Random matchings which induce Hamilton cycles, and hamiltonian decompositions of random regular graphs

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 VIC 3010
Australia
nick@ms.unimelb.edu.au

Abstract

Select four perfect matchings of $2n$ vertices, independently at random. We find the asymptotic probability that each of the first and second matchings forms a Hamilton cycle with each of the third and fourth. This is generalised to embrace any fixed number of perfect matchings, where a prescribed set of pairs of matchings must each produce Hamilton cycles (with suitable restrictions on the prescribed set of pairs). We also show how the result with four matchings implies that a random d-regular graph for fixed even $d \geq 4$ asymptotically almost surely decomposes into $d/2$ Hamilton cycles. This completes a general result on the edge-decomposition of a random regular graph into regular spanning subgraphs of given degrees together with Hamilton cycles, and verifies conjectures of Janson and of Robinson and Wormald.

1 Introduction

Our main concern in this paper is to address two questions. The first is whether a random d-regular graph for even d asymptotically almost surely (a.a.s.) has an edge-decomposition into Hamilton cycles. Here and in all our asymptotic statements, we take $n \to \infty$, with n restricted to even integers in statements about d-regular graphs if d is odd. Related to this, we also consider algorithms for finding these edge-decompositions in random regular graphs. The second question can be phrased in general as follows: given a set of randomly generated perfect matchings of an even number of vertices, what is the probability that each of a prescribed set of pairs of those matchings induces a connected graph (that is, a Hamilton cycle) on those vertices? This question was actually motivated by the first, although the relationship between the two questions will only appear in the proof of Theorem 1 below.

*Research supported in part by the Australian Research Council and performed in part while visiting Microsoft Research.

1.1 Hamiltonian decompositions and contiguity

By a *complete hamiltonian decomposition* of a d-regular graph G we mean a set of $\lfloor d/2 \rfloor$ edge-disjoint Hamilton cycles of G. If d is odd, the left-over edges must form a perfect matching. Let $\mathcal{G}_{n,d}$ denote the uniform probability space whose elements are the d-regular graphs on the vertex set $[n] = \{1, \ldots, n\}$. Conjecture 1 in [9] was essentially that if $d \geq 4$ then $G \in \mathcal{G}_{n,d}$ a.a.s. has a complete hamiltonian decomposition. (Here we say "essentially" because of the unimportant restriction to even n in [9].) It is already known (see [11, Section 4]) that this conjecture follows from the case $d = 4$, which is given explicitly as Corollary 1 below.

We prove Corollary 1 using the small subgraph conditioning method, which originated in the papers of Robinson and Wormald [8, 9], from which the general method was set forth by Molloy et al. [6] and by Janson [4]; see [11] for a full development. This method is of use in certain cases when most of the variance of the random variable of interest is caused by the influence of a set of other random variables. In all cases where this method has been applied, the variable of interest has been a count of large subgraphs (such as Hamilton cycles) and the other variables have been counts of small subgraphs (such as cycles of given fixed lengths).

The results we obtain are best described in terms of contiguity. Let $\mathcal{F}_n$ and $\hat{\mathcal{F}}_n$ be two discrete probability spaces which have the same underlying set for each $n \geq 1$. We say that the two sequences of spaces $(\mathcal{F}_n)$ and $(\hat{\mathcal{F}}_n)$ are *contiguous* if any sequence of events A_n $(n \geq 1)$ is true a.a.s. in $(\mathcal{F}_n)$ if and only if it is true a.a.s. in $(\hat{\mathcal{F}}_n)$. In this case we write $\mathcal{F}_n \approx \hat{\mathcal{F}}_n$ and by a slight abuse of notation say that $\mathcal{F}_n$ and $\hat{\mathcal{F}}_n$ are contiguous. As in [11], if $\mathcal{F}$ and $\hat{\mathcal{F}}$ are probability spaces of random d-regular graphs or multigraphs on the same vertex set, we define their *sum* $\mathcal{F} + \hat{\mathcal{F}}$ to be the space whose elements are defined by the random multigraph $G \cup \hat{G}$ where $G \in \mathcal{F}$ and $\hat{G} \in \hat{\mathcal{F}}$ are generated independently. Similarly, define the *graph-restricted sum* of $\mathcal{F}$ and $\hat{\mathcal{F}}$, denoted by $\mathcal{F} \oplus \hat{\mathcal{F}}$, to be the space which is the restriction of $\mathcal{F} + \hat{\mathcal{F}}$ to simple graphs (i.e. with no multiple edges). In order to maintain identical underlying sets for spaces which are to be related, the sum space $\mathcal{F} + \hat{\mathcal{F}}$ is extended to include all d-regular multigraphs on the same vertex set, with all multigraphs not already appearing being given probability 0. Similarly, $\mathcal{F} \oplus \hat{\mathcal{F}}$ is extended to include the underlying set of $\mathcal{G}_{n,d}$. Defining $\mathcal{H}_n$ to be a random Hamilton cycle on the same vertex set as $\mathcal{G}_{n,d}$, we can now state the following.

