \rho|A||B|| \leq \sqrt{(\rho n_2 + \mu|A|)|A||B|} + |B| \mathbb{I}_{\{|A| \rho < 1\}}.$$

**Remark 16.** The proof in [14] is given only in the case  $n_1 = n_2 = n$ . However, it carries over in this more general setting, as in [14]  $n$  always refers to  $|V_2|$ .

Recall that  $K \setminus \mathbb{R}(t)$  has precisely  $N - t = N - (1 - \tau)M = (\tau p + q)N$  edges. The following technical result states that, under the conditions of Lemma 6, with high probability,  $K \setminus \mathbb{R}(t)$  is jumbled for parameters which are tailored for Lemma 20.

**Lemma 17.** Let  $\alpha = \min\{\tau p, q\}$  and  $\pi := \tau p + q$ . For every constant  $C > 0$  and  $p > 0.49$ , if assumptions (27) and (28) hold, then, with probability  $1 - O(N^{-C})$ , for all  $t < t_0$

$$K \setminus \mathbb{R}(t) \text{ is } (\pi, \alpha/16)\text{-jumbled}$$

and for any  $e \in K \setminus \mathbb{R}(t)$

$$K \setminus (\mathbb{R}(t) \cup \{e\}) \text{ is } (\pi, \alpha/16)\text{-jumbled.}$$

**Proof.** W.l.o.g., we assume that  $n_1 \geq n_2$ . Let  $\lambda := 3(C + 4) \log N$ , and

$$\delta := 20 (\lambda/n_2)^{1/4} + 10q^{3/2}\mathbb{I}. \tag{94}$$

The plan is to show that

$$\delta \leq \frac{\alpha}{16} \tag{95}$$

and that with the correct probability,  $K \setminus \mathbb{R}(t)$  and  $K \setminus \mathbb{R}(t) \cup \{e\}$  are in fact  $(\pi, \delta)$ -jumbled.

We start with the proof of (95). Notice that, by (28),

$$\sqrt{\lambda/n_2} \leq (\lambda/n_2)^{1/4} \leq \frac{q}{680} \leq \frac{\pi}{680} \leq \frac{1}{680}. \tag{96}$$

Since  $\hat{p} > \frac{49}{51}q$ , condition (27) implies that  $1/320 \geq q^{1/2}\mathbb{I}$ . After multiplying both sides by  $10q$ , we get

$$\frac{q}{32} \geq 10 \cdot q^{3/2}\mathbb{I},$$

which together with the second inequality in (96) implies

$$\frac{q}{16} = \frac{q}{32} + \frac{q}{32} \geq 20 (\lambda/n_2)^{1/4} + 10 \cdot q^{3/2}\mathbb{I} = \delta.$$

On the other hand, using  $p > 0.49$ ,  $\tau \geq \tau_0$  and the definition (25) of  $\tau_0$ , we infer that

$$\begin{aligned} \frac{\tau p}{16} &\geq \frac{0.49 \cdot \tau_0}{16} = \frac{0.49 \cdot 700 \cdot (3(C + 4))^{1/4}}{16} \left( \left( \frac{\log N}{n_2} \right)^{1/4} + q^{3/2}\mathbb{I} \right) \\ &\geq 20 (\lambda/n_2)^{1/4} + 10q^{3/2}\mathbb{I} = \delta. \end{aligned}$$

Hence  $\alpha/16 = \min\{\tau p/16, q/16\} \geq \delta$ , implying (95).

We now prove the jumbledness, first focusing on  $K \setminus \mathbb{R}(t)$  and then indicating the tiny change in calculation for  $K \setminus (\mathbb{R}(t) \cup \{e\})$ . Fixing an arbitrary  $t < t_0$ , we will first show that, with probability  $1 - O(N^{-C-1})$ , conditions (93) of Theorem 15 are satisfied by  $F = K \setminus \mathbb{R}(t)$  and  $F = K \setminus (\mathbb{R}(t) \cup \{e\})$  for suitably chosen  $\rho$  and  $\mu$ . Then we will apply Theorem 15 to deduce that

$K \setminus \mathbb{R}(t)$  is  $(\pi, \delta)$ -jumbled. Lemma 17 will follow by applying the union bound over all (at most  $t_0 \leq M \leq N$ ) choices of  $t$ .

By (147),  $\lambda \leq \tau_0 p \hat{n}$  with a big room to spare. Note that

$$(1 - \tau)p = 1 - \pi, \tag{97}$$

which implies that  $|\deg_{K \setminus \mathbb{R}(t)}(v) - \pi n_2| = |\deg_{\mathbb{R}(t)}(v) - (1 - \tau)d_1|$ . Hence, by Lemma 13, with probability  $1 - O(N^2 e^{-\lambda})$ ,

$$\max_{v \in V_1} |\deg_{K \setminus \mathbb{R}(t)}(v) - \pi n_2| \leq 3\sqrt{\frac{4}{9}\lambda\tau d_1} \leq 2\sqrt{\lambda n_2} \leq 3\sqrt{\lambda n_2}. \tag{98}$$

Moreover, recalling that  $n_1 \geq n_2$  and, again using (97), Lemma 14 implies that, with probability  $1 - O(N^3 e^{-\lambda/3})$ ,

$$\max_{u, v \in V_1, u \neq v} \text{cod}_{\mathbb{R}(t)}(u, v) \leq (1 - \pi)^2 n_2 + 20q^3 n_2 \mathbb{I} + 15\sqrt{\lambda n_2}. \tag{99}$$

Since  $\lambda = 3(C + 4) \log N$ , the intersection of events (98) and (99) holds with probability  $1 - O(N^3 e^{-\lambda/3}) = 1 - O(N^{-C-1})$ .

Note that for distinct  $u, v \in V_1$ , by (56) and (98)

$$\text{cod}_{K \setminus \mathbb{R}(t)}(u, v) \leq \text{cod}_{\mathbb{R}(t)}(u, v) + (2\pi - 1)n_2 + 6\sqrt{\lambda n_2}, \tag{100}$$

which, by (99), implies that

$$\max_{u, v \in V_1, u \neq v} \text{cod}_{K \setminus \mathbb{R}(t)}(u, v) \leq \pi^2 n_2 + 20q^3 n_2 \mathbb{I} + (15 + 6)\sqrt{\lambda n_2}. \tag{101}$$

Set

$$\rho := \pi - 3\sqrt{\lambda/n_2} \quad \text{and} \quad \mu := 20q^3 n_2 \mathbb{I} + (15 + 12)n_2 \sqrt{\lambda/n_2},$$

and note that by the inequality  $\pi \leq 1$  and by (96), we have  $0 < \rho < 1$ . Furthermore,  $\rho^2 \geq \pi^2 - 6\sqrt{\lambda/n_2}$ . Hence, (98) and (101) imply the assumptions (93) for  $F = K \setminus \mathbb{R}(t)$  with the above  $\rho$  and  $\mu$ . Consequently, by Theorem 15 (using  $a \leq n_1$  and  $\rho \leq \pi$ ),

$$|e_{K \setminus \mathbb{R}(t)}(A, B) - \rho ab| \leq \sqrt{\left(\pi n_2 + 20q^3 n_1 n_2 \mathbb{I} + (15 + 12)n_1 n_2 \sqrt{\lambda/n_2}\right) ab} + b.$$

Further, since  $N = n_1 n_2$ ,  $n_1 \geq n_2 \geq b$  and  $a, \lambda \geq 1$ , we have, with a big margin,

$$\pi \leq 1 \leq n_1 \sqrt{\lambda/n_2} \quad \text{and} \quad b \leq \sqrt{bn_2} \leq \sqrt{Nab}(\lambda/n_2)^{1/4}.$$

It follows, applying the inequality  $\sqrt{x+y} \leq \sqrt{x} + \sqrt{y}$  as well as (94), that

$$|e_{K \setminus \mathbb{R}(t)}(A, B) - \rho ab| \leq \left( (\sqrt{15 + 13} + 1)(\lambda/n_2)^{1/4} + \sqrt{20}q^{3/2} \mathbb{I} \right) \sqrt{Nab} \leq \frac{\delta}{2} \sqrt{Nab}. \tag{102}$$

Moreover, note that using  $ab \leq n_1 n_2 = N$  and the first inequality in (96),

$$(\pi - \rho)ab = 3\sqrt{\lambda/n_2} \cdot ab \leq 3(\lambda/n_2)^{1/4} \cdot \sqrt{Nab} \leq \frac{\delta}{2} \sqrt{Nab}, \tag{103}$$

Hence, (102) and (103) imply

$$|e_{K \setminus \mathbb{R}(t)}(A, B) - \pi ab| \leq |e_{K \setminus \mathbb{R}(t)}(A, B) - \rho ab| + (\pi - \rho)ab \leq \delta \sqrt{Nab},$$

meaning that  $K \setminus \mathbb{R}(t)$  is  $(\pi, \delta)$ -jumbled.

