Hamiltonian decompositions of random bipartite regular graphs

Catherine Greenhill*
Department of Mathematics and Statistics,
University of Melbourne
Parkville, VIC 3010, Australia
csg@ms.unimelb.edu.au

Jeong Han Kim†
Microsoft Research,
One Microsoft Way,
Redmond, WA 98052, USA
jehkim@microsoft.com

Nicholas C. Wormald‡
Department of Mathematics and Statistics,
University of Melbourne
Parkville, VIC 3010, Australia
nick@ms.unimelb.edu.au

12th March 2002

Abstract

We prove a complete hamiltonian decomposition theorem for random bipartite regular graphs, thereby verifying a conjecture of Robinson and Wormald. The main step is to prove contiguity (a kind of asymptotic equivalence) of two probabilistic models of 4-regular bipartite graphs; namely, the uniform model, and the model obtained by taking the union of two independent, uniformly chosen bipartite Hamilton cycles, conditioned on forming no multiple edges. The proof uses the small subgraph conditioning method to establish contiguity, while the differential equation method is used to analyse a critical quantity.

1 Introduction

For positive integers n and d, let $\mathcal{B}_{n,d}$ denote the probability space of random bicoloured d-regular graphs on $2n$ labelled vertices, where each graph occurs with equal probability. These are bipartite graphs with a fixed proper 2-colouring, or *bicolouring*, of the vertices.

*Research supported by an Australian Research Council Postdoctoral Fellowship.
†Research performed in part while visiting the University of Melbourne.
‡Research supported in part by the Australian Research Council.

For definiteness, let vertices $\{1, \dots, n\}$ have one colour and vertices $\{n+1, \dots, 2n\}$ have the second colour. By a slight abuse of terminology, we refer to $\mathcal{B}_{n,d}$ as a random bipartite graph.

Now suppose that d is a fixed positive integer with $d \geq 4$. In this paper we prove a conjecture of Robinson and Wormald [10]; namely, that the probability that a random d-regular bipartite graph on $2n$ vertices has $\lfloor d/2 \rfloor$ pairwise edge-disjoint Hamilton cycles tends to 1 as $n \to \infty$. Such a partition of the edges of a graph into disjoint unions of Hamilton cycles and at most one perfect matching is called a *hamiltonian decomposition* of the graph. (The corresponding result for arbitrary d-regular graphs was proven by Frieze et al. [1] and Kim and Wormald [5]. We extend this proof, but the extension required substantial new arguments. This contrasts with the proof of hamilitonicity of almost all cubic graphs, which was distinctly easier in the bipartite case.)

Presumably the Robinson–Wormald conjecture also holds when $d = d(n) \to \infty$ as $n \to \infty$. However, the current methods do not extend to this situation in general.

Our main result is best stated in terms of contiguity of models. Let $(\mathcal{G}_n)_{n\geq 1}$ and $(\widehat{\mathcal{G}}_n)_{n\geq 1}$ be two sequences of probability spaces such that $\mathcal{G}_n$ and $\widehat{\mathcal{G}}_n$ have the same underlying set Ω_n and differ only in the probabilities, for $n \geq 1$. We say that these sequences are *contiguous* if, for any sequence of events $(A_n)_{n\geq 1}$ where $A_n \subseteq \Omega_n$ for $n \geq 1$, we have

$$\lim_{n\to\infty} \mathbf{P}_{\mathcal{G}_n}(A_n) = 1 \qquad \text{if and only if} \qquad \lim_{n\to\infty} \mathbf{P}_{\widehat{\mathcal{G}}_n}(A_n) = 1.$$

In other words, an event A_n is a.a.s. true in $\mathcal{G}_n$ if and only if it is a.a.s. true in $\widehat{\mathcal{G}}_n$. (Here a.a.s. means "asymptotically almost surely", i.e. with probability tending to 1 as $n \to \infty$.) If $(\mathcal{G}_n)$ and $(\widehat{\mathcal{G}}_n)$ are contiguous then we write

$$\mathcal{G}_n \approx \widehat{\mathcal{G}}_n.$$

Further, if $\mathcal{G}_n$ and $\widehat{\mathcal{G}}_n$ are both probability spaces of random graphs or pseudographs on the same vertex set, define their *sum* $\mathcal{G}_n + \widehat{\mathcal{G}}_n$ to be the space whose elements are defined by the random pseudographs $G \cup \widehat{G}$, where $G \in \mathcal{G}$ and $\widehat{G} \in \widehat{\mathcal{G}}$ are generated independently. Similarly, define their *graph-restricted sum* $\mathcal{G}_n \oplus \widehat{\mathcal{G}}_n$ to be the space which is the restriction of $\mathcal{G}_n + \widehat{\mathcal{G}}_n$ to simple graphs (i.e. those with no loops or multiple edges). For a survey of known contiguity results for sums of graph models, see [12] or [4, Chapter 9].

Write $k\mathcal{G}_n$ for $\mathcal{G}_n \oplus \cdots \oplus \mathcal{G}_n$ where there are k terms in the right hand side summation ($k \geq 2$). Let $\mathcal{H}_n$ denote the uniform probability space of bicoloured Hamilton cycles on $2n$ vertices, with the fixed bipartition. The main result of this paper shows the following.

Theorem 1.1

$$\mathcal{B}_{n,4} \approx 2\mathcal{H}_n.$$

This result was the final fact required to prove the conjecture of Robinson and Wormald, as follows.

Corollary 1.1 *For $d \geq 3$ we have*

$$\mathcal{B}_{n,d} \approx \begin{cases} \frac{d}{2}\mathcal{H}_n & \text{if } d \text{ is even,} \\ \frac{d-1}{2}\mathcal{H}_n \oplus \mathcal{B}_{n,1} & \text{if } d \text{ is odd.} \end{cases}$$

Proof. First note (as is well known) that

$$\mathcal{B}_{n,d} \approx d\mathcal{B}_{n,1} \tag{1}$$

for $d \geq 3$. This follows since

$$\mathcal{B}_{n,d} \approx \mathcal{B}_{n,3} \oplus (d-3)\mathcal{B}_{n,1} \approx 3\mathcal{B}_{n,1} \oplus (d-3)\mathcal{B}_{n,1} = d\mathcal{B}_{n,1}.$$

Here the first contiguity result follows from the proof of [10, Theorem 3] and the second result can be found in [7, Theorem 4]. That the results can be combined follows from basic properties of contiguity, see [12, Section 4.3]. Similarly,

$$\mathcal{B}_{n,d} \approx \mathcal{H}_n \oplus \mathcal{B}_{n,d-2} \tag{2}$$

for $d \geq 5$. This follows from the chain of results

$$\begin{aligned} \mathcal{B}_{n,d} &\approx \mathcal{B}_{n,3} \oplus (d-3)\mathcal{B}_{n,1} \\ &\approx \mathcal{B}_{n,1} \oplus \mathcal{H}_n \oplus (d-3)\mathcal{B}_{n,1} \\ &= \mathcal{H}_n \oplus (d-2)\mathcal{B}_{n,1} \\ &\approx \mathcal{H}_n \oplus \mathcal{B}_{n,d-2}. \end{aligned}$$

The first line follows from the proof of [10, Theorem 3], the second line from the proof of [8, Theorem 2.3] and the last line follows by applying (1) in reverse.

If d is even then the theorem follows by repeatedly applying (2) if $d \geq 6$, and using Theorem 1.1 to conclude that

$$\mathcal{B}_{n,d} \approx \frac{d-4}{2}\mathcal{H}_n \oplus \mathcal{B}_{n,4} \approx \frac{d}{2}\mathcal{H}_n.$$

If $d = 3$ then the result can be found in [8], as above. If d is odd and $d \geq 5$ then

$$\begin{aligned} \mathcal{B}_{n,d} &\approx d\mathcal{B}_{n,1} \\ &= \mathcal{B}_{n,1} \oplus (d-1)\mathcal{B}_{n,1} \\ &\approx \mathcal{B}_{n,1} \oplus \mathcal{B}_{n,d-1} \\ &\approx \mathcal{B}_{n,1} \oplus \frac{d-1}{2}\mathcal{H}_n. \end{aligned}$$

Here the first two contiguity results follow from (1) and the final line follows from the argument for d even, given above. ∎

We will use the small subgraph conditioning method of Robinson and Wormald [9, 10] (see [12]) to prove Theorem 1.1. This method requires the computation of two constants, one relating directly to variance and the other related to a conditional distribution depending on short cycles. The (apparently miraculous) equality of these constants implies the desired result. Before stating the theorem we introduce some notation. Let $\mathcal{G}$ be a probability space with underlying set Ω. Given any nonnegative random variable Y on $\mathcal{G}$, denote by $\mathcal{G}^{(Y)}$ the probability space with underlying set Ω and probabilities given by

$$\mathbf{P}_{\mathcal{G}^{(Y)}}(X) = \frac{Y(X)\mathbf{P}_{\mathcal{G}}(X)}{Z}$$

for all $X \in \Omega$, where $Z = \sum_{X\in\Omega} Y(X)$ is the normalising constant. The notation $[X]_k$ denotes the falling factorial, $[X]_k = X(X-1)\cdots(X-k+1)$. (Later we use $[x]$ with no subscript to denote extraction of coefficients.)

The following statement of the small subgraph conditioning method is taken from [12]. A similar theorem is given in [4, Theorem 9.12].

Theorem 1.2 ([12], Theorem 4.1) *Let $\lambda_i > 0$ and $\delta_i \geq -1$ be real numbers for $i = 1, 2, \ldots$ and suppose that for each n there are random variables $X_i = X_i(n)$, $i = 1, 2, \ldots$ and $Y = Y(n)$, all defined on the same probability space $\mathcal{G} = \mathcal{G}_n$ such that X_i is nonnegative integer valued, Y is nonnegative and $\mathbf{E}Y > 0$ (for n sufficiently large). Suppose furthermore that*

(i) *For each $k \geq 1$, the variables $X_1, \ldots, X_k$ are asymptotically independent Poisson random variables with $\mathbf{E}X_i \to \lambda_i$,*

(ii)

$$\frac{\mathbf{E}(Y[X_1]_{j_1}\cdots[X_k]_{j_k})}{\mathbf{E}Y} \to \prod_{i=1}^{k}(\lambda_i(1+\delta_i))^{j_i}$$

for every finite sequence $j_1, \ldots, j_k$ of nonnegative integers,

(iii) $\sum_i \lambda_i\, \delta_i^{\,2} < \infty$,

(iv) $\mathbf{E}Y^2/(\mathbf{E}Y)^2 \leq \exp(\sum_i \lambda_i\, \delta_i^{\,2}) + o(1)$ *as $n \to \infty$.*

Then

$$\bar{\mathcal{G}}^{(Y)} \approx \bar{\mathcal{G}}$$

where $\bar{\mathcal{G}}$ is the probability space obtained from $\mathcal{G}$ by conditioning on the event $\wedge_{\delta_i=-1}(X_i = 0)$.