Theorem 1 $\mathcal{H}_n \oplus \mathcal{H}_n \approx \mathcal{G}_{n,4}$.

Since the event of having two edge-disjoint Hamilton cycles occurs with probability 1 in $\mathcal{H}_n \oplus \mathcal{H}_n$, this theorem immediately implies the following.

Corollary 1 *$G \in \mathcal{G}_{n,4}$ a.a.s. has two edge-disjoint Hamilton cycles.* ▮

Whenever it works, the small subgraph conditioning method gives the asymptotic distribution of the numbers of graph decompositions concerned. Thus, the number of pairs of edge-disjoint Hamilton cycles in $G \in \mathcal{G}_{n,4}$ is asymptotically the exponential of a linear combination of infinitely many independent Poisson variables (see [4, Theorem 1] or [11, Theorem 4.3]).

We approach random d-regular graphs by using the standard pairing model (see Bollobás [1]). Consider dn points, with d points in each of n buckets, and take a random pairing of all the points. We call this uniform probability space $\mathcal{P}_{n,d}$. Letting the buckets be vertices and each pair represent an edge (which joins the buckets containing its two vertices), we obtain a random regular pseudograph (possibly having loops and multiple edges). Denote this probability space by $\mathcal{G}^*_{n,d}$. Graphs with no loops or multiple edges occur with equal probabilities, so the restiction of $\mathcal{G}^*_{n,d}$ to (simple) graphs gives the probability space $\mathcal{G}_{n,d}$.

As was widely expected, our proof of Corollary 1 simultaneously gives a proof the following result, which answers affirmatively a conjecture of Janson [4, Conjecture 2]. For this we need to define the probability space $\mathcal{G}'_{n,4}$ to be the restriction of $\mathcal{G}^*_{n,4}$ to loopless multigraphs.

Theorem 2

$$\mathcal{H}_n + \mathcal{H}_n \approx \mathcal{G}'_{n,4}.$$

Contiguity is clearly an equivalence relation. So by Theorem 2 we may deduce contiguity between $\mathcal{H}_n + \mathcal{H}_n$ and the uniform space of d-regular multigraphs. The intermediate contiguity is given by [4, Theorem 12].

We next show how the desired decomposition of random d-regular graphs follows easily from Theorem 1 and previously known results. (The corresponding multigraph version also follows in the same way, by the analogous results for multigraphs.)

Corollary 2 *If $d \geq 4$ is even then $G \in \mathcal{G}_{n,d}$ a.a.s. has a complete hamiltonian decomposition.*

Proof. Let $k\mathcal{F}$ denote the graph-restricted sum of k copies of a random graph space $\mathcal{F}$. First consider the case that d is even. It was proved by Frieze et al. [2] that

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

Repeated use of this fact together with simple properties of contiguity (see [11, Lemma 4.14]) and Theorem 1 gives

$$\mathcal{G}_{n,d} \approx \mathcal{G}_{n,d-2} \oplus \mathcal{H}_n \approx \mathcal{G}_{n,d-4} \oplus 2\mathcal{H}_n \approx \cdots \approx \tfrac{d}{2}\mathcal{H}_n.$$

The corollary now follows, since a random element of the last space trivially has a complete hamiltonian decomposition with probability 1. For odd d, the conjecture was already known to be true by a similar argument using only the previously known results that $d\mathcal{G}_{n,1} \approx \mathcal{G}_{n,d}$ for $d \geq 3$, and that $\mathcal{G}_{n,3} \approx \mathcal{G}_{n,1} \oplus \mathcal{H}_n$. (See [11, Section 4.3] for details.) ∎

In [11, Corollary 4.17] it was shown that combining Theorem 1 with previously known contiguity results produces a general fact about graph-restricted sums of random regular graphs: provided $2j + \sum_{i=1}^{d-1} ik_i = d \geq 3$ for non-negative k_i,

$$\mathcal{G}_{n,d} \approx j\mathcal{H}_n \oplus k_1\mathcal{G}_{n,1} \oplus \cdots \oplus k_{d-1}\mathcal{G}_{n,d-1}.$$

This implies the existence a.a.s. of a decomposition of $G \in \mathcal{G}_{n,d}$ into given numbers of Hamilton cycles and regular factors of prescribed degrees. (This is not true for $d = 2$.)