If above we replace  $K \setminus \mathbb{R}(t)$  by  $K \setminus (\mathbb{R}(t) \cup \{e\})$ , the upper bound in (98) and the bound in (101) still hold trivially, while the lower bound  $\pi n_2 - 3\sqrt{\lambda n_2}$  in (98) remains correct, since  $\deg_{K \setminus (\mathbb{R}(t) \cup \{e\})}(v) \geq \deg_{K \setminus \mathbb{R}(t)}(v) - 1$  and we have plenty of room in (98). Hence Theorem 15 applies with the same  $\rho$  and  $\mu$ , implying that  $K \setminus (\mathbb{R}(t) \cup \{e\})$  is also  $(\pi, \delta)$ -jumbled.  $\square$

**6.2. A technical inequality for blue-red graphs**

We find it convenient to introduce *relative* counterparts of basic graph quantities. Below  $i \in \{1, 2\}$ . As before, let  $K := K_{n_1, n_2}$  be the complete bipartite graph with partition classes  $V_1$  and  $V_2$ , where  $|V_i| = n_i$ . The *relative size* of a subset of vertices  $S \subseteq V_i$  is

$$s(S) := \frac{|S|}{n_i}.$$

Further, for  $X \subseteq V_1$  and  $Y \subseteq V_2$  and a subgraph  $F \subseteq K$ , we define *the relative edge count*

$$\varepsilon_F(X, Y) = \varepsilon_F(Y, X) = \frac{e_F(X, Y)}{n_1 n_2}.$$

Moreover, for  $v \in V_i$  and  $Y \subseteq V_{3-i}$ , we define *the relative degree*

$$d_F(v, Y) = \frac{e_F(\{v\}, Y)}{n_{3-i}}.$$

If  $Y = V_{3-i}$ , we shorten  $d_F(v, Y)$  to  $d_F(v)$ .

In this notation, a graph  $F$  is  $(\pi, \delta)$ -jumbled if for every  $X \subseteq V_1, Y \subseteq V_2$

$$|\varepsilon_F(X, Y) - \pi s(X)s(Y)| \leq \delta \sqrt{s(X)s(Y)}. \tag{104}$$

Now, let the edges of a graph  $F$  be 2-coloured by blue and red, and let  $B$  and  $R$  be the subgraphs of  $F$  induced by the edges of colour, resp., blue and red. We then call  $F = B \cup R$  a *blue-red graph*. Note that

$$\varepsilon_{B \cup R}(X, Y) = \varepsilon_F(X, Y) = \varepsilon_R(X, Y) + \varepsilon_B(X, Y).$$

We say that a blue-red graph  $F$  is  $(r, b, \delta)$ -regular, if

$$b - \delta \leq d_B(v) \leq b + \delta, \quad \text{and} \quad r - \delta \leq d_R(v) \leq r + \delta \quad \text{for every } v \in V_1 \cup V_2. \tag{105}$$

If  $F$  is at the same time  $(r + b, \delta)$ -jumbled and  $(r, b, \delta)$ -regular, as in the technical lemma below, we will sometimes loosely refer to such a graph as *regularly jumbled*.

Finally, for every  $S \subseteq V_i, i = 1, 2$ , set  $\bar{S} := V_i \setminus S$ .

**Lemma 18.** *Let  $r, b \in (0, 1)$  be real numbers and define  $\alpha := \min\{r, b\}$ . Let  $\nu < \alpha/16$  and  $\delta \leq \alpha/16$  be positive reals and let  $F \subseteq K$  be a  $(r, b, \delta)$ -regular,  $(b + r, \delta)$ -jumbled bipartite blue-red graph. If sets  $X \subseteq V_i, Y \subseteq V_{3-i}$  satisfy*

$$\varepsilon_B(X, \bar{Y}) + \varepsilon_R(\bar{X}, Y) \leq \nu, \tag{106}$$

and

$$\min\{bs(X), rs(Y)\} \leq \frac{rb}{r + b}, \tag{107}$$

then

$$\max\{bs(X), rs(Y)\} \leq \frac{\nu}{1 - 7\delta/\alpha}. \tag{108}$$

**Proof.** Since

$$\varepsilon_R(X, Y) = \frac{e_R(X, Y)}{n_1 n_2} = \frac{\sum_{v \in X} e_R(\{v\}, Y)}{n_1 n_2} = \frac{1}{n_1} \sum_{v \in X} d_R(v, Y),$$

from (105) we have

$$\varepsilon_R(X, Y) + \varepsilon_R(X, \bar{Y}) \leq (r + \delta)s(X) \tag{109}$$

and, similarly,

$$\varepsilon_B(X, Y) + \varepsilon_B(\bar{X}, Y) \leq (b + \delta)s(Y). \tag{110}$$

By summing (106), (109) and (110), we infer that

$$\varepsilon_F(X, Y) + \varepsilon_F(X, \bar{Y}) + \varepsilon_F(\bar{X}, Y) \leq \nu + (r + \delta)s(X) + (b + \delta)s(Y). \tag{111}$$

On the other hand, by (105),

$$\varepsilon_F(X, \bar{Y}) = \varepsilon_F(X, V_2) - \varepsilon_F(X, Y) \geq (b + r - 2\delta)s(X) - \varepsilon_F(X, Y)$$

and

$$\varepsilon_F(\bar{X}, Y) = \varepsilon_F(V_1, Y) - \varepsilon_F(X, Y) \geq (b + r - 2\delta)s(Y) - \varepsilon_F(X, Y).$$

Hence, by (104) with  $\pi = b + r$ ,

$$\begin{aligned} \varepsilon_F(X, Y) + \varepsilon_F(X, \bar{Y}) + \varepsilon_F(\bar{X}, Y) &\geq (b + r - 2\delta)(s(X) + s(Y)) - \varepsilon_F(X, Y) \\ &\geq (b + r - 2\delta)(s(X) + s(Y)) - (b + r)s(X)s(Y) - \delta\sqrt{s(X)s(Y)}. \end{aligned}$$

Comparing with (111), we obtain the inequality

$$bs(X) + rs(Y) - \frac{(b + r)bs(X)rs(Y)}{br} \leq \nu + \delta \left( 3s(X) + 3s(Y) + \sqrt{s(X)s(Y)} \right).$$

Denoting  $x := bs(X)$  and  $y := rs(Y)$  and  $h := rb/(b + r)$ , this becomes

$$x + y - \frac{xy}{h} \leq \nu + \delta \left( \frac{3x}{b} + \frac{3y}{r} + \sqrt{\frac{xy}{br}} \right) =: \psi. \tag{112}$$

Trivially, by the definitions of  $s(X)$  and  $\alpha$ , and by our assumptions on  $\nu$  and  $\delta$ , we have

$$\psi \leq \nu + 7\delta < \frac{1}{2}\alpha \leq \frac{1}{1/b + 1/r} = h.$$

Since our goal—inequality (108)—now reads as

$$\max\{x, y\} \leq \frac{\nu}{1 - 7\delta/\alpha},$$

to complete the proof it is enough to assume, without loss of generality, that  $\max\{x, y\} = y$  and show, equivalently, that

$$y \leq \nu + 7\delta y/\alpha. \tag{113}$$

By (107) we have  $x = \min\{x, y\} \leq h$ . Note that  $x = h$  cannot hold, since then the LHS of (112) would equal  $h$ , contradicting the fact that  $\psi < h$ . Hence, we have  $x < h$ , which, together with  $\psi < h$  and (112), implies that

$$y \leq \frac{\psi - x}{1 - x/h} = h - \frac{h - \psi}{1 - x/h} \leq h - (h - \psi) = \psi = \nu + \delta \left( \frac{3x}{b} + \frac{3y}{r} + \sqrt{\frac{xy}{br}} \right) \leq \nu + \frac{7\delta y}{\alpha},$$

and (113) is proved. □

### 6.3. Alternating walks and cycles

A cycle in a blue-red bipartite graph is said to be *alternating* if it is a union of a red matching and a blue matching, that is, every other edge is blue and the remaining edges are red. The ultimate goal of this subsection is to show that for every edge in a regularly jumbled blue-red bipartite graph, there is an alternating cycle of bounded length containing that edge. We are going to achieve it by utilising walks.

Given  $x, y \in V_1 \cup V_2$ , an *alternating walk from  $x$  to  $y$*  in a blue-red graph  $F$  is a sequence of (not necessarily distinct) vertices  $(v_1 = x, \dots, v_s = y)$  such that for each  $i = 1, \dots, s - 1$ ,  $v_i v_{i+1} \in F$ ,

every other edge is blue and the remaining edges are red. There is no restriction on the colour of the initial edge  $v_1v_2$ . The length of a walk is defined as the number of edges, or  $s - 1$ .

If the vertices  $v_1, \dots, v_s$  are all distinct, an alternating walk is called an *alternating path*. Note also that if  $F$  is bipartite,  $x \in V_1, y \in V_2$ , and the edge  $xy$  is, say, blue, then every alternating path from  $x$  to  $y$  which begins (and thus ends) with a red edge together with  $xy$  forms an alternating cycle containing  $xy$ .

The first result of this section asserts that regularly jumbled blue-red bipartite graphs have a short alternating walk between any pair of vertices.

**Lemma 19.** *Given  $r, b \in (0, 1)$  such that  $\alpha := \min\{r, b\}$ , let  $\delta \in (0, \alpha/16]$  and let  $F \subseteq K$  be an  $(r, b, \delta)$ -regular,  $(b + r, \delta)$ -jumbled blue-red graph. Let  $L = 4\lceil 16/rb \rceil + 1$ . For any  $x \in V_i$  and  $y \in V_{3-i}$  there exist at least two alternating walks from  $x$  to  $y$  of length at most  $L$ , one starting with a blue edge and another starting with a red edge.*

**Proof.** For  $w \in V_1 \cup V_2$  and an integer  $k \geq 1$ , define  $R_k^w$  and  $B_k^w$  as the sets of vertices  $v \in V(F)$  such that there is an alternating walk from  $v$  to  $w$  of length  $\ell \leq k, \ell \equiv k \pmod{2}$ , starting with, respectively, a red edge and a blue edge. (Note that these definitions concern walks ending with  $w$ .)