As in many contiguity proofs for random graphs, it is convenient to perform calculations in the pairing model. Let $\mathcal{P}_{n,d}$ denote the pairing model corresponding to $\mathcal{B}_{n,d}$. Here there are n buckets of each colour, and each bucket contains d points. A perfect bicoloured matching is chosen uniformly at random, and the edges of the matching are called the *pairs* of the pairing. Thus, $|\mathcal{P}_{n,d}| = (dn)!$. In particular,

$$|\mathcal{P}_{n,4}| = (4n)! \sim \sqrt{8\pi n}\, 4^{4n} \left(\frac{n}{e}\right)^{4n}.$$

(Usually $\mathcal{P}_{n,d}$ denotes the normal pairing model on n buckets of d points each, with no colours. However, we only refer to bicoloured models in this paper, so this notation should not cause confusion.)

An H-decomposition of $P \in \mathcal{P}_{n,4}$ is an ordered partition of the pairs of P into two sets, each of which corresponds to a bicoloured Hamilton cycle. One such set of pairs is called an *H-cycle*. For $P \in \mathcal{P}_{n,4}$, let $Y(P)$ be the number of H-decompositions of P. Let $\widehat{\mathcal{H}}_n$ denote the pairings version of $\mathcal{H}_n$. That is, $\widehat{\mathcal{H}}_n$ is the uniform probability space of pairings on $2n$ bicoloured buckets, each containing 2 points, with n buckets of each colour. We show that Y satisfies the conditions of Theorem 1.2, with $\delta_i > -1$ for all $i \geq 1$, and with X_i equal to the number of bicoloured cycles of length $2i$. Let a H-decomposition of $G \in \mathcal{B}_{n,4}$ be an ordered partition of the edges of G into two Hamilton cycles. Conditioning on no multiple edges, we obtain

$$\mathcal{B}_{n,4}^{(Y)} \approx \mathcal{B}_{n,4},$$

where here $Y(G)$ is the number of H-decompositions of $G \in \mathcal{B}_{n,4}$. But it is easy to see that

$$\mathcal{B}_{n,4}^{(Y)} = \mathcal{H}_n \oplus \mathcal{H}_n.$$

To apply the method we need to calculate the expectation and variance of Y, as well as the interaction between the number of short cycles and Y. The calculations are presented in Sections 2 and 3 below, and combined in Section 4. As is usual in the application of the small subgraph conditioning method, the calculation of the variance is by far the most difficult part. In this paper we employ the differential equation method described in [13] to calculate a critical quantity which contributes to the variance of Y.

It is well known that the probability that a random bipartite regular graph has given girth tends towards a non-zero constant. Thus, a simple corollary of Theorem 1.1 is that there exist 4-regular bipartite graphs which decompose into two Hamilton cycles and have arbitrarily large girth. An application of this has been found in topological group theory, where McCammond and Wise [6] use this to deduce the existence of graphs which provide examples of complexes with incoherent fundamental group.

2 Expectation and variance

Let $h(n)$ denote the number of bicoloured Hamilton cycles on $2n$ vertices. It is not difficult to see that

$$h(n) = \frac{n!^2}{2n}. \tag{3}$$

To form an H-decomposition, first select the adjacencies of the vertices in the two Hamilton cycles, then for each vertex choose one of the 4! ways to assign the four points to the four pairs involved. Hence

$$\mathbf{E}Y = \frac{h(n)^2\, 4!^{2n}}{(4n)!} \sim \frac{\pi^{3/2}}{\sqrt{8n}} \left(\frac{3}{2}\right)^{2n}. \tag{4}$$

We must also calculate the expected value of $Y(P)^2$. The argument begins with the same steps used by Kim and Wormald [5], but requires new arguments on random matchings.

We compute $Y(P)^2$ by viewing it as the number of ordered pairs of H-decompositions of a pairing P. Given $P \in \mathcal{P}_{n,4}$, let $((H_1, H_2), (H_3, H_4))$ be an ordered pair of H-decompositions of P. That is, each H_i is an H-cycle and $P = H_1 \cup H_2 = H_3 \cup H_4$. A pair in P is of type (i, j) if it belongs to H_i and H_{2+j}, for $1 \le i, j \le 2$. As in [5], a vertex is said to be of

- type A if it is incident with a pair of each type,
- type B if it is incident with two pairs of type $(1, 1)$ and two pairs of type $(2, 2)$,
- type C if it is incident with two pairs of type $(1, 2)$ and two pairs of $(2, 1)$.

(Note that these are the only possibilities.) Pairs of a given type (i, j) must form either closed cycles, or disjoint paths which start and end in type A vertices. The type $(1, 1)$ pairs (or type $(2, 2)$ pairs) can only form a closed cycle if they form a Hamilton cycle. This occurs if and only if $(H_1, H_2) = (H_3, H_4)$. The contribution to $\mathbf{E}(Y(P)^2)$ from such pairs shall be seen to be neglible. Similarly, if the type $(1, 2)$ pairs (or type $(2, 1)$ pairs) form a closed cycle then $(H_1, H_2) = (H_4, H_3)$. We ignore these cases. Therefore, pairs of a given type form disjoint paths which start and end in type A vertices. Moreover, each type A vertex is the endpoint of such a path. Therefore the number of type A vertices must be even. Each type B vertex must lie on a path of type $(1, 1)$ pairs, and a path of type $(2, 2)$ pairs. Similarly, each type C vertex must lie on a path of type $(1, 2)$ pairs, and a path of type $(2, 1)$ pairs.

By ignoring the type B and type C vertices, the (paths of) pairs of type (i, j) form a matching on the type A vertices. Relabel these four perfect matchings as M_1, M_2, M_3, M_4 where $M_1 = H_1 \cap H_3$, $M_2 = H_1 \cap H_4$, $M_3 = H_2 \cap H_4$ and $M_4 = H_2 \cap H_3$. So each type B vertex must lie on an edge of M_1 and of M_3, while each type C vertex must lie on an edge of M_2 and M_4. Let the colours used in the bicolouring be γ, δ. Suppose that there are $2k$ type A vertices, of which a_γ are coloured γ and a_δ are coloured δ. Suppose that there are b_γ type B vertices (c_γ type C vertices, respectively) coloured γ, and b_δ type B vertices (respectively, c_δ type C vertices) coloured δ. For brevity, a vertex coloured γ is called a γ-vertex (and similarly for vertices coloured δ), and an edge with endpoints coloured γ, δ is called an edge of end-type $\{\gamma, \delta\}$ (and similarly for other endpoint colours).

Let x_i be the number of bichromatic edges in M_i, for $1 \le i \le 4$. Then there are

$$w_i = \frac{a_\gamma - x_i}{2}$$

edges in M_i joining two γ-vertices, and

$$y_i = \frac{a_\delta - x_i}{2}$$

edges in M_i joining two δ-vertices, for $1 \le i \le 4$. In particular, $a_\delta - x_i$ must be even (which implies that $a_\gamma - x_i = 2k - a_\delta - x_i$ is also even). We now show that all the parameters can be written in terms of seven:

$$k,\ a_\delta,\ b_\delta,\ x_1,\ x_2,\ x_3,\ x_4.$$

Since $a_\gamma = 2k - a_\delta$, we can write w_i and y_i in terms of k, a_δ and x_i, for $1 \le i \le 4$. It is also clear that $c_\delta = n - a_\delta - b_\delta$. Moreover, we have

$$b_\delta - b_\gamma = w_1 - y_1 = w_3 - y_3$$

and

$$c_\delta - c_\gamma = w_2 - y_2 = w_4 - y_4.$$

Solving these gives $b_\gamma = a_\delta + b_\delta - k$ and $c_\delta = n - b_\delta - k$. For readability, we will not make these substitutions until later.

For fixed k and a_δ, let $p_H(\text{all} \mid k, a_\delta, x_1, x_2, x_3, x_4)$ denote the probability that $M_1 \cup M_2$, $M_2 \cup M_3$, $M_3 \cup M_4$ and $M_1 \cup M_4$ are all hamiltonian, when each M_i is a randomly chosen matching on $2k$ vertices, of which a_δ are δ-vertices and the rest are γ-vertices, conditional on fixed values of (x_1, x_2, x_3, x_4) satisfying the parity conditions mentioned above. The quantity $p_H(\text{all} \mid k, a_\delta, x_1, x_2, x_3, x_4)$ is analysed in Section 5 for certain values of the parameters (see Theorem 5.1). The number of ways to select $M_1, \dots, M_4$, such that $M_i \cup M_{i+1}$ is hamiltonian and M_i has x_i bicoloured edges, for $1 \le i \le 4$, is then

$$\begin{aligned} & p_H(\text{all} \mid k, a_\delta, x_1, x_2, x_3, x_4) \prod_{i=1}^{4} \binom{a_\gamma}{2w_i} \frac{(2w_i)!}{2^{w_i}\, w_i!} \binom{a_\delta}{2y_i} \frac{(2y_i)!}{2^{y_i}\, y_i!}\, x_i! \\ = \quad & \frac{p_H(\text{all} \mid k, a_\delta, x_1, x_2, x_3, x_4)\, 2^{x_1+x_2+x_3+x_4}\, a_\gamma!^4\, a_\delta!^4}{2^{4k} \prod_{i=1}^{4} x_i\,! y_i!\, w_i!}. \end{aligned} \tag{5}$$

There are

$$\binom{n}{a_\gamma}\binom{n}{a_\delta} \tag{6}$$

ways to choose the labels of the type A vertices. Now we count the number of ways to add the type B vertices onto matchings M_1 and M_3. To make sure the resulting graph is bicoloured, each edge of end-type $\{\gamma, \gamma\}$ requires an extra δ-vertex of type B, and each edge of end-type $\{\delta, \delta\}$ requires an extra γ-vertex of type B. The number of configurations is

$$\binom{b_\delta - w_1 + k - 1}{k - 1}\binom{b_\delta - w_3 + k - 1}{k - 1}. \tag{7}$$

Similarly, there are

$$\binom{c_\delta - w_2 + k - 1}{k - 1}\binom{c_\delta - w_4 + k - 1}{k - 1} \tag{8}$$

ways to add the type C vertices onto matchings M_2 and M_4. Then there are

$$b_\gamma! b_\delta! c_\gamma! c_\delta! \tag{9}$$

ways to identify the two copies of each of the type B and type C vertices, respecting colours. There are

$$(n - a_\gamma)!(n - a_\delta)! \tag{10}$$

ways to order the type B and type C vertices. Finally there are

$$4!^{2n} \tag{11}$$

ways to decide, for each edge, which particular points in the buckets corresponding to the endpoints are the ones joined by a pair corresponding to that edge. This determines the pairing in full.