1.2 Finding complete hamiltonian decompositions

In Section 3 we prove the following.

Theorem 3 *For fixed $d \geq 4$, there is a randomised algorithm which runs in polynomial time and which for $G \in \mathcal{G}_{n,d}$ a.a.s. finds a complete hamiltonian decomposition of G.*

Note that the a.a.s. here refers to the randomisation over G and not the randomised algorithm. The existence of the algorithm is proved by induction on d.

In [2] it was also shown that there is a polynomial time near uniform generator for the set of Hamilton cycles of $G \in \mathcal{G}_{n,d}$ a.a.s. (see Section 3) and also an FPRAS (fully polynomial randomised approximation scheme — see [2] for the definition). We conjecture that the same statements are true for complete hamiltonian decompositions rather than just Hamilton cycles.

1.3 Random perfect matchings

The main difficulty in using the small subgraph conditioning method is invariably the computation of the variance of the number of decompositions. For the case of decompositions into Hamilton cycles, we will eventually be interested in sets of matchings on the same vertex set. Here and in Section 4, we consider matchings on the set of vertices $[2n] = \{1, 2, , \ldots, 2n\}$. Given two such matchings A and B, we denote by AB the multigraph on that vertex set which has edge set $A \cup B$. We denote the number of components of a graph by κ, and note that AB is a union of $\kappa(AB)$ disjoint cycles. (Throughout this paper, if A and B have an edge in common, we regard that as a cycle of length 2 in the multigraph AB.) Let HAM denote the set of 2-regular connected graphs, i.e. those graphs whose edge set induces a Hamilton cycle on the vertex set. We say that A and B are *H-compatible* if $AB \in \text{HAM}$.

Let $h(n)$ denote the number of Hamilton cycles on n vertices and $m(n)$ the number of perfect matchings of n vertices. Then

$$h(2n) = \frac{(2n-1)!}{2}, \qquad m(2n) = \frac{(2n)!}{n!2^n}. \tag{1}$$

We are interested in the probability that a set of random (independent, uniformly distributed) perfect matchings on $[2n]$ are such that prescribed pairs of the matchings are H-compatible. It is easily computed from (1) that for two such random matchings, or equivalently for one fixed and one random, the probability that they are H-compatible is

$$p_{\mathrm{H}}(2n) \sim \sqrt{\pi/4n}. \tag{2}$$

When we consider one uniformly distributed random perfect matching S we denote probability in this space by $\mathbf{P}_S$. Our main result in Section 4 is the following.

Theorem 4 *For given perfect matchings B and R of the same set of vertices such that $\kappa(BR) \leq n/40$,*

$$\mathbf{P}_S(BS, SR \in \text{HAM}) = (1 + O((l+1)/n) + O((\kappa(BR) + \log n)/n^2))p_{\mathrm{H}}(2n)^2$$

where $l = |B \cap R|$.

We do not attempt to optimise the upper bound on $\kappa(BR)$ in the hypothesis. This is already far stronger than we need. For we will show that Theorem 4 has the following easy consequence. The special case of this to be used in Section 2 is when the graph G is just a 4-cycle. Let $E(G)$ denote the edge set of G.

Corollary 3 *Let $G = G(n)$ be a graph on r vertices where $r = r(n) = o\left(\sqrt{n/\log n}\right)$, and assume that for all n, G has no subgraph of minimum degree greater than 2. Let $M_1, \ldots, M_r$ be independent and uniformly distributed random perfect matchings of $[2n]$. Then*

$$\mathbf{P}(M_iM_j \in \mathrm{HAM} \quad \forall ij \in E(G)) = (1 + O(r^2 \log n/n))p_{\mathrm{H}}(2n)^m$$

where $m = |E(G)|$. Moreover, if r is fixed and G is an r-cycle, then

$$\mathbf{P}(M_iM_{i+1} \in \mathrm{HAM} \quad \forall i) = (1 + O(r \log n/n))p_{\mathrm{H}}(2n)^r.$$

2 Proof of Theorems 1.1 and 1.2

The small subgraph conditioning method calls for computation of expectation and variance of the number of decompositions under examination, as well as computation of joint moments of the numbers of decompositions and other variables which affect its distribution. For most applications so far, including the present one, these other variables are the numbers of short cycles.