Clearly, for every  $k \geq 3, B_{k-2}^w \subseteq B_k^w$  and  $R_{k-2}^w \subseteq R_k^w$ . Observe also that for any  $k \geq 2$ , by definition the sets  $R_{k-1}^w$  and  $B_k^w$  are contained in opposite sides of the bipartition  $(V_1, V_2)$  and, moreover,

$$\varepsilon_B(\overline{B_k^w}, R_{k-1}^w) = 0. \tag{114}$$

By symmetry,  $B_{k-1}^w$  and  $R_k^w$  are contained in opposite sides of  $(V_1, V_2)$  and

$$\varepsilon_R(\overline{R_k^w}, B_{k-1}^w) = 0. \tag{115}$$

Set  $\nu = rb/16$  and note that, since  $r, b < 1$ , we have  $\nu < \alpha/16$ . There exists an integer  $t \leq T := \lceil 1/\nu \rceil = \lceil 16/rb \rceil$  such that

$$s(R_{2t+1}^w \setminus R_{2t-1}^w) \leq \nu, \tag{116}$$

since otherwise  $1 \geq s(R_{2T+1}^w) \geq \sum_{i=1}^T s(R_{2i+1}^w \setminus R_{2i-1}^w) > \nu T \geq 1$ , which is a contradiction.

By (114) and (116),

$$\varepsilon_B(R_{2t+1}^w, \overline{B_{2t}^w}) = \varepsilon_B(R_{2t-1}^w, \overline{B_{2t}^w}) + \varepsilon_B(R_{2t+1}^w \setminus R_{2t-1}^w, \overline{B_{2t}^w}) \leq s(R_{2t+1}^w \setminus R_{2t-1}^w) \leq \nu. \tag{117}$$

Combining (115) for  $k = 2t + 1$  and (117), we get

$$\varepsilon_B(R_{2t+1}^w, \overline{B_{2t}^w}) + \varepsilon_R(\overline{R_{2t+1}^w}, B_{2t}^w) \leq \nu. \tag{118}$$

Set  $X = R_{2t+1}^w, Y = B_{2t}^w$ , for convenience. We claim that

$$s(X) > r/(r + b) \quad \text{and} \quad s(Y) > b/(r + b). \tag{119}$$

Assuming the contrary, we have

$$\min\{bs(X), rs(Y)\} \leq br/(r + b),$$

which, together with (118), constitute the assumptions of Lemma 18. Applying it, we get

$$\max\{bs(X), rs(Y)\} \leq \frac{\nu}{1 - 7\delta/\alpha} = \frac{rb}{16(1 - 7\delta/\alpha)} \leq \frac{rb}{16(1 - 7/16)} = \frac{rb}{9}, \tag{120}$$

where the second inequality follows by our assumption  $\delta \leq \alpha/16$ .

On the other hand, since  $X$  contains the set  $R_1^w = \Gamma_R(w)$  of red neighbours of  $w$ , by  $(r, b, \delta)$ -regularity of  $F$  we have  $s(X) \geq r - \delta \geq \frac{15}{16}r$ , a contradiction with (120). Hence, we have shown

(119). Since  $X = R_{2t+1}^w \subseteq R_{2T+1}^w$  and  $Y = B_{2t}^w \subseteq B_{2T}^w$ , we also have

$$s(R_{2T+1}^w) > \frac{r}{r+b}, \quad s(B_{2T}^w) > \frac{b}{r+b}. \tag{121}$$

Since we chose  $w$  arbitrarily, (121) holds for  $w \in \{x, y\}$ , implying

$$s(R_{2T+1}^y) > \frac{r}{r+b}, \quad s(B_{2T}^x) > \frac{b}{r+b}. \tag{122}$$

Let us assume, without loss of generality, that  $x \in V_1$  and  $y \in V_2$ . Then,  $B_{2T}^x, R_{2T+1}^y \subseteq V_1$  and, in particular, for every vertex  $v \in B_{2T}^x$  there is a walk (of even length at most  $2T$ ) from  $v$  to  $x$  starting with a blue and thus ending with a red edge. By (122),  $s(B_{2T}^x) + s(R_{2T+1}^y) > 1$ , so there exists  $v \in B_{2T}^x \cap R_{2T+1}^y$ . This means that there is an alternating walk of length at most  $2T + (2T + 1) = 4\lceil 16/rb \rceil + 1 = L$  from  $x$  to  $y$  (through  $v$ ) that starts with a red edge.

By an analogous reasoning with the roles of the colours red and blue swapped,  $R_{2T}^x \cap B_{2T+1}^y \neq \emptyset$ , and thus there is an alternating walk of length at most  $2T + (2T + 1) = L$  from  $x$  to  $y$  which starts with a blue edge.  $\square$

The following result is an easy consequence of Lemma 19.

**Lemma 20.** *Let  $r, b \in (0, 1)$ ,  $\alpha := \min\{r, b\}$ , and  $\delta \in (0, \alpha/16]$ . If  $F \subseteq K$  is an  $(r, b, \delta)$ -regular,  $(r + b, \delta)$ -jumbled blue-red bipartite graph, then every edge of  $F$  belongs to an alternating cycle of length at most  $2D$ , where  $D = 2\lceil 16/rb \rceil + 1$ .*

**Proof.** Let  $xy$  be an edge with  $x \in V_1, y \in V_2$ . Without loss of generality, we assume that  $xy$  is blue. Then, by Lemma 19, there exists at least one alternating walk from  $x$  to  $y$  of length at most  $4\lceil 16/rb \rceil + 1 = 2D - 1$  starting with a red edge. Consider a shortest such walk  $W$ . We claim  $W$  is a path. Indeed, assume that  $W$  is not a path and let  $w$  be the first repeated vertex on  $W$ . If we remove the whole segment of the walk between the first two occurrences of  $w$ , what remains is still an alternating walk from  $x$  to  $y$  starting with a red edge (since this segment has an even number of edges), contradicting the minimality of  $W$ . The path  $W$  is not just a single edge  $xy$  (since  $xy$  is blue) and therefore  $W$  and  $xy$  form an alternating cycle of length at most  $2D$ .  $\square$

**6.4. Proof of Lemma 10**

Let  $\alpha = \min\{\tau p, q\}$ . Applying Lemma 13 with  $\lambda = (C + 1) \log N$  (note that the condition  $\lambda \leq \tau_0 p \hat{n}$  follows generously from (147)) we have that, with probability  $1 - O(N^{-C})$ , for every  $t < t_0$

$$\forall v_i \in V_i \quad \tau d_i(1 - \delta) \leq d_i - \deg_{\mathbb{R}(t)}(v_i) \leq \tau d_i(1 + \delta), \quad i \in \{1, 2\}, \tag{123}$$

where

$$\delta = 3\sqrt{(C + 1) \log N / (\tau p \hat{n})}$$

Whenever  $p > 0.49$ , by Lemma 17, with probability  $1 - O(N^{-C})$  for every  $t < t_0$  we have that

$$\begin{aligned} K \setminus \mathbb{R}(t) & \text{ is } (\tau p + q, \alpha/16)\text{-jumbled,} \\ \forall e \in K \setminus \mathbb{R}(t) \quad K \setminus (\mathbb{R}(t) \cup \{e\}) & \text{ is } (\tau p + q, \alpha/16)\text{-jumbled.} \end{aligned} \tag{124}$$

Fix  $\mathbb{R}(t) = G$  satisfying (123) and (124). It remains to prove that for every pair of distinct edges  $e, f \in K \setminus G$

$$\mathbb{P}(e \in \mathbb{R}, f \notin \mathbb{R} \mid \mathbb{R}(t) = G) \geq e^{-\lambda(t)} = N^{-2D}, \tag{125}$$

where  $D = 3$  if  $p \leq 0.49$  and  $D = 32/\tau pq + 3$ , if  $p > 0.49$ . For this, fix distinct edges  $e, f \in K \setminus G$ . Aiming to apply Proposition 4, we need to verify the assumption on the existence of alternating cycles containing a given edge.

Given a graph  $G'$  and  $H \in \mathcal{R}_{G'}$ , recall our convention to call the edges of  $H \setminus G'$  blue and the edges of  $K \setminus H$  red.

**Claim 21.** *Let  $G' \in \{G, G \cup \{e\}\}$ . For every  $H \in \mathcal{R}_{G'}$ , every edge  $g \in K \setminus G'$  is contained in an alternating cycle of length at most  $2D$ .*

From Claim 21 we complete the proof of (125) as follows. Since  $G$  is admissible, we have  $\mathcal{R}_G \neq \emptyset$ . Therefore Proposition 4 implies

$$\mathcal{R}_{G,e} \neq \emptyset, \quad \text{and} \quad \frac{|\mathcal{R}_{G,-e}|}{|\mathcal{R}_{G,e}|} \leq N^D - 1. \tag{126}$$

Since (126) implies  $\mathcal{R}_{G \cup \{e\}} = \mathcal{R}_{G,e} \neq \emptyset$ , Proposition 4, applied to graph  $G \cup \{e\}$  and edge  $f$ , implies

$$\mathcal{R}_{G \cup \{e\},-f} \neq \emptyset, \quad \text{and} \quad \frac{|\mathcal{R}_{G \cup \{e\},f}|}{|\mathcal{R}_{G \cup \{e\},-f}|} \leq N^D - 1. \tag{127}$$

Using (126) and (127), we infer

$$\begin{aligned} \frac{1}{\mathbb{P}(e \in \mathbb{R}, f \notin \mathbb{R} \mid \mathbb{R}(t) = G)} &= \frac{|\mathcal{R}_G|}{|\mathcal{R}_{G,e,-f}|} = \frac{|\mathcal{R}_G|}{|\mathcal{R}_{G,e}|} \cdot \frac{|\mathcal{R}_{G,e}|}{|\mathcal{R}_{G,e,-f}|} \\ &= \left(1 + \frac{|\mathcal{R}_{G,-e}|}{|\mathcal{R}_{G,e}|}\right) \cdot \left(1 + \frac{|\mathcal{R}_{G,e,f}|}{|\mathcal{R}_{G,e,-f}|}\right) \\ &= \left(1 + \frac{|\mathcal{R}_{G,-e}|}{|\mathcal{R}_{G,e}|}\right) \cdot \left(1 + \frac{|\mathcal{R}_{G \cup \{e\},f}|}{|\mathcal{R}_{G \cup \{e\},-f}|}\right) \leq N^{2D}, \end{aligned}$$

which implies (125).