Multiplying (5) – (11) and summing over all allowable values of $(k, a_\delta, b_\delta, x_1, x_2, x_3, x_4)$, (i.e. within the allowed range and satisfying the parity conditions) we obtain

$$\begin{aligned}\mathbf{E}Y^2 &= \frac{1}{|\mathcal{P}_{n,4}|} \sum_{(k,a_\delta,b_\delta,x_1,x_2,x_3,x_4)} \frac{p_H(\text{all} \mid k, a_\delta, x_1, x_2, x_3, x_4)\, k^4\, 4!^{2n}\, 2^{x_1+x_2+x_3+x_4}}{2^{4k}\,(b_\delta - w_1 + k)(b_\delta - w_3 + k)(c_\delta - w_2 + k)(c_\delta - w_4 + k)} \\ &\quad \times \frac{n!^2\, b_\gamma!\, b_\delta!\, c_\gamma!\, c_\delta!\, a_\gamma!^3\, a_\delta!^3\, (b_\delta - w_1 + k)!(b_\delta - w_3 + k)!(c_\delta - w_2 + k)!(c_\delta - w_4 + k)!}{k!^4\, \left(\prod_{i=1}^4 x_i!\, y_i!\, w_i!\right) (b_\delta - w_1)!\,(b_\delta - w_3)!\,(c_\delta - w_2)!\,(c_\delta - w_4)}.\end{aligned}$$

Assuming that all arguments of the factorial function tend to infinity (justified below), we obtain by Stirling's formula

$$\begin{aligned}\mathbf{E}Y^2 &\sim \sum_{(k,a_\delta,b_\delta,x_1,x_2,x_3,x_4)} \frac{p_H(\text{all} \mid k, a_\delta, x_1, x_2, x_3, x_4)\, n\, k^2\, \sqrt{b_\gamma b_\delta c_\gamma c_\delta}\, {a_\gamma}^{3/2}\, {a_\delta}^{3/2}}{8\sqrt{2}\pi^2\, \left(\prod_{i=1}^4 \sqrt{x_i y_i w_i}\right) \sqrt{(b_\delta - w_1)(b_\delta - w_3)(c_\delta - w_2)(c_\delta - w_4)}} \\ &\quad \times \frac{f(\kappa, \alpha, \beta, \chi_1, \chi_2, \chi_3, \chi_4)^n}{\sqrt{(b_\delta - w_1 + k)(b_\delta - w_3 + k)(c_\delta - w_2 + k)(c_\delta - w_4 + k)}}\end{aligned} \tag{12}$$

where

$$\begin{aligned}&f(\kappa, \alpha, \beta, \chi_1, \chi_2, \chi_3, \chi_4) \\ &\quad = \frac{9}{4} \times \frac{2^{\chi_1+\chi_2+\chi_3+\chi_4}\,(2\kappa - \alpha)^{3(2\kappa-\alpha)}\, \alpha^{3\alpha}\,(\beta + \alpha - \kappa)^{(\beta+\alpha-\kappa)}\, \beta^\beta (1 - \alpha - \beta)^{1-\alpha-\beta}}{2^{4\kappa}\, \kappa^{4\kappa} \left(\prod_{i=1}^4 {\chi_i}^{\chi_i}\, (\frac{\alpha-\chi_i}{2})^{(\alpha-\chi_i)/2}\, (\kappa - \frac{\alpha+\chi_i}{2})^{\kappa-(\alpha+\chi_i)/2}\right)} \\ &\quad\quad \times \frac{(1 - \kappa - \beta)^{1-\kappa-\beta}\, (\beta + \frac{\alpha+\chi_1}{2})^{(\beta+(\alpha+\chi_1)/2)}\, (\beta + \frac{\alpha+\chi_3}{2})^{\beta+(\alpha+\chi_3)/2}}{(\beta - \kappa + \frac{\alpha+\chi_1}{2})^{\beta-\kappa+(\alpha+\chi_1)/2}\, (\beta - \kappa + \frac{\alpha+\chi_3}{2})^{\beta-\kappa+(\alpha+\chi_3)/2}} \\ &\quad\quad \times \frac{(1 - \beta - \frac{\alpha-\chi_2}{2})^{(1-\beta-(\alpha-\chi_2)/2}\, (1 - \beta - \frac{\alpha-\chi_4}{2})^{1-\beta-(\alpha-\chi_4)/2}}{(1 - \beta - \kappa - \frac{\alpha-\chi_2}{2})^{1-\beta-\kappa-(\alpha-\chi_2)/2}\, (1 - \beta - \kappa - \frac{\alpha-\chi_4}{2})^{1-\beta-\kappa-(\alpha-\chi_4)/2}}.\end{aligned} \tag{13}$$

Here $\kappa = k/n$, $\alpha = a_\delta/n$, $\beta = b_\delta/n$, $\chi_i = x_i/n$ for $1 \le i \le 4$, and all other variables have been substituted out. We look at f over the domain $D \subseteq [0,1]^7$ defined by

$$\begin{aligned}D \;=\; &\{(\kappa, \alpha, \beta, \chi_1, \chi_2, \chi_3, \chi_4) \in [0,1]^7 \mid \text{ every factor of the form } t^t \text{ in} \\ &\quad f(\kappa, \alpha, \beta, \chi_1, \chi_2, \chi_3, \chi_4) \text{ has } t \ge 0 \,\}\end{aligned} \tag{14}$$

(with the convention that $0^0 = 1$).

We now justify our assumption that the arguments of factorials tend to infinity as $n \to \infty$. In Section 6 we will show that the unique global maximum of f over the closed domain D is attained when

$$(\kappa, \alpha, \beta, \chi_1, \chi_2, \chi_3, \chi_4) = (\tfrac{2}{3}, \tfrac{2}{3}, \tfrac{1}{6}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}), \tag{15}$$

which is an interior point of D. Any term of the summation in (12) in which the argument of a factorial is bounded corresponds, after scaling, to a point arbitrarily close to the boundary of D. In this case the asymptotic relation in (12) is not valid, but the error is at most a constant factor. Since f is raised to the power n, the contribution of all such terms is negligible.

Now set $\chi_i = \frac{1}{2} + z_i$ for $1 \le i \le 4$, $\kappa = \frac{2}{3} + z_5$, $\alpha = \frac{2}{3} + z_6$ and $\beta = \frac{1}{6} + z_7$. Expanding $\log f$ around $(z_1, \dots, z_7) = (0, \dots, 0)$ we obtain

$$\begin{aligned}
f(z_1, \dots, z_7) &= \frac{81}{16} \exp\left(\left[\sum_{i=1}^{4} -\frac{16}{3} {z_i}^2 + 12 z_i z_5 + (-1)^i \frac{8}{3} z_i z_6 + (-1)^i \frac{16}{3} z_i z_7 \right] \right. \\
&\qquad \left. -36 {z_5}^2 + 9 z_5 z_6 - \frac{41}{6} {z_6}^2 - \frac{28}{3} z_6 z_7 - \frac{28}{3} {z_7}^2 + T \right) \\
&= \frac{81}{16} \exp\left(\left[\sum_{i=1}^{4} -\frac{1}{48} (16 z_i - 18 z_5 - (-1)^i 4 z_6 - (-1)^i 8 z_7)^2 \right] \right. \\
&\qquad \left. -\frac{9}{2} {z_5}^2 - \frac{9}{2} (z_5 - z_6)^2 - (z_6 - 2 z_7)^2 + T \right)
\end{aligned}$$

where $T = T(z_1, \dots, z_7)$ is a sum of terms $O(z_i z_j z_k)$. Consider the cube

$$\mathcal{C} = \left\{ (z_1, \dots, z_7) \mid -n^{-2/5} \le z_i \le n^{-2/5} \text{ for } 1 \le i \le 7 \right\}$$

centred at the origin. Since $\exp(O(n(n^{-2/5})^3)) = 1 + o(1)$, the contribution to $f(z_1, \dots, z_7)^n$ from T is negligible for all points $(z_1, \dots, z_7) \in \mathcal{C}$. Using standard arguments, since the unique maximum of f on the closed domain D is attained at (15), points outside $\mathcal{C}$ make negligible contribution to $f(z_1, \dots, z_7)^n$ and

$$\begin{aligned}
&\sum_{(k, a_\delta, b_\delta, x_1, x_2, x_3, x_4)} f(\kappa, \alpha, \beta, \chi_1, \chi_2, \chi_3, \chi_4)^n \\
&\sim \left(\frac{81}{16} \right)^n \frac{n^7}{16} \int_{-n^{-2/5}}^{n^{-2/5}} \cdots \int_{-n^{-2/5}}^{n^{-2/5}} \exp\left(\left[\sum_{i=1}^{4} -\frac{n}{48} (16 z_i - 18 z_5 - (-1)^i 4 z_6 - (-1)^i 8 z_7)^2 \right] \right. \\
&\qquad \left. -\frac{9n}{2} {z_5}^2 - \frac{9n}{2} (z_5 - z_6)^2 - n (z_6 - 2 z_7)^2 \right) dz_1 \cdots dz_7 \\
&\sim \left(\frac{81}{16} \right)^n \frac{n^7}{16} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left(\left[\sum_{i=1}^{4} -\frac{n}{48} (16 z_i - 18 z_5 - (-1)^i 4 z_6 - (-1)^i 8 z_7)^2 \right] \right. \\
&\qquad \left. -\frac{9n}{2} {z_5}^2 - \frac{9n}{2} (z_5 - z_6)^2 - n (z_6 - 2 z_7)^2 \right) dz_1 \cdots dz_7 \\
&= \frac{(\pi n)^{7/2}}{4096} \left(\frac{81}{16} \right)^n.
\end{aligned}$$

(We divide by 16 because the parity of x_i is determined by k and a_δ, for $1 \le i \le 4$.)

When $k = \frac{2n}{3} + O(n^{-2/5})$, $a_\delta = \frac{2n}{3} + O(n^{-2/5})$, and $x_i = \frac{n}{2} + O(n^{-2/5})$ for $1 \le i \le 4$, we conclude by Theorem 5.1 that

$$p_H(\text{all} \mid k,\, a_\delta,\, x_1,\, x_2,\, x_3,\, x_4) \sim \frac{81\pi^2}{1024n^2}.$$

The trivial upper bound $p_H(\text{all} \mid k,\, a_\delta,\, x_1,\, x_2,\, x_3,\, x_4) \le 1$ holds for all other values of the parameters. As mentioned above, only values corresponding to points in the cube $\mathcal{C}$ make a nonnegligible contribution to (12). Hence we can substitute $k = \frac{2n}{3}$, $a_\delta = \frac{2n}{3}$, and so on (see (15)), into the non-exponential part of (12) and conclude that

$$\mathbf{E}Y^2 \sim \frac{3\sqrt{2}\pi^3}{32\,n}\left(\frac{81}{16}\right)^n.$$

Dividing this by $(\mathbf{E}Y)^2$ using (4) gives

$$\frac{\mathbf{E}Y^2}{(\mathbf{E}Y)^2} \sim \frac{3\sqrt{2}}{4}. \tag{16}$$

3 Interaction with short cycles

We must calculate

$$\frac{\mathbf{E}(YC_k)}{\mathbf{E}Y}$$

where C_k is the number of $2k$-cycles in $\mathcal{P}_{n,4}$. To do this we calculate

$$|\mathcal{P}_{n,4}|\,\mathbf{E}(YC_k) = \sum_{C} \sum_{\substack{P \in \mathcal{P}_{n,4} \\ C \subseteq P}} Y(P),$$

where the first sum is over all possible labelled bicoloured $2k$-cycles C. Let C be such a cycle which has been endowed with a direction. There are

$$\frac{{[n]_k}^2}{k} \sim \frac{n^{2k}}{k}$$

choices for C (compare with (3)). Edges of the cycle will correspond to pairs in a pairing $P \in \mathcal{P}_{n,4}$, with an H-decomposition (H_1, H_2). We calculate the number of ways to complete $C \cap H_1$ and $C \cap H_2$ to give bicoloured Hamilton cycles on $n+n$ vertices. Then there are $(4!)^{2n}$ ways to assign endpoints in the vertices (buckets) to the edges so determined.

Now $C \cap H_1$ is the union of i disjoint paths, for some $i \ge 1$. Then $C \cap H_2$ is also the union of i disjoint paths. As we shall see, the value of i will affect the number of ways the the Hamilton cycles can be chosen. We wish to count the $2k$-cycles in which each edge is specified to be in H_1 or H_2, such that edges in H_j form i disjoint paths of length at least 1, for $j = 1, 2$. To count these, start at an arbitrary vertex v at the start of a path induced by H_1 and proceed around the cycle in the chosen direction. There are $2k$ ways to choose the vertex to begin at, and the number of ways to determine the lengths of paths $P_1, Q_1, P_2, Q_2, \ldots, P_i, Q_i$ is

$$[x^{2k}]\left(\frac{x}{1-x}\right)^{2i}.$$

Here square brackets denote the extraction of coefficients, and the formula comes from considering the concatenation of $2i$ paths of length at least 1 each, so that the generating function for each individual path is $\frac{x}{(1-x)}$. This expression must be multiplied by

$$\frac{k}{i}$$

since there are $2k$ ways of choosing the starting vertex v, and then every configuration has been counted $2i$ times, since there are i paths P_l and two orientations of the cycle.

After C and its partition into paths is decided, let F denote the graph with vertex set $V(C)$ whose edge set consists of the edges in both C and H_1. Then F is a set of paths and isolated vertices. Let s (for "same") denote the number of paths in F with γ-vertices at both ends, plus the number of isolated γ-vertices in F. Then the number of paths with δ-vertices at each end, plus the number of isolated δ-vertices, is also equal to s. Let d (for "different") denote the number of paths with endpoints coloured differently. Then the rest of the edges in H_1 can be chosen in asymptotically

$$\frac{1}{2}(n-k+s)!(n-k+s-1)!(2n)^d 2^{i-d} \sim (n!)^2 n^{d-2k+2s-1} 2^{i-1}$$

ways. This comes from a cyclic biparitite ordering of the $(2n-2k+2s)$-set whose elements are the vertices not in C and the $\{\gamma,\gamma\}$ and $\{\delta,\delta\}$ paths in C, into which the d $\{\gamma,\delta\}$ paths have been inserted. There are $(2n-2k+2s)^d \sim (2n)^d$ ways to insert all the $\{\gamma,\delta\}$ paths, and the factor of $2^i - d$ accounts for the two choices of the end vertex of a $\{\gamma,\gamma\}$ or $\{\delta,\delta\}$ path.

Similarly, the rest of the edges in H_2 can be chosen in asymptotically

$$(n!)^2 n^{d'-2k+2s'-1} 2^{i-1}$$

ways, where s' and d' are defined like s and d. It is easily seen that the total number of edges $(2k)$ in the cycle C is $2k-d-2s+2k-d'-2s'$, and so the product of the two factors above, regardless of d, d', s and s', is asymptotic to

$$(n!)^4 n^{-2k-2} 4^{i-1}.$$

Putting all this together, we obtain that

$$\mathbf{E}(YC_k) \sim \frac{n^{2k}}{k}\,[x^{2k}]\sum_{i\geq 1}\left(\frac{2x}{1-x}\right)^2 \cdot \frac{k}{i} \cdot \frac{n!^4}{4\,n^{2k+2}} \cdot 4!^{2n}.$$

Dividing by $\mathbf{E}Y$ gives

$$\frac{\mathbf{E}(YC_k)}{\mathbf{E}Y} \sim \sum_{i\geq 1}[x^{2k}]\left(\frac{4x}{(1-x)^2}\right)^i \frac{1}{i} = \frac{3^{2k}-1}{2k} = \rho_k \tag{17}$$

for $k \geq 1$. A direct generalisation of this argument, applied to an ordered set of i_1 cycles of length 2, i_2 bicoloured cycles of length 4, and so on, shows that

$$\frac{\mathbf{E}\left(Y[C_1]_{i_1}\cdots[C_j]_{i_j}\right)}{\mathbf{E}Y} \sim \prod_{k=1}^{j} {\rho_k}^{i_k}.$$

4 Synthesis

We now combine the results of Sections 2 and 3. One further piece of information is required, namely the short cycle distribution in $\mathcal{P}_{n,4}$. Let C_k be the number of bicoloured cycles of length $2k$ in $P \in \mathcal{P}_{n,4}$. As is well known (see for example [10]), the C_k are asymptotically independent Poisson random variables with expectations

$$\mathbf{E}C_k \sim \lambda_k = \frac{3^{2k}}{2k}.$$

Recall ρ_k defined in (17), and define δ_k by

$$\delta_k = \frac{\rho_k}{\lambda_k} - 1 = -\frac{1}{3^{2k}}.$$

Note that $\delta_k > -1$ for $k \geq 1$. Then

$$\begin{aligned} \exp\left(\sum_{k=1}^{\infty} \lambda_k\, {\delta_k}^2\right) &\sim \exp\left(\frac{1}{2}\sum_{k=1}^{\infty}\frac{1}{k\,3^{2k}}\right) \\ &= \exp\left(-\frac{1}{2}\log\frac{8}{9}\right) \\ &= \frac{3\sqrt{2}}{4} \\ &\sim \frac{\mathbf{E}Y^2}{(\mathbf{E}Y)^2}, \end{aligned}$$

using (16). Thus by Theorem 1.2, we conclude that

$$\mathcal{P}_{n,4} \approx \mathcal{P}_{n,4}^{(Y)}.$$

By conditioning on no multiple edges, we obtain Theorem 1.1, as explained in Section 1.

5 Random matchings

Assume that, as in Section 2, there are $2k$ type A vertices, of which a_δ are coloured δ and the remaining a_γ are coloured γ. Recall that $p_H(\text{all} \mid k,\, a_\delta,\, x_1,\, x_2,\, x_3,\, x_4)$ is the probability that $M_1 \cup M_2$, $M_2 \cup M_3$, $M_3 \cup M_4$ and $M_4 \cup M_1$ all form Hamilton cycles, where M_i is a perfect matching of the $2k$ type A vertices which contains x_i bichromatic edges, chosen uniformly at random and independently, for $1 \leq i \leq 4$.

The aim of this section is to prove the following result, used in Section 2.

Theorem 5.1 *If* $k = \frac{2n}{3} + O(n^{-2/5})$, $a_\delta = \frac{2n}{3} + O(n^{-2/5})$, *and* $x_i = \frac{n}{2} + O(n^{-2/5})$ *for* $1 \le i \le 4$, *then*

$$p_H(\text{all} \mid k,\, a_\delta,\, x_1,\, x_2,\, x_3,\, x_4) \sim \frac{81\pi^2}{1024n^2}.$$

Throughout this section we assume that k, a_δ and the x_i satisfy the requirements of Theorem 5.1. (These values are the only significant ones in the variance calculations of Section 2, and the value of all other parameters can be determined from these values.)

Suppose that M is a perfect matching on the given $2k$ vertices with x bichromatic edges. We refer to M simply as a *matching with parameter* x. We can generate a random M with given parameter x, uniformly at random using the following simple stochastic process with three phases. Start with $M = \emptyset$ and add an edge $\{u, v\}$ to M at each time step, chosen as follows. For steps $1, \ldots, (a_\gamma - x)/2$, let u and v be distinct unmatched γ-vertices, chosen uniformly at random (A vertex is *unmatched* if it does not belong to any edge currently in M.) This constitutes phase 1. For the next x steps, let u be an unmatched γ-vertex chosen uniformly at random, and let v be an unmatched δ-vertex chosen uniformly at random. After these steps, all γ-vertices are matched in M. This is phase 2. For the last $(a_\delta - x)/2$ steps, let u and v be distinct unmatched δ-vertices, chosen uniformly at random. This is phase 3, and is similar to the process used in [5]. Clearly the resulting perfect matching M has parameter x, and is distributed uniformly at random over all such perfect matchings.

When B and R are matchings, we will refer to their edges as being blue and red, respectively. The number of cycles (and 2-cycles) of the graph $BR = B \cup R$ will be of interest.

Lemma 5.1 *Fix* x *such that* $0 \le x \le k$. *Let* B *be a fixed perfect matching on the given* $2k$ *vertices, and let* R *be such a matching chosen unformly at random from those with parameter* x. *Let* $\mathbf{P}_R$ *denote probability with respect to this choice of* R, *and let* $\kappa(BR)$ *be the number of cycles in* BR. *Then*

$$\mathbf{P}_R(\kappa(BR) > 5 \log n) = O(n^{-3}).$$

Proof. Edges of R are generated one by one, using the procedure described above. Assume that during phase 1, the vertex u is chosen first (uniformly at random from all unmatched γ-vertices) and then the vertex v is chosen uniformly, such that v is unmatched and distinct from u. At the ith step, for any choice of u, there are $a_\gamma - 2i + 1$ choices for v of which at most one causes the ith red edge to create a cycle with blue and previous red edges. So the probability of creating a cycle at the ith step of phase 1 is at most $1/(a_\gamma - 2i + 1)$. It is easy to see that the event is independent of the previous choices and hence, letting X denote the number of cycles created in phase 1, we have that

$$X \le \sum_{i=1}^{(a_\gamma - x)/2} \varphi_i,$$

for appropriate independent random variables φ_i with

$$\mathbf{P}(\varphi_i = 1) = 1 - \mathbf{P}(\varphi_i = 0) = \frac{1}{a_\gamma - 2i + 1}.$$

Similarly, the number Y of cycles created in phases 2 and 3 satisfies

$$Y \leq \sum_{i=1}^{x} \tau_i + \sum_{i=1}^{(a_\delta - x)/2} \sigma_i$$

for appropriate independent random variables τ_i, σ_i with

$$\mathbf{P}(\tau_i = 1) = 1 - \mathbf{P}(\tau_i = 0) = \frac{1}{a_\delta - i + 1}$$

and

$$\mathbf{P}(\sigma_i = 1) = 1 - \mathbf{P}(\sigma_i = 0) = \frac{1}{a_\delta - x - 2i + 1}.$$

Thus for

$$Z = \sum_{i=1}^{(a_\gamma - x)/2} \varphi_i + \sum_{i=1}^{x} \tau_i + \sum_{i=1}^{(a_\delta - x)/2} \sigma_i,$$

we have that

$$\mathbf{P}(X + Y \geq 5 \log k) \leq \mathbf{P}(Z \geq 5 \log k).$$

Since $a_\gamma, a_\delta, k \sim 2n/3$ and Z is a sum of k independent random variables with $\mathbf{E}(Z) = (1 + o(1)) \log k$, a standard argument shows that

$$\mathbf{P}(Z \geq 5 \log n) = O(n^{-3}).$$

For example, apply the argument of [5, Lemma 1]. $\square$

Let HAM denote the event "is a Hamilton cycle". For fixed perfect matchings B and R on the $2k$ vertices with parameters x_1, x_3, respectively, let

$$\mathbf{P}_{(x)}(BS, RS \in \text{HAM})$$

denote the probability that both BS and RS form Hamilton cycles, where S is a uniformly chosen perfect matching of the same set of vertices with parameter x. Note that $\mathbf{P}_{(x)}(BS \in \text{HAM})$ depends only on k, a_δ, x_1, x_3 and x. We call S the silver matching.