It remains to prove Claim 21. As a preparation, we derive bounds on the vertex degrees in  $G \cup \{e\}$ . Note that the inequality (147) implies

$$\delta \leq 0.001, \tag{128}$$

and

$$\delta \tau d_i \geq 3\sqrt{C\tau p \hat{n} \log N} \geq 3C \log N \geq 1. \tag{129}$$

The latter, together with (123), implies that, for an arbitrary  $e \in K \setminus G$ ,

$$\forall v_i \in V_i \quad \tau d_i(1 - 2\delta) \leq d_i - \deg_{G \cup \{e\}}(v_i) \leq \tau d_i(1 + \delta), \quad i \in \{1, 2\} \tag{130}$$

We consider two cases with respect to  $p$ .

**Case  $p \leq 0.49$ .** We first claim that for any two vertices  $x_i \in V_i, i = 1, 2$ , there is an alternating path  $x_1 y_2 y_1 x_2$  such that  $x_1 y_2$  is red (and thus  $y_1 y_2$  is blue and  $y_1 x_2$  is red). The number of ways to choose a blue edge  $y_1 y_2$  is

$$M - |G'| \geq M - t - 1 = \tau p N - 1.$$

We bound the bad choices of  $y_1 y_2$  which do not give a desired alternating path. These correspond to the walks (we must permit  $y_1 = x_1$  and  $y_2 = x_2$ )  $x_1 y_2 y_1$  and  $x_2 y_1 y_2$  whose first edge in non-red, i.e., it belongs to  $H$ , while the second one is blue, i.e., it belongs to  $H \setminus G'$ . By second inequalities in (123) and in (130), there are at most  $d_1 \cdot \tau(1 + \delta)d_2$  choices of such  $x_1 y_2 y_1$  and at most  $d_2 \cdot \tau(1 +$

$\delta)d_1$  choices of such  $x_2y_1y_2$ , so altogether there are at most

$$2p^2\tau(1 + \delta)N \leq 0.98(1 + \delta)\tau pN$$

bad choices of  $y_1y_2$ . Thus, noting that  $1/\tau pN \leq \delta \leq 0.001$  (cf. (129) and (128)), the number of good choices of  $y_1y_2$  is

$$\tau pN - 1 - 2p^2\tau(1 + \delta)N \geq \tau pN(1 - 2p - 3\delta) > 0,$$

implying that there exists a desired path  $x_1y_2y_1x_2$ .

This immediately implies that if  $g = x_1x_2$  is blue, then  $g$  is contained in an alternating 4-cycle. If  $g = u_1u_2$  is red, then we choose blue neighbours  $x_1 \in \Gamma_{H \setminus G'}(u_2)$  and  $x_2 \in \Gamma_{H \setminus G'}(u_1)$  (which exist due to the lower bounds in (123) and (130) being positive). Since there exists an alternating path  $x_1y_2y_1x_2$  starting with a red edge, we obtain an alternating 6-cycle containing  $g$ . This proves Claim 6.4 in the case  $p \leq 0.49$ .

**Case  $p > 0.49$ .** We aim to apply Lemma 20. We first verify that, for  $G' \in \{G, G \cup \{e\}\}$  and every  $H \in \mathcal{R}_{G'}$ , blue-red graph  $K \setminus G'$  is  $(q, \tau p, \alpha/16)$ -regular. This assumption is trivial for the red graph  $K \setminus H$ , which is  $q$ -biregular regardless of  $G'$ . In view of (123) and (130), the relative degrees  $d_{H \setminus G'}(v)$  in the blue graph lie in the interval  $[\tau p - 2\delta\tau p, \tau p + 2\delta\tau p]$ . Since (28) implies

$$\delta\tau p = 3\sqrt{\frac{(C+1)\tau p \log N}{\hat{n}}} \leq 3\sqrt{\frac{(C+1)\log N}{\hat{n}}} \leq \sqrt{3} \left(\frac{q}{680}\right)^2 \leq \frac{q}{32}$$

and (128) implies  $\delta\tau p \leq \tau p/32$ , we obtain  $2\delta\tau p \leq \alpha/16$ . Hence, indeed,  $K \setminus G'$  is  $(q, \tau p, \alpha/16)$ -regular.

On the other hand, by (124)  $K \setminus G$  is also  $(\tau p + q, \alpha/16)$ -jumbled. Hence, by Lemma 20 with  $F = K \setminus G'$ ,  $r = q$  and  $b = \tau p$ , the edge  $g$  belongs to an alternating cycle of length at most  $2D$  with  $D = 2\lceil 16/\tau p q \rceil + 1 \leq 32/\tau p q + 3$ . Claim 21 is proven.  $\square$

### 7. Extension to non-bipartite graphs

Given integers  $n$  and  $d$ ,  $0 \leq d \leq n - 1$  such that  $nd$  is even, define the random regular graph  $\mathbb{R}(n, d)$  as a graph selected uniformly at random from the family  $\mathcal{R}(n, d)$  of all  $d$ -regular graphs on an  $n$ -vertex set  $V$ . To make the comparison with the binomial model  $\mathbb{G}(n, p)$  easier, similarly as in the bipartite case, we set  $p = \frac{d}{n-1}$  and define  $\mathbb{R}(n, p) := \mathbb{R}(n, d)$ . In what follows, we often suppress the parameter  $d$  and instead just assume that  $0 \leq p \leq 1$ ,  $p(n - 1)$  is an integer, and  $p(n - 1)n$  is even.

As described below, our proof of Theorem 2 can be adjusted to yield its non-bipartite version and, consequently, also a non-bipartite version of Corollary 3. Instead of formulating these two quite technical results, we limit ourselves to just stating their abridged version, analogous to Theorem 1. It confirms the Sandwich Conjecture of Kim and Vu [6] whenever  $d \gg (n \log n)^{3/4}$  and  $n - d \gg n^{3/4}(\log n)^{1/4}$ .

**Theorem 1'.** *If*

$$p \gg \frac{\log n}{n} \quad \text{and} \quad 1 - p \gg \left(\frac{\log n}{n}\right)^{1/4}, \tag{131}$$

*then for some  $m \sim p\binom{n}{2}$ , there is a joint distribution of random graphs  $\mathbb{G}(n, m)$  and  $\mathbb{R}(n, p)$  such that*

$$\mathbb{G}(n, m) \subseteq \mathbb{R}(n, p) \quad \text{a.a.s.}$$

If

$$p \gg \left(\frac{\log^3 n}{n}\right)^{1/4},$$

then for some  $m \sim p\binom{n}{2}$ , there is a joint distribution of random graphs  $\mathbb{G}(n, m)$  and  $\mathbb{R}(n, p)$  such that

$$\mathbb{R}(n, p) \subseteq \mathbb{G}(n, m) \quad \text{a.a.s.}$$

Moreover, in both inclusions, one can replace  $\mathbb{G}(n, m)$  by the binomial random graph  $\mathbb{G}(n, p')$ , for some  $p' \sim p$ .

To obtain a proof of Theorem 1', one would modify the proof of Theorem 2 and its prerequisites fixing, say,  $C = 1$ . For the bulk of the proof (see Sections 3–6) the changes are straightforward and consist mainly of replacing  $K = K_{n_1, n_2}$  by  $K_n$ , both  $n_1$  and  $n_2$  by  $n$ ,  $N = n_1 n_2$  by  $\binom{n}{2}$ , both  $d_1$  and  $d_2$  by  $p(n - 1)$ , as well as of setting  $\mathbb{I} = 0$ . In particular, we redefine (cf. (25))

$$\tau_0 := C_1 \begin{cases} \frac{\log n}{pn}, & p \leq 0.49, \\ \left(\frac{\log n}{n}\right)^{1/4}, & p > 0.49, \end{cases} \tag{132}$$

$$\gamma_t = C_2 \begin{cases} \sqrt{\frac{\log n}{\tau pn}}, & p \leq 0.49, \\ \sqrt{\frac{\log n}{\tau^2 q^2 n}}, & p > 0.49, \end{cases}$$

and

$$\gamma := C_3 \begin{cases} \sqrt{\frac{\log n}{pn}}, & p \leq 0.49, \\ \left(\frac{\log n}{n}\right)^{1/4}, & p > 0.49, \end{cases}$$

for some appropriately chosen constants  $C_1, C_2, C_3 > 0$  and replace assumptions (26)–(28) by conditions (131). Some other constants appearing in various definitions, might need to be updated, too.

The proofs of non-bipartite versions of Theorem 2, Claim 7, Lemmas 8, 9, 13, and 14 follow the bipartite ones in a straightforward way.