Lemma 5.2 *Let $x \in \{x_2, x_4\}$. Suppose that*

$$\mathbf{P}_{(x)}(BS, RS \in \text{HAM}) \sim \frac{9\pi}{32n}, \tag{18}$$

for perfect matchings B and R with parameters x_1, x_3 respectively, such that $\kappa(BR) \leq 5 \log n$, with asymptotics uniform over all B, R and x such that x_1, x_3 and x are all $n/2 + O(n^{-2/5})$. Then Theorem 5.1 *holds.*

Proof. Choose M_1 uniformly with parameter x_1. Then consider M_1 fixed and choose M_3 uniformly with parameter x_3. We will apply Lemma 5.1 with $B = M_1$, $R = M_3$, letting $\mathbf{E}_B$, $\mathbf{E}_R$ denote expectation with respect to choices of B and R, respectively. For brevity, let $P_2(B,R) = \mathbf{P}_{(x_2)}(BS, RS \in \text{HAM})$ and $P_4(B,R) = \mathbf{P}_{(x_4)}(BS, RS \in \text{HAM})$. In the first of these, S stands for M_2 and in the second, M_4. Then

$$\begin{aligned}
& p_H(\text{all} \mid k,\, a_\delta,\, x_1,\, x_2,\, x_3,\, x_4) \\
= \;& \mathbf{E}_B\, \mathbf{E}_R \Big(P_2(B,R) P_4(B,R) \Big) \\
= \;& \mathbf{E}_B \Big(\mathbf{P}_R(\kappa(BR) \le 5\log n) \mathbf{E}_R \Big(P_2(B,R)P_4(B,R) \mid \kappa(BR) \le 5 \log n \Big) \\
& \qquad + \mathbf{P}_R(\kappa(BR) > 5\log n) \mathbf{E}_R \Big(P_2(B,R)P_4(B,R) \mid \kappa(BR) > 5\log n \Big) \Big) \\
= \;& \mathbf{E}_B \left(\Big(1 - O(n^{-3})\Big) \frac{81\pi^2}{1024\, n^2} + O(n^{-3}) \right) \\
\sim \;& \frac{81\pi^2}{1024 n^2},
\end{aligned}$$

using (18). ∎

It remains to prove that (18) holds for fixed matchings B, R with parameters x_1, x_3 respectively, such that $\kappa(BR) \le 5\log n$ and $x_1, x_3 = n/2 + O(n^{-2/5})$. We must choose a silver matching S uniformly at random from those with parameter x, where $x = n/2 + O(n^{-2/5})$ also. This is achieved using the stochastic process with three phases, slightly modified from the one described above, so as to keep track of cycles produced in BS and RS. (This modification is similar to one used in [5] for a simpler process.)

After t edges of S have been determined, we define a matching $B(t)$ such that two vertices are adjacent in $B(t)$ if they are endpoints of the same path in BS. Define $R(t)$ similarly using paths in RS. Set $BR(t) = B(t) \cup R(t)$. At time $t \ge 1$, randomly select two vertices u_t and v_t from $BR(t-1)$, with the colours of u_t and v_t determined by the phase of the process, as described above. (For example, in phase 1, u_t and v_t are distinct γ-vertices, chosen uniformly at random.) Then $\{u_t, v_t\}$ becomes an edge in S. When $t = k$, the set S is a perfect matching.

It may help to consider the following alternative definition of $BR(t)$. For brevity, a blue edge in $BR(t)$ is called a B-edge, and similarly for R. At time t, a vertex u has a unique B-neighbour and a unique R-neighbour, which is other endpoint of the B-edge (respectively R-edge) of $BR(t)$ containing u. The graph $BR(t)$ is formed from $BR(t-1)$, u_t and v_t, as follows. If $\{u_t, v_t\}$ is a B-edge in $BR(t-1)$ then we delete this edge from $BR(t-1)$. Otherwise, we delete u_t and v_t, and join the B-neighbour of u_t to the B-neighbour of v_t with a new B-edge. This is called "contracting" the edge $\{u_t, v_t\}$. Perform this operation with respect to R as well. The resulting graph $BR(t)$ is the union of two perfect matchings on $2k - 2t$ points. A cycle is formed in BS (respectively RS) at step t if and only if $\{u_t, v_t\}$ is an existing B-edge (respectively, R-edge) of $BR(t-1)$. Thus, BS and RS are both Hamilton cycles if and only if $\{u_t, v_t\}$ is not equal to any edge of $BR(t-1)$, for all $t < k$.

Denote by T_i the number of steps in phase i, for $1 \le i \le 3$. That is,

$$T_1 = (a_\gamma - x)/2, \;\; T_2 = x, \;\; T_3 = (a_\delta - x)/2.$$

Note that S has exactly $x = T_2$ edges of end-type $\{\gamma, \delta\}$. (Hence it will have exactly T_1 edges of end-type $\{\gamma, \gamma\}$ and exactly T_3 edges of end-type $\{\delta, \delta\}$.) We need to calculate the asymptotic value that no cycle is produced in BS or RS in all but the last step of this process. To analyse phases 1 and 2, we model the number of edges of certain end-types by a continuous function, using the differential equation method given in [13, Theorem 5.1], stated below. In phase 3, there are only vertices of one colour left, so we can use arguments similar to those of [5].

The theorem stated below is a slightly simplified version of that given in [13, Theorem 5.1]. (In particular, readers who consult [13, Theorem 5.1] can check that we take $a = 1$, $\beta = C_0 \geq 1$ and $\gamma = 0$.) Consider any discrete-time random process, which forms a probability space which may be denoted by $(Q_0, Q_1, \dots)$, where each Q_i is a (random) element of some set S. Let $H_t = (Q_0, \dots, Q_t)$ be the history of the process up to time t.

Now consider a sequence of random processes indexed by n for $n = 1, 2, \dots$. Thus $Q_t = Q_t^{(n)}$ and $S = S^{(n)}$, but the dependence on n is often dropped from the notation. Asymptotics are for $n \to \infty$ but are uniform over all other variables. Let $S^{(n)+}$ denote the set of all $H_t = (Q_0, \dots, Q_t)$ where $Q_i \in S^{(n)}$, for $t = 0, 1, 2, \dots$.

We say that a function $f(u_1, \dots, u_j)$ satisfies a *Lipschitz* condition on $D \subseteq \mathbb{R}^j$ if there exists a constant $L > 0$ such that

$$|f(u_1, \dots, u_j) - f(v_1, \dots, v_j)| \leq L \max_{1 \leq i \leq j} |u_i - v_i|$$

for all $(u_1, \dots, u_j)$, $(v_1, \dots, v_j) \in D$. Note that $\max_{1 \leq i \leq j} |u_i - v_i|$ is the distance between $(u_1, \dots, u_j)$ and $(v_1, \dots, v_j)$ in the ℓ^∞ metric.

For a variable Y defined on components of the process, and for $D \subseteq \mathbb{R}^2$, define the *stopping time* $T_D(Y)$ to be the minimum t such that $(t/n, Y(t)/n) \notin D$.

Theorem 5.2 ([13], Theorem 5.1) *Let $Y : S^{(n)+} \to \mathbb{R}$ and $f : \mathbb{R}^2 \to \mathbb{R}$ be functions such that for some constant $C_0 \geq 1$, we have $|Y(H_t)| < C_0 n$ for all $H_t \in S^{(n)+}$ and for all n. For simplicity, denote $Y(H_t)$ by $Y(t)$. Assume the following three conditions hold, where in (ii) and (iii) D is some bounded connected open set containing the closure of*

$$\{(0, z) : \mathbf{P}(Y(0) = zn) \neq 0 \text{ for some } n\}.$$

(i) *For all t we have*

$$|Y(t+1) - Y(t)| \leq C_0.$$

(ii) *For some function $\lambda_1 = \lambda_1(n) = o(1)$, we have*

$$|\mathbf{E}(Y(t+1) - Y(t) \mid H_t) - f(t/n, Y(t)/n)| \leq \lambda_1$$

for $t < T_D(Y)$.

(iii) *The function f is continuous and satisfies a Lipschitz condition on*

$$D \cap \{(t, z) : t \geq 0\}.$$

Then the following are true.

(a) *For $(0, \hat{z}) \in D$, the differential equation*

$$\frac{dz}{dx} = f(x, z)$$

has a unique solution in D for $z : \mathbb{R} \to \mathbb{R}$ passing through

$$z(0) = \hat{z}$$

which extends to points arbitrarily close to the boundary of D;

(b) *Let $\lambda > \lambda_1$ with $\lambda = o(1)$. For a sufficiently large constant C, with probability $1 - O(\lambda^{-1} \exp(-n\lambda^3))$, we have*

$$Y(t) = nz(t/n) + O(\lambda n)$$

uniformly for $0 \le t \le \sigma n$, where $z(x)$ is the solution in (a) *with $\hat{z} = Y(0)/n$, and $\sigma = \sigma(n)$ is the supremum of those x to which the solution can be extended before reaching within ℓ^∞-distance $C\lambda$ of the boundary of D.*

5.1 Phase 1

By a slight abuse of notation, let γ_t be the number of γ-vertices in $BR(t)$ for $t \ge 0$. So $\gamma_0 = a_\gamma = 2n/3 + O(n^{-2/5})$. At each step of phase 1, two γ-vertices are deleted. Therefore $\gamma_t = \gamma_0 - 2t$.

Let $w_t^{(B)}$ be the number of B-edges of end-type $\{\gamma, \gamma\}$ in $BR(t)$, for $0 \le t \le T_1$. Now $w_{t+1}^{(B)} = w_t^{(B)} - 1$ unless the B-neighbour of both u_{t+1} and v_{t+1} is coloured δ, in which case $w_{t+1}^{(B)} = w_t^{(B)}$. Thus $|w_t^{(B)}|/n$ and $|w_{t+1}^{(B)} - w_t^{(B)}|$ are both bounded above by a constant, as required for Theorem 5.2. Let $\Gamma_0 = \gamma_0/n$, and let $H_t = (BR(0), BR(1), \ldots, BR(t))$ be the history of the process for steps 1 to t. The above discussion shows that

$$\begin{aligned} \mathbf{E}(w_{t+1}^{(B)} - w_t^{(B)} \mid H_t) &= \frac{\binom{\gamma_0 - 2t - 2w_t^{(B)}}{2}}{\binom{\gamma_0 - 2t}{2}} - 1 \\ &= f(t/n, w_t^{(B)}/n) + \lambda_1 \end{aligned}$$

where $\lambda_1 = O(n^{-2/5})$ and

$$f(s, z) = \frac{(1 - 3s - 3z)^2}{(1 - 3s)^2} - 1.$$

Fix small positive constant η and define the open set

$$D = \{(s, z) \mid -\eta < s < 1/12 - \eta,\ -\eta < z < 1/3 + \eta\}.$$

Then f is continuous on the closure of D and hence satisfies the Lipschitz condition required for Theorem 5.2.

Now $\Gamma_0 = 2n/3 + O(n^{3/5})$ and $T_1 = n/12 + O(n^{3/5})$ by assumption, so $(t/n, w_t^{(B)}/n)$ must meet the boundary of D at $t/n = 1/12 - \eta$. Then the requirements of Theorem 5.2 are satisfied, with λ_1 as above and $\lambda = O(n^{-1/10})$, say.

Solving $z'(s) = f(s, z)$, with initial condition $z(0) = z_0$, gives the solution

$$z(s) = \frac{z_0(1-3s)^2}{1-9z_0 s}.$$

By Theorem 5.2, this solution is unique in D for all z_0 such that $(0, z_0) \in D$, and extends to points arbitrarily close to the boundary of D. Setting $z_0 = w_0^{(B)}/n$, we conclude that, with probability $1 - O(n^{-2})$,

$$\begin{aligned} w_t^{(B)} &= n\, z(t/n) + o(n) \\ &= \frac{(n-3t)^2}{3(4n-3t)} + o(n) \end{aligned} \tag{19}$$

for $0 \le t \le n/12 - \eta n$. This uses the fact that $w_0^{(B)} = n/12 + O(n^{3/5})$. (The error probability given in Theorem 5.2 is $O(\lambda^{-1}\exp(-n\lambda^3))$, and

$$O(\lambda^{-1}\exp(-n\lambda^3)) = O(n^{1/10}\exp(-n^{7/10})) = O(n^{-2})$$

by choice of λ.) Noting that in ηn steps the change in $w_t^{(B)}$ is $O(\eta)$, we can let $\eta \to 0$ slowly and deduce (19) uniformly for all $t \le T_1$, with probability $1 - O(n^{-2})$. The same conclusion can be reached with R in place of B.

For $1 \le \tau \le T_1$, define the event $\mathcal{O}_\tau^{(B)}$ that (19) holds for $1 \le t \le \tau$, with some fixed error function in mind. This is an event over the set of all possible τ-step histories of the process, which says that the variable $w_t^{(B)}$ remains within distance $o(n)$ of the solution given by the differential equation, up until time τ, for a function $o(n)$ determined from the error function implicit in (19). In other words, the variables $w_t^{(B)}$ are well-behaved, for $1 \le t \le \tau$. Similarly define $\mathcal{O}_\tau^{(R)}$, and let

$$\mathcal{O}_\tau = \mathcal{O}_\tau^{(B)} \wedge \mathcal{O}_\tau^{(R)}.$$

This is the probability that both $w_t^{(B)}$ and $w_t^{(R)}$ are well behaved for $1 \le t \le \tau$. We know that $\mathcal{O}_{T_1}^{(B)}$ and $\mathcal{O}_{T_1}^{(R)}$ both hold with probability $1 - O(n^{-2})$. Therefore $\mathcal{O}_{T_1}$ holds with probability $1 - O(n^{-2})$ as well. If $\mathcal{O}_{T_1}$ holds then so do $\mathcal{O}_\tau$, $\mathcal{O}_\tau^{(B)}$ and $\mathcal{O}_\tau^{(R)}$, for $1 \le \tau \le T_1$.

Now define $\text{cycle}^{(B)}(t)$ to be the event that a cycle of BS is formed at step t of phase 1, and similarly $\text{cycle}^{(R)}(t)$. Let $\text{cycle}(t) = \text{cycle}^{(B)} \vee \text{cycle}^{(R)}(t)$ be the event that a cycle was formed either BS or RS at step t, and finally let $\text{both}(t) = \text{cycle}^{(B)} \wedge \text{cycle}^{(R)}(t)$ be the event that a cycle was formed in both BS and RS at step t. This only happens if $\{u_t, v_t\}$ is a 2-cycle of $BR(t-1)$.

A cycle is formed in BS at step $t+1$ of phase 1 if and only if one of the existing $w_t^{(B)}$ edges of end-type $\{\gamma,\gamma\}$ in $BR(t)$ was chosen as $\{u_{t+1},v_{t+1}\}$. This has probability $w_t^{(B)}/\binom{\gamma_0-2t}{2} = 9(1+o(1))\,w_t^{(B)}/2(n-3t)^2$. Therefore

$$\mathbf{P}(\text{cycle}^{(B)}(t+1)\mid \mathcal{O}_t^{(B)}) = \frac{3}{2(4n-3t)} + o(n^{-1}). \tag{20}$$

Similarly with B replaced by R. Now $\mathbf{P}(\text{both}(t+1))$ depends on the number $\kappa_2(BR(t))$ of 2-cycles in $BR(t)$.

Lemma 5.3 *Assume that $\kappa_2(BR)\le 5\log n$, where B and R are fixed perfect matchings with parameters x_1, x_3 respectively. Then for some constant C_2,*

$$\mathbf{P}(\kappa_2(BR(t)) > C_2\log n) = O(n^{-2}),$$

for $1\le t\le T_1$.

Proof. At each step $i\le t$ and any fixed u_i, there are at most three γ-vertices v_i which may create at least one 2-cycle in $BR(t+1)$. (There are at most two such vertices if u_i belongs to a cycle of length at least 8, one if the cycle length is 6, three if it is 4, none if it is 2. Each such vertex creates one 2-cycle unless the cycle length is 6, in which case two 2-cycles would be created.) Since there are at least $x_3\sim n/2$ γ-vertices remaining, it is easy to define independent random variables τ_i with

$$\mathbf{P}(\tau_i=1) = 1-\mathbf{P}(\tau_i=0) = 3/x_3 \sim 6/n$$

such that the number of 2-cycles created at step i is at most $2\tau_i$. Then $Y=\sum_{i=1}^t \tau_i$ is binomially distributed with expected value asymptotic to $6t/n\le 6T_1/n\sim 1/2$. Hence by standard bounds

$$\mathbf{P}(\kappa_2(BR(t)) > C_2\log n) \le \mathbf{P}\Big(2\sum_{i=1}^{t}\tau_i > C_2\log n\Big) = O(n^{-2}).$$

□

Define the event $\mathcal{L}_t=\{\kappa_2(BR(t))\le C_2\log n\}$. Since both$(t+1)$ requires $\{u_{t+1},v_{t+1}\}$ to be a 2-cycle in $BR(t)$, we have

$$\mathbf{P}(\text{both}(t+1)\mid \mathcal{O}_t\wedge\mathcal{L}_t) = O(\log n/n^2) = o(n^{-1})$$

for $0\le t<T_1$, by Lemma 5.3. Moreover,

$$\begin{aligned}\mathbf{P}(\text{cycle}^{(B)}(t+1)\mid \mathcal{O}_t\wedge\mathcal{L}_t) &= \mathbf{P}(\text{cycle}^{(B)}(t+1)\mid \mathcal{O}_t)(1-O(n^{-2}))+O(n^{-2})\\ &= \frac{3}{2(4n-3t)}+o(n^{-1})\end{aligned}$$

by equation (20) (and similarly with B replaced by R). Therefore

$$\begin{aligned}
&\mathbf{P}(\text{cycle}(t+1) \mid \mathcal{O}_t \wedge \mathcal{L}_t) \\
&\quad = \mathbf{P}(\text{cycle}^{(B)}(t+1) \mid \mathcal{O}_t^{(B)} \wedge \mathcal{L}_t) + \mathbf{P}(\text{cycle}^{(R)}(t+1) \mid \mathcal{O}_t^{(R)} \wedge \mathcal{L}_t) \\
&\qquad\qquad - \mathbf{P}(\text{both}(t+1) \mid \mathcal{O}_t \wedge \mathcal{L}_t) \\
&\quad = \frac{3}{4n-3t} + o(n^{-1}).
\end{aligned}$$

For any event $\mathcal{F}_t$ depending only on steps $1, \ldots, t$ and such that $\mathcal{O}_t \wedge \mathcal{L}_t \wedge \mathcal{F}_t \neq \emptyset$, the same argument gives

$$\mathbf{P}\Big(\text{cycle}(t+1) \Big| \mathcal{O}_t \wedge \mathcal{L}_t \wedge \mathcal{F}_t\Big) = \frac{3}{4n-3t} + o(n^{-1}). \tag{21}$$

Let cycle(phase 1) be the event that a cycle is created at some step of phase 1. For the events $\mathcal{E}_s = \neg\text{cycle}(s)$,

$$\mathbf{P}(\neg\text{cycle}(\text{phase } 1)) = \mathbf{P}\Big(\bigwedge_{t=0}^{T_1-1} \mathcal{E}_{t+1}\Big) = \prod_{t=0}^{T_1-1} \mathbf{P}\Big(\mathcal{E}_{t+1} \mid \bigwedge_{s=1}^{t} \mathcal{E}_s\Big).$$

Let $\mathcal{F}_t = \bigwedge_{s=1}^{t} \mathcal{E}_s$ for $t \geq 0$. Then

$$\begin{aligned}
\mathbf{P}(\mathcal{E}_{t+1} \mid \mathcal{F}_t) &= \mathbf{P}(\mathcal{E}_{t+1} \mid \mathcal{O}_t \wedge \mathcal{L}_t \wedge \mathcal{F}_t))\mathbf{P}(\mathcal{O}_t \wedge \mathcal{L}_t \mid \mathcal{F}_t) \\
&\quad + \mathbf{P}(\mathcal{E}_{t+1} \mid \neg(\mathcal{O}_t \wedge \mathcal{L}_t) \wedge \mathcal{F}_t)\mathbf{P}(\neg(\mathcal{O}_t \wedge \mathcal{L}_t) \mid \mathcal{F}_t).
\end{aligned}$$

For any fixed u_{t+1} there are at most three γ-vertices in $BR(t)$ which, together with u_{t+1} create a cycle. Since there are at least $x \sim n/2$ γ-vertices remaining at any step of phase 1, we have

$$\mathbf{P}(\mathcal{F}_t) \geq (1 - 3/x)^t > e^{-1} + o(1),$$

and so, by the statement at (19),

$$\mathbf{P}(\neg(\mathcal{O}_t \wedge \mathcal{L}_t) | \mathcal{F}_t) < e\, \mathbf{P}(\neg(\mathcal{O}_t \wedge \mathcal{L}_t)) = O(n^{-2}).$$

It follows from all this that

$$\begin{aligned}
\mathbf{P}(\mathcal{E}_{t+1} \mid \mathcal{F}_t) &= \mathbf{P}(\mathcal{E}_{t+1} \mid \mathcal{O}_t \wedge \mathcal{L}_t \wedge \mathcal{F}_t)\left(1 - O(n^{-2})\right) + O(n^{-2}) \\
&= \mathbf{P}(\mathcal{E}_{t+1} \mid \mathcal{O}_t \wedge \mathcal{L}_t)(1 - O(n^{-2})) + O(n^{-2}) \\
&= 1 - \frac{3}{4n-3t} + o(n^{-1})
\end{aligned}$$

using (21). Therefore

$$\begin{aligned}
\mathbf{P}(\neg\text{cycle(phase 1)} &= \prod_{t=0}^{T_1-1} \mathbf{P}(\mathcal{E}_{t+1} \mid \mathcal{F}_t) \\
&= \prod_{t=0}^{T_1-1} \left(1 - \frac{3}{4n-3t} + o(n^{-1})\right) \\
&\sim \exp\left(-\sum_{t=0}^{T_1-1} \frac{3}{4n-3t}\right) \\
&\sim \exp\left(-\int_{s=0}^{T_1/n} \frac{3}{4-3s}\, ds\right) \\
&\sim \frac{15}{16}, \qquad (22)
\end{aligned}$$

using the fact that $T_1 \sim n/12$.