The proof of Lemma 6 is modified also in a straightforward way except for one technical change. In the switching graph  $B$  we consider 6-circuits rather than 6-cycles (that is, we allow the vertices  $x_1$  and  $x_2$  in Figure 2 coincide). With this definition the degrees in the switching graph  $B$  (cf. (39)–(40)) are now as follows. If edges  $f = u_1 u_2$  and  $e = v_1 v_2$  are disjoint, then

$$\deg_B(H) = \theta_{G,H}(u_1, v_1)\theta_{G,H}(u_2, v_2) + \theta_{G,H}(u_1, v_2)\theta_{G,H}(u_2, v_1), \quad H \in \mathcal{R} \tag{133}$$

and

$$\deg_B(H') = \theta_{G,H}(v_1, u_1)\theta_{G,H}(v_2, u_2) + \theta_{G,H}(v_1, u_2)\theta_{G,H}(v_2, u_1), \quad H' \in \mathcal{R}'. \tag{134}$$

If  $e$  and  $f$  share a vertex (without loss of generality,  $u_1 = v_1$ ), then

$$\deg_B(H) = \theta_{G,H}(u_2, v_2), \quad H \in \mathcal{R}, \quad \text{and} \quad \deg_B(H') = \theta_{G,H'}(v_2, u_2), \quad H' \in \mathcal{R}'.$$

We modify the definition (42) of a  $\delta$ -typical graph by taking the maximum over all pairs  $(u, v) \in f \times e$  of distinct vertices. Since we now have two terms in (133) and (134), the bounds in (49) and (50) get some extra factors 2, which cancel out, leading to a bound similar to (51), but with constant 9 inflated.

There is also a little inconvenience related to the analog of formula (56). Namely, now

$$\begin{aligned} \text{cod}_F(u, v) &= |\Gamma_F(u) \setminus \{v\}| + |\Gamma_F(v) \setminus \{u\}| - |\Gamma_F(u) \cup \Gamma_F(v) \setminus \{u, v\}| \\ &= \text{deg}_F(u) + \text{deg}_F(v) - 2\mathbb{I}_{uv \in F} - (n - 2 - \text{cod}_{K_n \setminus F}(u, v)), \end{aligned} \tag{135}$$

so, the formula gets an extra factor  $O(1)$ , which turns out to be negligible.

Set  $\mathbb{R} := \mathbb{R}(n, p)$ . More substantial modifications needed to prove Theorem 1' (which we discuss in detail below) are the following.

- Lemma 11 (only the case  $\mathbb{I} = 0$  left). Instead of using the asymptotic enumeration formula of Canfield, Greenhill and McKay (Theorem 5), we apply the one of Liebenau and Wormald [8].
- Lemma 10 in the case  $p > 0.49$ . Instead of directly showing the existence of short alternating cycles in  $K_n \setminus \mathbb{R}(t) = (K_n \setminus \mathbb{R}) \cup (\mathbb{R} \setminus \mathbb{R}(t))$ , we create a blue-red auxiliary bipartite graph from  $K_n \setminus \mathbb{R}(t)$  and apply unchanged Lemma 20 to it.

**7.1. Sketch of co-degree concentration for regular graphs**

The non-bipartite version of Lemma 11 below is obtained by just setting  $\mathbb{I} = 0$  and replacing  $\hat{n}$  by  $n$  (the resulting term  $\hat{p}$  is swallowed by  $\lambda$ ).

**Lemma 11'.** *Suppose that  $\hat{p}n \rightarrow \infty$  and  $\lambda = \lambda(n) \rightarrow \infty$ . Then, for any distinct  $u, v \in [n]$ ,*

$$\mathbb{P}\left(|\text{cod}_{\mathbb{R}}(u, v) - p^2n| \leq 20\hat{p}\sqrt{\lambda n} + \lambda\right) = O(ne^{-\lambda}). \tag{136}$$

As before, it is sufficient to assume that  $p \leq 1/2$  and thus replace  $\hat{p}$  by  $p$  in (136). Indeed, by (135) and the identity  $2p - 1 = p^2 - q^2$ ,

$$\begin{aligned} \text{cod}_{\mathbb{R}}(u, v) - p^2n &= 2(p(n - 1) - \mathbb{I}_{uv \in \mathbb{R}}) - (n - 2 - \text{cod}_{K_n \setminus \mathbb{R}}(u, v)) - p^2n \\ &= \text{cod}_{K_n \setminus \mathbb{R}}(u_1, v_1) - q^2n + O(1). \end{aligned}$$

This allows, for  $p > 1/2$ , to replace  $p$  by  $q$ , as explained in the proof of Lemma 11 (the error  $O(1)$  is absorbed by the term  $\lambda$ ).

The proof of (136) for  $p \leq 1/2$  is based on the following enumeration result by Liebenau and Wormald [8] (see Cor. 1.5 and Conj. 1.2 therein), proved for some ranges of  $d$  already by McKay and Wormald [10,11].

Given a sequence  $\mathbf{d} = (d_1, \dots, d_n)$ , let  $g(\mathbf{d})$  denote the number of graphs  $G$  on the vertex set  $V = (v_1, \dots, v_n)$  whose degree sequence is  $\mathbf{d}$ , that is,  $\text{deg}_G(v_i) = d_i, i = 1, \dots, n$ . Further, let

$$\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i, \quad \mu = \frac{\bar{d}}{n-1}, \quad \gamma_2 = \frac{1}{(n-1)^2} \sum_{i=1}^n (d_i - \bar{d})^2, \quad \hat{d} = \min(\bar{d}, n-1-\bar{d}).$$

**Theorem 22 ([8]).** *For some absolute constant  $\varepsilon > 0$ , if  $\mathbf{d} = \mathbf{d}(n)$  satisfies  $\max_i |d_i - \bar{d}| = o(n^\varepsilon \hat{d})$ ,  $n\hat{d} \rightarrow \infty$ , as  $n \rightarrow \infty$ , and  $\sum_{i=1}^n d_i$  is even, then*

$$g(\mathbf{d}) \sim \sqrt{2} \exp\left(-\frac{1}{4} - \frac{\gamma_2^2}{4\mu^2(1-\mu)^2}\right) (\mu^\mu(1-\mu)^{1-\mu})^{\binom{n}{2}} \prod_{i=1}^n \binom{n-1}{d_i}.$$

The proof of Lemma 11' follows the lines of the proof of Lemma 11 in the case  $\mathbb{I} = 0$ . For fixed distinct  $u, v \in V$  we define, as before,  $\mathcal{R}_k = \{G \in \mathcal{R}(n, d) : \text{cod}_G(u, v) = k\}$ ,  $0 \leq k \leq d$ . Using Theorem 22, it is tedious but straightforward to find a sequence  $r_k := r_k(n, d)$  such that  $|\mathcal{R}_k| \sim r_k$ . We have  $r_k = r_k^0 + r_k^1$ , as the formula for  $|\mathcal{R}_k|$  breaks into two, according to whether  $uv$  is an edge ( $r_k^1$ ) or not ( $r_k^0$ ). It can be checked that  $r_0, \dots, r_d > 0$  and, for  $1 \leq k \leq d$ ,

$$\frac{r_k^0}{r_{k-1}^0} = \frac{(d - k + 1)^2}{k(n - 2 - 2d + k)} \left(1 - \frac{1}{d}\right) \left(1 - \frac{1}{n - 1 - d}\right)$$

as well as

$$\frac{r_k^1}{r_{k-1}^1} = \frac{(d - k)^2}{k(n - 2d + k)} \left(1 - \frac{1}{d}\right) \left(1 - \frac{1}{n - 1 - d}\right).$$

Let  $\mu^0 := (d + 1)^2/n$  and  $\mu^1 := d^2/n$ . For  $\mu^0 \leq k \leq d$  and, respectively, for  $\mu^1 \leq k \leq d$ ,

$$\frac{r_k^0}{r_{k-1}^0} \leq \frac{\mu^0}{k} \quad \text{and} \quad \frac{r_k^1}{r_{k-1}^1} \leq \frac{\mu^1}{k} \leq \frac{\mu^0}{k}.$$

Thus, for  $\mu^0 \leq k \leq d$ ,

$$\frac{r_k}{r_{k-1}} = \frac{r_k^0 + r_k^1}{r_{k-1}^0 + r_{k-1}^1} \leq \frac{\mu^0}{k}.$$

Similarly, setting  $\rho = (1 - \frac{1}{d}) \left(1 - \frac{1}{n-1-d}\right)$ , for  $1 \leq k \leq \mu^0$  and, respectively,  $1 \leq k \leq \mu^1$ ,

$$\frac{r_{k-1}^0}{r_k^0} \leq \frac{k}{\mu^0 \rho} \leq \frac{k}{\mu^1 \rho} \quad \text{and} \quad \frac{r_{k-1}^1}{r_k^1} \leq \frac{k}{\mu^1 \rho}.$$

So, for  $1 \leq k \leq \mu^1$ ,

$$\frac{r_{k-1}}{r_k} = \frac{r_{k-1}^0 + r_{k-1}^1}{r_k^0 + r_k^1} \leq \frac{k}{\mu^1 \rho}.$$

Setting conveniently  $\mu_+ := \mu^0$  and  $\mu_- := \mu^1 \rho$ , we may now apply Claim 12, which extends straightforwardly to the non-bipartite setting.

We have, using that  $d \geq 1$ ,

$$\mu_+ \leq \frac{4d^2}{n} \leq 4p^2 n. \tag{137}$$

Also, using  $p \leq 1/2$ ,

$$\mu_+ \leq \frac{(pn + 1)^2}{n} \leq p^2 n + 2 \quad \text{and} \quad \mu_- = \mu^1 \rho \geq \frac{d^2}{n} \left(1 - \frac{1}{d}\right)^2 \geq \frac{(pn - 2)^2}{n} \geq p^2 n - 2,$$

whence, setting  $X = \text{cod}_{\mathbb{R}}(u_1, v_1)$ , and  $x = 20\sqrt{p^2 n \lambda} + \lambda - 2$ ,

$$\mathbb{P}\left(|X - p^2 n| \geq 20p\sqrt{\lambda n} + \lambda\right) \leq \mathbb{P}(X \geq \mu_+ + x) + \mathbb{P}(X \leq \mu_- - x). \tag{138}$$

We now bound the RHS of (138) using Claim 12 with  $\lambda - 2$  instead of  $\lambda$ . Noting that (137) implies  $x \geq \sqrt{2\mu_+(\lambda - 2)} + (\lambda - 2)$ , we have

$$\mathbb{P}(X \geq \mu_+ + x) + \mathbb{P}(X \leq \mu_- - x) = O\left(\sqrt{N}e^{-(\lambda-2)}\right).$$

Since  $\sqrt{N}e^{-(\lambda-2)} = \Theta(ne^{-\lambda})$ , this completes the proof of Lemma 11'.

**7.2. Sketch of the proof of new Lemma 10, case  $p > 0.49$**

For completeness, we state here the non-bipartite counterpart of Lemma 10 which on the surface looks almost identical.

**Lemma 10'.** *Assuming (131), we have, a.a.s.,*

$$\min_{e,f \in K_n \setminus \mathbb{R}(t), e \neq f} \mathbb{P}(e \in \mathbb{R}, f \notin \mathbb{R} \mid \mathbb{R}(t)) \geq e^{-\lambda(t)}, \quad \text{for } t \leq t_0.$$

Looking at the diagram in Figure 1, we see that the proof of Lemma 10 relies on several other results, most notably Lemmas 17 and 20. It would be, however, a very tedious task to come up with non-bipartite counterpart of Lemma 20, and, consequently, ones of Lemmas 19 and 18. Instead, we rather convert the non-bipartite case into the bipartite one by a standard probabilistic construction and use Lemmas 17 and 20 practically unchanged.

Given a blue-red graph  $H \subseteq K_n$ , let  $\text{bip}(H)$  be a random blue-red bipartite graph with bipartition  $V_1 = \{u_1, \dots, u_n\}$ ,  $V_2 = \{v_1, \dots, v_n\}$ , such that for each edge  $ij \in E(H)$  we flip a fair coin and include into  $\text{bip}(H)$  either  $u_i v_j$  or  $u_j v_i$  (coloured the same colour as  $ij$ ), with the flips being independent. In particular,  $|E(\text{bip}(H))| = |E(H)|$ , so the density of  $\text{bip}(H)$  is exactly half of that of  $H$ , while, if the densities of the blue and red subgraphs of  $H$  are  $b$  and  $r$ , then the expected densities of the blue and red subgraphs of  $\text{bip}(H)$  are  $b/2$  and  $r/2$ .

Note that if there is an instance of  $\text{bip}(H)$  in which an edge  $u_i v_j$  is contained in an alternating cycle of length at most  $D$ , then  $ij$  is contained in an alternating circuit of the same length. It is not, in general, a cycle, since some vertices can be repeated, but edges are not, as the edges of  $\text{bip}(H)$  correspond to different edges of  $H$ . Luckily, Proposition 4 actually works for alternating circuits too, since in a circuit the blue and red degrees of each vertex equal each other (see the paragraph following equation (15)) and, similarly as for cycles, in  $K_n$  there are at most  $n^{2\ell-2}$  circuits of length  $\ell$  containing a given edge  $e$ . Defining  $\mathcal{R}_G$ ,  $\mathcal{R}_{G,e}$ , and  $\mathcal{R}_{G,-e}$  analogously to the bipartite case (cf. (15) and (16)), and making obvious modifications of the proof of Proposition 4, we obtain the following.

**Proposition 4'.** *Let a graph  $G \subseteq K_n$  be such that  $\mathcal{R}_G \neq \emptyset$  and let  $e \in K_n \setminus G$ . Assume that for some number  $D > 0$  and every  $H \in \mathcal{R}_G$  the edge  $e$  is contained in an alternating circuit of length at most  $2D$ . Then  $\mathcal{R}_{G,-e} \neq \emptyset$ ,  $\mathcal{R}_{G,e} \neq \emptyset$ , and*

$$\frac{1}{n^{2D} - 1} \leq \frac{|\mathcal{R}_{G,-e}|}{|\mathcal{R}_{G,e}|} \leq n^{2D} - 1.$$

The plan to adapt the proof of Lemma 10 is to condition on  $K_n \setminus \mathbb{R}(t)$  having its degrees and co-degrees concentrated (using the non-bipartite counterparts of Lemmas 13 and 14) and then show that there is an instance  $F$  of  $\text{bip}(K_n \setminus \mathbb{R}(t))$  in which the degrees and co-degrees are similarly concentrated, with just a negligibly larger error.

In particular, such an  $F$  is  $(q/2, \tau p/2, \alpha/32)$ -regular (as before, we denote  $\alpha = \min\{\tau p, q\}$ ). Then, applying Theorem 15 along the lines of the proof of Lemma 17, we show that  $F$  is also  $(\pi/2, \alpha/32)$ -jumbled, where, as before,  $\pi = \tau p + q$ . Hence, we are in position to apply Lemma 20, obtaining for every edge of  $F$  an alternating cycle of length  $O(1/(\tau p q))$  in  $F$ . As explained above, this implies alternating circuits in  $K_n \setminus \mathbb{R}(t)$  of the same length, and so we may complete the proof of Lemma 10', based on Proposition 4', in the way we did it in the bipartite case.

Let us now present some more details. We condition on  $\mathbb{R}(t) = G$  such that (cf. (75) and (78))

$$\forall v \in [n] \quad |\deg_G(v) - (1 - \tau)pn| = O(\sqrt{n \log n}) \tag{139}$$

and

$$\max_{u,v \in [n], u \neq v} \text{cod}_G(u, v) = (1 - \tau)^2 p^2 n + O(\sqrt{n \log n}). \tag{140}$$

Fix an arbitrary  $H \in \mathcal{R}_G$ . Since the red graph  $K_n \setminus H$  is  $q(n-1)$ -regular, we have  $\deg_{\text{bip}(K_n \setminus H)}(v) \sim \text{Bin}(q(n-1), 1/2)$ . Thus, by a routine application of the Chernoff bound a.a.s.

$$\max_{v \in [n]} \left| \deg_{\text{bip}(K_n \setminus H)}(v) - qn/2 \right| = O(\sqrt{n \log n}). \tag{141}$$

By (139), the blue graph  $H \setminus G$  has degrees  $pn - (1 - \tau)pn + O(\delta\tau pn) = \tau pn + O(\sqrt{n \log n})$ , so the Chernoff bound implies that (using  $p > 0.49$ ) a.a.s.

$$\max_{v \in [n]} \left| \deg_{\text{bip}(H \setminus G)}(v) - \tau pn/2 \right| = O(\sqrt{n \log n}). \tag{142}$$

By (131) and (132), since  $p > 0.49$  and  $\tau \geq \tau_0$ ,

$$\alpha = \min\{\tau p, q\} \gg \left(\frac{\log n}{n}\right)^{1/4}.$$

Consequently,  $\sqrt{n \log n} \ll \alpha n$  and, with a big margin, we conclude that a.a.s.  $\text{bip}(K_n \setminus G)$  is  $(q/2, \tau p/2, \alpha/32)$ -regular.

Next, we check that  $\text{bip}(K_n \setminus G)$  is  $(\pi/2, \alpha/32)$ -jumbled. Inequalities (141) and (142) imply

$$\deg_{\text{bip}(K_n \setminus G)}(v) = \frac{\pi n}{2} + O(\sqrt{n \log n}). \tag{143}$$

Further, by (135), (139), and (140), we have that (cf. (100))

$$\begin{aligned} \max_{u \neq v} \text{cod}_{K_n \setminus G}(u, v) &\leq 2 \max_{v \in [n]} \deg_{K_n \setminus G}(v) - (n - 2 - \text{cod}_G(u, v)) \\ &= (2\pi - 1)n + (1 - \pi)^2 n + O(\sqrt{n \log n}) = \pi^2 n + O(\sqrt{n \log n}). \end{aligned}$$

Since for every  $u, v \in [n], u \neq v$ ,  $\text{cod}_{\text{bip}(K_n \setminus G)}(u, v) \sim \text{Bin}(\text{cod}_{K_n \setminus G}(u, v), \frac{1}{4})$ , by a simple application of Chernoff's inequality and the union bound we show that a.a.s.

$$\max_{u, v \in V_1} \text{cod}_{\text{bip}(K_n \setminus G)}(u, v) = \frac{\pi^2 n}{4} + O(\sqrt{n \log n}). \tag{144}$$

Now, fix an instance  $F$  of the graph  $\text{bip}(K_n \setminus G)$  for which (141), (142) and (144) hold.

Applying Theorem 15, by calculations similar to those in the proof of Lemma 17, from (143) and (144) we deduce that  $F$  is  $(\pi/2, \alpha/32)$ -jumbled. Since we earlier showed that  $F$  is  $(q/2, \tau p/2, \alpha/32)$ -regular, Lemma 20 implies that in  $F$  every edge is contained in an alternating cycle of length  $O(1/(\tau pq))$  and therefore in  $K_n \setminus G$  every edge is contained in an alternating circuit of the same length. The same argument implies alternating cycles in  $K_n \setminus (G \cup \{e\})$  for an arbitrary edge  $e$ , since only the lower bound in (139) has to be decreased by a negligible quantity 1.

The rest of the proof of Lemma 10' in the case  $p > 0.49$  follows along similar lines.

### 8. Technical facts

Here we collect a few technical or very plausible facts with their easy proofs. Most of them have been already utilised in the paper. An exception is Proposition 24 to be used only in Remark 26, Section 10.