5.2 Phase 2

For $1 \le t \le T_2$, write $\widehat{BR}(t) = BR(T_1 + t)$. This is the same as restarting the clock at the start of phase 2. Let $\widehat{\gamma}_t$ be the number of γ-vertices in $\widehat{BR}(t)$, for $0 \le t \le T_2$. Similarly, let δ_t denote the number of δ-vertices in $\widehat{BR}(t)$, for $0 \le t \le T_2$. Since we delete one γ-vertex and one δ-vertex at each step of phase 1, we have $\widehat{\gamma}_t = \widehat{\gamma}_0 - t$ and $\delta_t = \delta_0 - t$ for $0 \le t \le T_2$.

Note that $\widehat{\gamma}_0 = \gamma_{T_1} = n/2 + O(n^{3/5})$ and $\delta_0 = a_\delta = 2n/3 + O(n^{3/5})$.

Now let $x_t^{(B)}$ be the number of B-edges of type $\{\gamma, \delta\}$ in $\widehat{BR}(t)$, for $0 \le t \le T_2$. Now $x_0^{(B)} = \gamma_0 - 2w_{T_1}^{(B)}$, which depends on the outcome of Phase 1.

Note that $x_{t+1}^{(B)} = x_t^{(B)} - 1$, except when the B-neighbour of u_{t+1} is coloured γ and the B-neighbour of v_{t+1} is coloured δ, in which case $x_{t+1}^{(B)} = x_t^{(B)} + 1$. Thus $|x_t^{(B)}|/n$ and $|x_{t+1}^{(B)} - x_t|$ are both bounded above by a constant, as required for Theorem 5.2. Let $\widehat{\Gamma}_0 = \gamma_{T_1}/n$ and $\Delta_0 = \delta_0/n$, and let $H_t = (\widehat{BR}(0), \widehat{BR}(1), \dots, \widehat{BR}(t))$ be the history of the first t steps of phase 2. Since there are $\widehat{\gamma}_0 - t - x_t^{(B)}$ γ-vertices incident with B-edges of type $\{\gamma, \gamma\}$, and $\delta_0 - t - x_t^{(B)}$ such δ-vertices, the above discussion shows that

$$\begin{aligned}
\mathbf{E}(x_{t+1}^{(B)} - x_t^{(B)} \mid H_t) &= \frac{2(\widehat{\gamma}_0 - t - x_t^{(B)})(\delta_0 - t - x_t^{(B)})}{(\widehat{\gamma}_0 - t)(\delta_0 - t)} - 1 \\
&= f(t/n, x_t^{(B)}/n) + \lambda_1
\end{aligned}$$

where

$$f(s,z) = \frac{2(1-2s-2z)(2-3s-3z)}{(1-2s)(2-3s)} - 1$$

and $\lambda_1 = O(n^{-2/5})$ provided $t/n < 1/2 - \eta$ for some positive constant η. Define the open set D by

$$D = \{(s,z) \mid -\eta < s < 1/2 - \eta,\ -\eta < z < 1/2 + \eta\}.$$

We may now argue as for phase 1. Solving $z'(s) = f(s,z)$ with initial condition $z(0) = z_0$ gives

$$z(s) = \frac{(1-2s)(2-3s)(2z_0 + 2s - 7z_0 s)}{2(2 - 6s^2 - 12z_0 s + 21z_0 s^2)}.$$

By Theorem 5.2, this solution is unique in D for all z_0 such that $(0, z_0) \in D$, and extends to points arbitrarily close to the boundary of D. Now let $z_0 = x_0^{(B)}/n$. By Theorem 5.2, we conclude that, with probability $1 - O(n^{-2})$,

$$\begin{aligned} x_t^{(B)} &= n\, z(t/n) + o(n) \\ &= \frac{(n-2t)(2n-3t)(2x_0^{(B)}n + 2tn - 7x_0^{(B)}t)}{2(2n^3 - 6t^2 n - 12x_0^{(B)}tn + 21x_0^{(B)}t^2)} + o(n). \end{aligned} \tag{23}$$

The argument as in phase 1 shows that this applies for $0 \le t \le n/2$ (after letting η tend to zero slowly), and that the corresponding result applies also to $x_t^{(R)}$, defined analogously.

Now for $1 \le \tau \le T_2$, define the event $\mathcal{O}^{(B)}_{T_1+\tau}$ to be the event that $\mathcal{O}^{(B)}_{T_1}$ holds *and* that equation (23) holds for $1 \le t \le \tau$. This is an event over the set of all possible $(T_1+\tau)$-step histories of the process, and says that the variables $w_t^{(B)}$ were all well-behaved in phase 1, and the variables $x_t^{(B)}$ are well-behaved for $1 \le t \le \tau$. Define $\mathcal{O}^{(R)}_{T_1+\tau}$ similarly. As in phase 1, define

$$\mathcal{O}_{T_1+\tau} = \mathcal{O}^{(B)}_{T_1+\tau} \wedge \mathcal{O}^{(R)}_{T_1+\tau}$$

for $1 \le \tau \le T_2$. We know that $\mathcal{O}^{(B)}_{T_1+T_2}$ and $\mathcal{O}^{(R)}_{T_1+T_2}$ both hold with probability $1 - O(n^{-2})$. Then also $\mathcal{O}_{T_1+T_2}$ holds with probability $1 - O(n^{-2})$. Clearly if $\mathcal{O}_{T_1+T_2}$ holds then so do $\mathcal{O}_{T_1+\tau}$, $\mathcal{O}^{(B)}_{T_1+\tau}$ and $\mathcal{O}^{(R)}_{T_1+\tau}$ for $1 \le \tau \le T_2$.

As in phase 1, define $\text{cycle}^{(B)}(T_1+t)$ to be the event that a cycle in BS is formed at step t of phase 2, and similarly $\text{cycle}^{(R)}(T_1+t)$, $\text{cycle}(T_1+t)$ and $\text{both}(T_1+t)$. A cycle is created in BS at step T_1+t+1 if and only if the edge $\{u_{T_1+t+1}, v_{T_1+t+1}\}$ chosen is identical to an existing B-edge in $BR(T_1+t)$ of type $\{\gamma, \delta\}$, for $0 \le t < T_2$. The probability of this is $x_t^{(B)}/(\widehat{\gamma}_0 - t)(\delta_0 - t)$. If $\mathcal{O}^{(B)}_{T_1}$ holds then, using (19) and the fact that $T_1 \sim n/12$, we have

$$x_0^{(B)} = \widehat{\gamma}_0 - 2w^{(B)}_{T_1} \sim \frac{n}{2} - \frac{(n-3T_1)^2}{3(4n-3T_1)} \sim \frac{2n}{5}.$$

Therefore

$$\begin{aligned} \mathbf{P}(\text{cycle}^{(B)}(T_1+t+1) \mid \mathcal{O}^{(B)}_{T_1+t}) &= \frac{x_t^{(B)}}{(\widehat{\gamma}_0 - t)(\delta_0 - t)} \\ &= \frac{6(n-t)}{5n^2 - 12nt + 6t^2} + o(n^{-1}), \end{aligned}$$

by (23) and using the fact that $\mathcal{O}^{(B)}_{T_1}$ holds. (Similarly with B replaced by R.)

Define $\mathcal{L}_{T_1+t}$ as for phase 1, but with the constant C_2 chosen so that the argument in Lemma 5.3 gives $\mathbf{P}(\mathcal{L}_{T_1+t}) = 1 - O(n^{-2})$ for $0 \le t < T_2$. We have

$$\begin{aligned}
& \mathbf{P}(\mathrm{cycle}^{(B)}(T_1+t+1) \mid \mathcal{O}^{(B)}_{T_1+t} \wedge \mathcal{L}_{T_1+t}) \\
= \;& \mathbf{P}(\mathrm{cycle}^{(B)}(T_1+t+1) \mid \mathcal{O}^{(B)}_{T_1+t})(1+O(n^{-2})) + O(n^{-2}) \\
= \;& \frac{6(n-t)}{5n^2-12nt+6t^2} + o(n^{-1}).
\end{aligned}$$

The same statement holds with B replaced by R throughout. Let $\mathcal{F}_{T_1+t} = \bigwedge_{s=1}^{T_1+t} \mathcal{E}_s$, where $\mathcal{E}_s = \neg\mathrm{cycle}(s)$, as defined in phase 1. Then, arguing as in phase 1 we obtain

$$\begin{aligned}
& \mathbf{P}(\mathcal{E}_{T_1+t+1} \mid \mathcal{F}_{T_1+t}) \\
& \quad = \; 1 - 2\,\mathbf{P}\Big(\mathrm{cycle}^{(B)}(T_1+t+1) \mid \mathcal{O}^{(B)}_{T_1+t} \wedge \mathcal{L}_{T_1+t}\Big) + O\left(\frac{\log n}{n^2}\right) \\
& \quad = \; 1 - \frac{12(n-t)}{5n^2-12nt+6t^2} + o(n^{-1})
\end{aligned}$$

for $0 \le t < T_2$.

Let cycle(phase 2) be the event that a cycle was created in phase 2. Then, arguing as in phase 1 we find

$$\mathbf{P}(\neg\mathrm{cycle}(\text{phase } 2)) = \prod_{t=0}^{T_2-1} \mathbf{P}(\mathcal{E}_{T_1+t+1} \mid \mathcal{F}_{T_1+t}) \sim \frac{1}{10} \tag{24}$$

using $T_2 \sim n/2$.

5.3 Phase 3 and conclusion

At the end of phase 1, there are

$$\delta_{T_1} = \delta_0 - T_1 \sim \frac{n}{6}$$

vertices remaining, all coloured δ. We will apply the following result for $m \sim n/12$. (This is implicitly proved in [5], from Lemma 3 and the argument on pages 42 and 43, so a separate proof is omitted here.)

Lemma 5.4 *For given perfect matchings B and R of the same set of $2m$ vertices such that $\kappa(BR) \le m/40$*

$$\mathbf{P}_S(BS, SR \in \mathrm{HAM}) = \Big(1 + O\Big(\frac{\kappa_2(BR)+1}{m}\Big)\Big)\frac{\pi}{4m}.$$

We first have to show that with high probability, the number of cycles at the end of phase 2 is small enough.