We begin with convenient consequences of the assumptions of Lemma 6.

**Proposition 23.** For  $t = 0, \dots, t_0 - 1$ , the conditions of Lemma 6, namely, (26), (27), and (28), imply that

$$\tau q \geq \tau_0 q \geq 700 \cdot 680 \sqrt{\frac{3(C+4) \log N}{\hat{n}}}, \quad \text{whenever } p > 0.49, \tag{145}$$

$$\gamma_t \leq \gamma_{t_0} \leq 1, \tag{146}$$

and

$$\tau \hat{p} \hat{n} \geq \tau_0 \hat{p} \hat{n} \geq 3000^2 (C+4) \log N. \tag{147}$$

**Proof.** Since  $\tau = \tau(t) \geq \tau(t_0) \geq \tau_0$  and  $\gamma_t$  is increasing in  $t$ , the first inequalities in (145)–(147) are immediate and it is enough to prove the second inequalities.

Inequality (145) follows from the definition (25) of  $\tau_0$  and (28).

To show (146), for  $p \leq 0.49$ , using the definition of  $\tau_0$  (see (25)) and (27), we get

$$\gamma_{t_0} \leq \frac{1080}{340^4} + \sqrt{\frac{2}{3}} < 0.01 + 0.82 < 1,$$

while for  $p > 0.49$

$$\gamma_{t_0} \leq 0.01 + \frac{25,000}{680 \cdot 700} < 1.$$

To see (147) first note that for  $p \leq 0.49$  it is straightforward from the definition of  $\tau_0$  in (25). For  $p > 0.49$ , observing that (26) implies  $\sqrt{\frac{3(C+4) \log N}{\hat{n}}} \leq \sqrt{\hat{p}/3240} \leq 1/3240$ , we argue that

$$\tau_0 \hat{p} \geq \frac{49}{51} \tau_0 q \stackrel{(145)}{\geq} \frac{49}{51} \cdot 700 \cdot 680 \sqrt{\frac{3(C+4) \log N}{\hat{n}}} \geq 3000^2 \frac{(C+4) \log N}{\hat{n}},$$

whence (147) follows. □

Next, we give a proof of Claim 7 which was instrumental in deducing Theorem 2 from Lemma 6 in Section 3.2.

**Proof of Claim 7.** Writing  $X = t_0 - |S|$ , we have  $\mathbb{E} X = \sum_{t=0}^{t_0-1} \gamma_t$ . Denoting  $\alpha := 1080 \hat{p}^2 \mathbb{I}$  and

$$\beta := \begin{cases} 3240 \sqrt{\frac{2(C+3) \log N}{p \hat{n}}}, & p \leq 0.49, \\ 25,000 \sqrt{\frac{(C+3) \log N}{q^2 \hat{n}}}, & p > 0.49, \end{cases}$$

we have

$$\gamma_t = \alpha + \begin{cases} \beta \tau^{-1/2}, & p \leq 0.49, \\ \beta \tau^{-1}, & p > 0.49. \end{cases}$$

Since, trivially,  $\sum_{t=0}^{t_0-1} \alpha = \alpha t_0 \leq \alpha M = 1080 \hat{p}^2 \mathbb{I} \cdot M$ , to prove (35), it suffices to show that

$$\sum_{t=0}^{t_0-1} \tau^{-1/2} \leq 2M, \quad \text{and} \quad \sum_{t=0}^{t_0-1} \tau^{-1} \leq \frac{M}{4} \log \frac{\hat{n}}{\log N}. \tag{148}$$

Since  $\tau = \tau(t) = 1 - t/M$  is positive and decreasing on  $[0, M)$ ,

$$\sum_{t=0}^{t_0-1} \tau^{-1/2} \leq \sum_{t=0}^{M-1} \tau^{-1/2} \leq \int_0^M \tau^{-1/2} dt = M \int_0^1 \tau^{-1/2} d\tau = 2M,$$

implying the first inequality in (148). On the other hand, recalling  $t_0 = \lfloor (1 - \tau_0)M \rfloor$ ,

$$\sum_{t=0}^{t_0-1} \tau^{-1} \leq \int_0^{t_0} \tau^{-1} dt \leq \int_0^{(1-\tau_0)M} \tau^{-1} dt \leq M \int_{\tau_0}^1 \tau^{-1} d\tau = M \log \frac{1}{\tau_0} \leq \frac{M}{4} \log \frac{\hat{n}}{\log N},$$

which implies the second inequality in (148).

Checking (36) is a dull inspection of definitions using conditions (28)–(29), which the readers might prefer to do themselves. We nevertheless spell out the details starting with the case  $p \leq 0.49$ . Choosing  $C^*$  large enough,

$$\begin{aligned} \gamma &= C^* \left( p^2 \mathbb{I} + \sqrt{\frac{\log N}{p \hat{n}}} \right) \\ &\geq 1080 p^2 \mathbb{I} + 6480 \sqrt{\frac{(C+3) \log N}{p \hat{n}}} + 3240 \sqrt{\frac{3(C+4) \log N}{p \hat{n}}} + \sqrt{\frac{\log N}{p \hat{n}}} + \sqrt{\frac{2C \log N}{p \hat{n}}} \\ &\geq \theta + \sqrt{\tau_0} + 2\sqrt{1/M} + \sqrt{\frac{2C \log N}{M}} \stackrel{(29)}{\geq} \theta + \tau_0 + 2/M + \sqrt{\frac{2C \log N}{M}}. \end{aligned}$$

In the case  $p > 0.49$ , note that  $q^{3/2} \geq q^2 \geq \hat{p}^2$ . Therefore, assuming  $C^*$  is large enough,

$$\gamma/3 \geq 700 (3(C+4))^{1/4} \left( q^{3/2} \mathbb{I} + \left( \frac{\log N}{\hat{n}} \right)^{1/4} \right) = \tau_0,$$

and

$$\gamma/3 \geq 1080 q^{3/2} \mathbb{I} + 6250 \sqrt{\frac{(C+3) \log N}{q^2 \hat{n}}} \log \frac{\hat{n}}{\log N} \geq \theta.$$

Finally, because (28) implies  $\hat{n}/(\log N) \geq 1$  and  $d_1, d_2 \geq 1$  implies  $M \geq \hat{n}$ , for large enough  $C^*$ ,

$$\begin{aligned} \gamma/3 &\geq \frac{C^*}{3} \left( \frac{\log N}{\hat{n}} \right)^{1/4} \geq \frac{\log N}{\hat{n}} + \sqrt{\frac{2C \log N}{\hat{n}}} \\ &\stackrel{\boxed{p > 0.49, M \geq \hat{n}}}{\geq} \frac{2}{M} + \sqrt{\frac{2C \log N}{M}}, \end{aligned}$$

which, with the previous two inequalities, implies (36). □

We conclude this technical section with a lower bound on the maximum degree in a binomial random graph.

**Proposition 24.** *If  $p' \leq 1/4$  and  $n_2 \leq n_1$ , then a.a.s. the maximum degree of  $\mathbb{G}(n_1, n_2, p')$  in  $V_1$  is at least  $\kappa := \min\{\lceil p'n_2 + \sqrt{p'(1-p')n_2 \log n_1} \rceil, n_2\}$ .*

**Proof.** Writing  $Z \sim \mathcal{N}(n_2 p', n_2 p'(1-p'))$ , by Slud’s inequality [13, Theorem 2.1] we have

$$r := \mathbb{P}(\text{Bin}(n_2, p') \geq \kappa) \geq \mathbb{P}(Z \geq \kappa) \geq \mathbb{P}\left(Z \geq \sqrt{p'(1-p')n_2 \log n_1}\right),$$

hence by a standard approximation of the normal tail we obtain

$$r \geq \frac{1 + o(1)}{\sqrt{2\pi \log n_1}} e^{-\frac{1}{2} \log n_1} = e^{-(1/2+o(1)) \log n_1}.$$

Therefore, the probability that all vertices in  $V_1$  have degrees smaller than  $\kappa$  is

$$(1 - r)^{n_1} \leq e^{-n_1 r} \leq \exp\left(-n_1 e^{-(1/2+o(1)) \log n_1}\right) \rightarrow 0,$$

proving the proposition. □

### 9. Application: perfect matchings between subsets of vertices

Perarnau and Petridis in [12], in connection with a problem of Plünnecke, studied the existence of perfect matchings between fixed subsets of vertices in random biregular graphs. In particular, they proved the following result (which we state in our notation to make it easier to apply Theorem 2).

**Theorem 25.** (Theorem 2 in [12]) *Let  $k > 0$  be a constant, and assume  $n_2 = kn_1$  and  $pn_2 \leq n_1$ . Take subsets  $A \subseteq V_1$  and  $B \subseteq V_2$  of size  $pn_2$  and let  $\mathcal{M}_{A,B}$  denote the event that the subgraph of  $\mathbb{R}(n_1, n_2, p)$  induced by  $A$  and  $B$  contains a perfect matching. As  $\hat{n} \rightarrow \infty$ , we have*

$$\mathbb{P}(\mathcal{M}_{A,B}) \rightarrow 0, \quad \text{if } p^2 n_2 - \log pn_2 \rightarrow -\infty \text{ or } pn_2 \text{ is constant,}$$

and

$$\mathbb{P}(\mathcal{M}_{A,B}) \rightarrow 1, \quad \text{if } p^2 n_2 - \log pn_2 \rightarrow \infty. \tag{149}$$

Perarnau and Petridis [12] speculated that if the bipartite version of the Sandwich Conjecture was true, then Theorem 25 would follow straightforwardly from the classical result of Erdős and Rényi on perfect matchings in the random bipartite graph (see Theorem 4.1 in [5]). That result, in particular, implies that the bipartite binomial random graph  $\mathbb{G}(n', n', p')$  contains a perfect matching a.a.s. whenever  $p'n' - \log n' \rightarrow \infty$  as  $n' \rightarrow \infty$ . We show how this, together with Theorem 2, implies the 1-statement in (149), provided that condition (4) of Theorem 2 is satisfied, which in particular implies

$$q \geq \left(\frac{\log N}{\hat{n}}\right)^{1/4}. \tag{150}$$

By Theorem 2, the random graph  $\mathbb{R}(n_1, n_2, p)$  a.a.s. contains a random graph  $\mathbb{G}(n_1, n_2, p')$  with

$$p' = (1 - 2\gamma)p,$$

where  $\gamma$  is defined in (5). In particular, the subgraph of  $\mathbb{R}(n_1, n_2, p)$  induced by  $A$  and  $B$  contains a random graph  $\mathbb{G}(pn_2, pn_2, p')$ . To see that the latter random graph contains a perfect matching a.a.s., let us verify the Erdős–Rényi condition

$$p'pn_2 - \log pn_2 \rightarrow \infty.$$

From the assumption in (149) it follows that

$$p \geq n_2^{-1/2} = \Theta(\hat{n}^{-1/2}). \tag{151}$$

For our purposes, it is sufficient to check that

$$\gamma \ll (\log N)^{-1}, \tag{152}$$

since then

$$\gamma \log pn_2 \leq \gamma \log N \rightarrow 0, \quad \text{and } \gamma \rightarrow 0,$$

which together with the condition in (149) imply

$$p'pn_2 - \log pn_2 = (1 - 2\gamma)(p^2n_2 - \log pn_2) - 2\gamma \log pn_2 \rightarrow \infty.$$

To see that (152) holds, first note that  $\mathbb{I} = 1$  implies  $p = O((\log N)^{-1})$  for  $p \leq 0.49$  and  $q = O((\log N)^{-1})$  for  $p > 0.49$ ; hence, regardless of  $p$ , the first term in the definition of  $\gamma$  is  $o((\log N)^{-1})$ . The remaining terms are much smaller: for  $p \leq 0.49$  inequality (151) implies that the second term in the definition of  $\gamma$  is  $O((\log N)^{1/2}\hat{n}^{-1/4}) \ll (\log N)^{-1}$ , while in the case  $p > 0.49$  assumption (150) implies that the last two terms in the definition of  $\gamma$  are at most  $(\log N)^{O(1)}\hat{n}^{-1/4} \ll (\log N)^{-1}$ .

### 10. Concluding remarks

**Remark 26.** Assume  $p \leq 1/4$ . If a.a.s.  $\mathbb{G}(n_1, n_2, p') \subseteq \mathbb{R}(n_1, n_2, p)$ , then we must have

$$p' = p \left( 1 - \Omega \left( \min \left\{ \sqrt{(\log N)/(p\hat{n})}, 1 \right\} \right) \right). \tag{153}$$

To see this, assume, without loss of generality, that  $n_2 \leq n_1$ , and note that by Proposition 24 we must have  $p'n_2 + \sqrt{p'(1-p')n_2 \log n_1} \leq pn_2$ , and therefore  $p/p' \geq 1 + \sqrt{\frac{(1-p') \log n_1}{p'n_2}}$ . Since  $p' \leq p \leq 1/4$  and  $n_2 \leq n_1$ , we have  $p/p' = 1 + \Omega \left( \sqrt{\frac{\log N}{p\hat{n}}} \right)$ , whence (153) follows.

**Remark 27.** In view of Remark 26, the error  $\gamma$  in Theorem 2 has optimal order of magnitude, whenever  $p \leq 1/4$ , provided that also  $p^2\mathbb{I} = O \left( \sqrt{\log N/p\hat{n}} \right)$  (that is, if either  $p = O \left( ((\log N)/\hat{n})^{1/5} \right)$  or  $\mathbb{I} = 0$ ).

From Remark 26 it also follows that we cannot have  $\gamma = o(1)$  for  $\log N = \Omega(p\hat{n})$ . Theorem 2 does not apply to the case  $\log N \geq p\hat{n}/(C^*)^2$ , but we think it would be interesting to find the largest  $p'$  for which one can a.a.s. embed  $\mathbb{G}(n_1, n_2, p')$  into  $\mathbb{R}(n_1, n_2, p)$  even in this case. The tightest embedding one can expect is the one permitted by the maximum degree. We conjecture that such an embedding is possible.

**Conjecture 1.** Suppose that a sequence of parameters  $(n_1, n_2, p, p')$  is such that a.a.s. in  $\mathbb{G}(n_1, n_2, p')$  the maximum degree over  $V_1$  is at most  $d_1 = n_2p$  and the maximum degree over  $V_2$  is at most  $d_2 = n_1p$ . There is a joint distribution of  $\mathbb{G}(n_1, n_2, p')$  and  $\mathbb{R}(n_1, n_2, p)$  such that

$$\mathbb{G}(n_1, n_2, p') \subseteq \mathbb{R}(n_1, n_2, p) \quad \text{a.a.s.}$$

Note that if Conjecture 1 is true, then by taking complements we also have the tightest embedding  $\mathbb{R}(n_1, n_2, p) \subseteq \mathbb{G}(n_1, n_2, p')$  that the minimum degrees of  $\mathbb{G}(n_1, n_2, p')$  permit.

We also propose the following strengthening of the Kim–Vu Sandwich Conjecture.

**Conjecture 2.** Suppose that a sequence of parameters  $(n, p, p')$  is such that a.a.s. in  $\mathbb{G}(n, p')$  the maximum degree is at most  $d = (n - 1)p$ . There is a joint distribution of  $\mathbb{G}(n, p')$  and  $\mathbb{R}(n, p)$  such that

$$\mathbb{G}(n, p') \subseteq \mathbb{R}(n, p) \quad \text{a.a.s.}$$

**Remark 28.** For constant  $p$ , to obtain  $\gamma = o(1)$  in Theorem 2, we need to assume  $\mathbb{I} = 0$ , which requires a rather restricted ratio  $n_1/n_2$ . For example, one cannot afford  $n_1 = n_2^{1+\delta}$  for any constant  $\delta > 0$ . This restriction comes from an enumeration result we use in the proof, namely, Theorem 5 (see condition (iii) therein). Should enumeration be proven with a relaxed condition, it would automatically improve our Theorem 2.

**Remark 29.** The terms  $p^2\mathbb{I}$  and  $q^{3/2}\mathbb{I}$  in (5) are artifacts of the application of switchings in Lemma 11. In the sparse case ( $\hat{p} \rightarrow 0$ ) it is plausible that the condition  $\mathbb{I} = 0$  can be made much milder by using a very recent enumeration result of Liebenau and Wormald [9] instead of Theorem 5. Due to the schedule of this manuscript, we did not check what this would imply, but readers seeking smaller errors in (5) are encouraged to do so.

## Acknowledgement

We thank Noga Alon and Benny Sudakov who, at the conference Random Structures & Algorithms 2019, suggested a way to show the existence of alternating paths in non-bipartite graphs. We are also thankful to the anonymous referee for useful remarks.

## References

- [1] Canfield, E. R., Greenhill, C. and McKay, B. D. (2008) Asymptotic enumeration of dense 0-1 matrices with specified line sums. *J. Comb. Th. A* **115** 32–66.
- [2] Dudek, A., Frieze, A., Ruciński, A. and Šileikis, M. (2017) Embedding the Erdős-Rényi hypergraph into the random regular hypergraph and hamiltonicity. *J. Comb. Theory B* **122** 719–740.
- [3] Gao, P. Kim-Vu’s sandwich conjecture is true for all  $d = \Omega(\log^7 n)$ . [arXiv:2011.09449](https://arxiv.org/abs/2011.09449).
- [4] Gao, P., Isaev, M. and McKay, B. (2020) Sandwiching random regular graphs between binomial random graphs. In *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms*, pp. 690–701.
- [5] Janson, S., Łuczak, T. and Ruciński, A. (2000) *Random Graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York.
- [6] Kim, J. H. and Vu, V. H. (2004) Sandwiching random graphs: universality between random graph models. *Adv. Math.* **188**(2) 444–469.
- [7] Krivelevich, M., Sudakov, B., Vu, V. H. and Wormald, N. (2001) Random regular graphs of high degree. *Random Struct. Alg.* **18** 346–363.
- [8] Liebenau, A. and Wormald, N. Asymptotic enumeration of graphs by degree sequence, and the degree sequence of a random graph. [arXiv:1702.08373](https://arxiv.org/abs/1702.08373).
- [9] Liebenau, A. and Wormald, N. Asymptotic enumeration of digraphs and bipartite graphs by degree sequence. [arXiv:2006.15797](https://arxiv.org/abs/2006.15797).
- [10] McKay, B. D. and Wormald, N. C. (1990) Asymptotic enumeration by degree sequence of graphs of high degree. *Eur. J. Comb.* **11** 565–580.
- [11] McKay, B. D. and Wormald, N. C. (1991) Asymptotic enumeration by degree sequence of graphs with degrees  $o(n^{1/2})$ . *Combinatorica* **11** 369–382.
- [12] Perarnau, G. and Petridis, G. (2013) Matchings in random biregular bipartite graphs. *Electron. J. Combin.* **20**(1) #P60 1.
- [13] Slud, E. V. (1977) Distribution inequalities for the binomial law *Ann. Prob.* **5**(3) 404–412.
- [14] Thomason, A. (1989) Dense expanders and pseudo-random bipartite graphs. *Discrete Math.* **75** 381–386.