Lemma 5.5 *Assume $\kappa(BR) \le 5\log n$. With probability $1-o(n^{-1})$, $\kappa(BR(T_1+T_2)) < n/1000$.*

Proof. We examine the number of cycles of length at most 1000 in $BR(t)$, following the proof of Lemma 5.3 for cycles of length 2. At each step i in phase 1, there are at least cn possible vertices v_i to choose, and at most 2002 of these can increase the number of cycles of length at most 1000. (These are the ones in the same cycle as u_i and of distance at most 1001 from it around the cycle.) The same holds in phase 2, if u_i is chosen from the γ-vertices and then v_i from the δ-vertices. Hence at each step in phases 1 and 2, the probability of increasing the number of cycles of length at most 1000 is $O(1/n)$. As in the proof of Lemma 5.3, it follows that with probability $1-O(n^{-2})$, the graph $BR(T_1+T_2)$ has $O(\log n)$ cycles of length at most 1000. All other cycles have at least 1001 vertices, so this implies that $\kappa(BR(T_1+T_2)) < n/1000$ for sufficiently large n. ◻

For the event $\mathcal{L}_{T_1+T_2} = \{\kappa_2(BR(T_1+T_2)) \le m/40\}$, Lemma 5.4 gives

$$\mathbf{P}(\neg\text{cycle(phase 3)} \mid \mathcal{O}_{T_1+T_2} \wedge \mathcal{L}_{T_1+T_2}) \sim \mathbf{P}(\neg\text{cycle(phase 3)} \mid \mathcal{L}_{T_1+T_2}) \sim \frac{3\pi}{n}$$

as $\mathbf{P}(\mathcal{O}_{T_1+T_2}) = 1 - O(n^{-2})$ and $m \sim n/12$. Since $m \sim n/12$ yields $\mathbf{P}(\mathcal{L}_{T_1+T_2}) = 1 - o(n^{-1})$ by Lemma 5.5, we conclude that

$$\mathbf{P}(\neg\text{cycle(phase 3)}) \sim \frac{3\pi}{n}. \tag{25}$$

Multiplying (22), (24) and (25) together, we find that

$$\mathbf{P}_{(x)}(BS, RS \in \text{HAM}) \sim \frac{15}{16} \cdot \frac{1}{10} \cdot \frac{3\pi}{n} = \frac{9\pi}{32n},$$

establishing (18). This holds for all fixed matchings B and R with parameter x_1, x_3 respectively, such that $\kappa(BR) \le 5\log n$. This and Lemma 5.2 imply Theorem 5.1.

6 Finding the maximum

The aim of this section is to show that the function f, defined in (13), has a unique maximum in the domain D defined (14), and this maximum satisfies (15).

The partial derivative of $\log f$ with respect to χ_i is equal to zero if and only if

$$\alpha(2\alpha\kappa + 4\beta\kappa - 2\alpha\beta - \alpha^2) = \chi_i(\alpha^2 + 4\beta\kappa)$$

for $i = 1, 3$, or

$$\alpha(\alpha^2 - 2\alpha - 2\alpha\kappa + 2\alpha\beta + 4\kappa - 4\beta\kappa) = \chi_i(\alpha^2 - 4\alpha\kappa + 4\kappa - 4\beta\kappa)$$

for $i = 2, 4$. Assume for the moment that $\alpha^2 + 4\beta\kappa \ne 0$ and $\alpha^2 - 4\alpha\kappa + 4\kappa - 4\beta\kappa \ne 0$. Then we have a unique solution for each χ_i. After substituting these values, we find that the partial derivatives of $\log f$ with respect to κ, α and β are equal to zero if and only if

$$\begin{aligned}
16(1-\kappa-\beta)((\beta+\alpha-\kappa) &= (2\kappa-\alpha)^2, && (26)\\
(\alpha+\beta-1)(2\beta+\alpha)^2(\alpha-2\kappa) &= (\beta+\alpha-\kappa)(2\beta-2+\alpha)^2\alpha, && (27)\\
(\alpha+\beta-1)(\beta+\kappa-1)(2\beta+\alpha)^4 &= \beta(\beta+\alpha-\kappa) && \\
&\quad \times(4\alpha\beta-4\alpha+4-8\beta+4\beta^2+\alpha^2)^2, && (28)
\end{aligned}$$

respectively. We can rewrite (28) as

$$\begin{aligned}(1-\alpha-2\beta)(&16\beta^4+16\beta^4\kappa-32\beta^3\kappa+32\beta^3\alpha\kappa+32\beta^3\alpha-32\beta^2\alpha\kappa+32\beta^2\kappa\\ &+16\beta^2+24\beta^2\alpha^2\kappa+16\alpha^2\beta^2-48\beta^2\alpha+8\alpha^3\beta\kappa-16\beta\kappa-8\alpha^2\beta\kappa\\ &+16\alpha\beta\kappa+16\alpha\beta-16\alpha^2\beta-\alpha^4+\alpha^4\kappa)=0\end{aligned} \tag{29}$$

Suppose first that $1-\alpha-2\beta=0$, i.e., $\alpha=1-2\beta$. Substituting this expression for α into (27) gives $\kappa=1-2\beta$. Finally, substituting all these expressions into (26) gives $\beta=\frac{1}{6}$. This yields the solution

$$\kappa=\frac{2}{3},\ \alpha=\frac{2}{3},\ \beta=\frac{1}{6},\ \chi_1=\frac{1}{2},\ \chi_2=\frac{1}{2},\ \chi_3=\frac{1}{2},\ \chi_4=\frac{1}{2}$$

with $f(\frac{2}{3},\frac{2}{3},\frac{1}{6},\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2})=\frac{81}{16}$. We will show that this is the global maximum.

First suppose that (29) equals zero but $1-\alpha-2\beta\neq 0$. Then the long factor must equal zero. Note that this factor is linear in κ, and its coefficient of κ is equal to

$$\alpha^4+16\alpha\beta-8\alpha^2\beta-16\beta+8\alpha^3\beta+24\alpha^2\beta^2+32\beta^2-32\alpha\beta^2+32\alpha\beta^3-32\beta^3+16\beta^4.$$

If this expression is nonzero then we have a unique solution for κ. Substituting this into (26) gives $\alpha=2-2\beta$ as the only nonnegative real solution. Substituting this into the expression for κ gives $\kappa=1-\beta$. Now $\alpha=2\kappa$, which means that all type α vertices are coloured δ (see Section 2). Therefore there can be no bichromatic edges in any matching on these vertices, and so $\chi_i=0$ for $1\leq i\leq 4$. Using all this, we find that

$$\frac{\partial\log f}{\partial\beta}=2\log(1-\beta)-2\log(2)-\log(-1+\beta)-\log(\beta)$$

which, assuming $\beta\neq 1$, is only zero when $\beta=-1$. This is outside our range, by definition of D. If $\beta=1$ then f has the value $\frac{9}{4}$, which is not maximum.

Next suppose that $\alpha^2+4\beta\kappa=0$. Then α, β and κ must all equal zero, which forces $\chi_i=0$ for all i. Here the value of f is $\frac{9}{4}$, which is not maximum. Finally, suppose that $\alpha^2-4\beta\kappa-4\alpha\kappa+4\kappa=0$. If $\kappa\neq 0$ then this gives $\beta=1-\alpha+\alpha^2/(4\kappa)$. But $c_\delta=n(1-\alpha-\beta)=-n\alpha^2/(4\kappa)$, which is only nonnegative when $\alpha=0$. This forces $\beta=1$ and $\chi_i=0$ for $1\leq i\leq 4$ (since $\chi_i\leq\alpha$ for $1\leq i\leq 4$). Here

$$\frac{\partial\log f}{\partial\kappa}=\log(1-\kappa)+2\log(2)-2\log(\kappa)+\log(-\kappa)$$

which is not equal to zero for any κ such that $0\leq\kappa\leq 1$. Hence there is no local maximum here. However, if $\alpha^2-4\beta\kappa-4\alpha\kappa+4\kappa=0$ and $\kappa=0$ then $\alpha=0$, which forces $\chi_i=0$ for $1\leq i\leq 4$. Here

$$\frac{\partial\log f}{\partial\beta}=2\log(\beta)-2\log(1-\beta)$$

which is only zero when $\beta=\frac{1}{2}$. But $f(0,0,\frac{1}{2},0,0,0,0)=\frac{9}{16}$, which is a local minimum.

Hence there are no internal points which give a greater value of f than the claimed maximum. It remains to check that the same is true for all points on the boundary of D. This is routine, though tedious, and we omit the details. We merely note that the following idea helps (formalised in [2, Lemma 12]).

Ignoring constants and powers of constants, we can write $\log f$ as $\sum_{i \in I} c_i U_i / \log(U_i)$. Consider the derivative of $\log f$ along a path which approaches a path y on the boundary from within D. Unless $U_i = 0$ at y for some i such that $c_i > 0$, this derivative turns out to be $-\infty$, and thus such boundary points cannot determine a global maximum of f on D.

References

[1] A. Frieze, M. Jerrum, M. Molloy, R. Robinson and N. Wormald, *Generating and counting Hamilton cycles in random regular graphs*, Journal of Algorithms **21** (1996), pp. 176–198.

[2] H. Garmo, *Random Railways and Cycles in Random Regular Graphs.* PhD Thesis, Uppsala University, Sweden, 1998.

[3] S. Janson, *Random regular graphs: asymptotic distributions and contiguity*, Combinatorics, Probability and Computing **4** (1995), pp. 369–405.

[4] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, John Wiley and Sons, New York (2000).

[5] J. H. Kim and N. C. Wormald, *Random matchings which induce Hamilton cycles, and hamiltonian decompositions of random regular graphs*, Journal of Combinatorial Theory (Series B) **81** (2001), pp. 20–44.

[6] J. McCammond and D. Wise, *Locally convex small cancellation groups*, (Preprint, 2001).

[7] M. S. O. Molloy, H. Robalewska, R. W. Robinson and N. C. Wormald, *1-Factorisations of random regular graphs*, Random Structures and Algorithms **10** (1997), pp. 305–321.

[8] R. W. Robinson and N. C. Wormald, *Existence of long cycles in random cubic graphs*, in *Enumeration and design* (D. M. Jackson and S. A. Vanstone, eds.), Academic Press, Toronto (1984), pp 251–270.

[9] R. W. Robinson and N. C. Wormald, *Almost all cubic graphs are hamiltonian*, Random Structures and Algorithms **3** (1992), pp. 117–125.

[10] R. W. Robinson and N. C. Wormald, *Almost all regular graphs are hamiltonian*, Random Structures and Algorithms **5** (1994), pp. 363–374.

[11] N. C. Wormald, *Differential equations for random processes and random graphs*, Annals of Applied Probability **5** (1995), pp. 1217–1235.

[12] N. C. Wormald, *Models of random regular graphs*, in *Surveys in Combinatorics 1999* (J. D. Lamb and D. A. Preece, eds.), vol. 267 of LMS Lecture Note Series, Cambridge University Press, Cambridge, 1999, pp. 239–298.

[13] N. C. Wormald, *The differential equation method for random graph processes and greedy algorithms*, in *Lectures on Approximation and Randomized algorithms* (M. Karoński and H. J. Prömel, eds.), PWN, Warsaw, 1999, pp. 73–155.