It is often easiest to work directly with the random pairings in the space $\mathcal{P}_{n,d}$ defined in Section 1.1. It will be seen near the end of this section that calculations in $\mathcal{P}_{n,d}$ will imply what we need in Theorem 2 about $\mathcal{G}'_{n,4}$. Then Theorem 1 will be deduced at the end. So in this section we let G be a random pairing in $\mathcal{P}_{n,4}$. Define an *H-cycle* of G to be a set of pairs of G which induce a Hamilton cycle in the corresponding multigraph, and define an *H-decomposition* of G to be an ordered pair of disjoint H-cycles of G. Let $H_2(G)$ be the number of H-decompositions of G. (Although the H-cycles in the decomposition are disjoint, the multigraph corresponding to the pairing may have two parallel edges, one from each H-cycle.) Also for $k \geq 2$ define $C_k(G)$ to be the number of k-cycles of G (i.e. k-cycles of the corresponding multigraph). We will compute asymptotic values for $\mathbf{E}H_2$ and $\mathbf{E}(H_2^2)$ as well as the joint moments $\mathbf{E}(H_2[C_2]_{i_2} \cdots [C_j]_{i_j})$ for fixed j and $i_2, \ldots, i_j$. Here $[x]_j$ denotes $x(x-1)\cdots(x-j+1)$. As a starting point, note that the total number of pairings is

$$|\mathcal{P}_{n,4}| = \frac{(4n)!}{(2n)!2^{2n}} \sim \sqrt{2}\left(\frac{n}{e}\right)^{2n} 16^n. \tag{3}$$

(i) $\mathbf{E}H_2$

Select an H-decomposition of a pairing G as follows. First choose the adjacencies of the vertices in each of the two H-cycles (this determines all the edges of the corresponding multigraph), and then for each vertex choose one of the 4! ways to assign the four points to the four pairs involved. The whole pairing has now been determined. This shows that the number of H-decompositions of pairings is

$$h(n)^2(4!)^n \sim \frac{\pi}{2n}\left(\frac{n}{e}\right)^{2n} 24^n \tag{4}$$

by (1), and thus using (3)

$$\mathbf{E}H_2 \sim \frac{\pi}{\sqrt{8}n}\left(\frac{3}{2}\right)^n . \tag{5}$$

(ii) $\mathbf{E}(H_2^2)$

This is computed by counting pairings G for which there has been distinguished an ordered pair of H-decompositions. Each edge of G can be said to be type (i, j) if it is in the i'th H-cycle in the first H-decomposition and the j'th H-cycle in the second. So edges are of four types: $(1, 1)$, $(1, 2)$, $(2, 1)$, $(2, 2)$.

Vertices are now of three types: A if all four types of edge are incident, B two $(1, 1)$ and two $(2, 2)$ edges are incident, and C if two $(1, 2)$ and two $(2, 1)$ edges are incident. Let a, b and c denote the numbers of vertices of types A, B and C respectively.

Note that the edges of any particular type induce exactly $a/2$ paths which terminate only at the vertices of type A. In particular, a must be even. Some of these paths are just an edge between two A-type vertices, and some have other internal vertices of type B or C. However, the type of internal vertices does not change along the path, since paths of $(1, 1)$ edges and those of $(2, 2)$ edges use only the B-type internal vertices, and those of $(1, 2)$ edges and those of $(2, 1)$ edges each use only C-type internal vertices. Moreover, because of the existence of the four H-cycles, each B-type vertex must occur in a path of $(1, 1)$ edges and also in a path of $(2, 2)$ edges, each C-type vertex is similarly in a path of $(1, 2)$ edges and in a path of $(2, 1)$ edges, and the paths must induce four Hamilton cycles on the A-type vertices. In particular, the paths of $(i, 1)$ and $(i, 2)$ edges combine to give a Hamilton cycle on the A-type vertices (for $i = 1$ and 2), as do the paths of $(1, i)$ and $(2, i)$ edges. So we proceed by counting the connection schemes of the paths to the A-type vertices, and then multiplying by the number of ways of choosing which are the A-type vertices and also the number of ways to run through all the B and C type vertices (which is determined by b and c).

Let $S(a)$ denote the number of connection schemes of the paths to the A-type vertices, and take $a = 2k$. Then $S(a)$ is the number of quadruples (M_1, M_2, M_3, M_4) of perfect matchings of the type A vertices (for paths of edges of types $(1, 1)$, $(1, 2)$, $(2, 2)$ and $(2, 1)$ respectively), such that $M_1 \cup M_2$, $M_2 \cup M_3$, $M_3 \cup M_4$ and $M_4 \cup M_1$ are all Hamilton cycles. By Corollary 3 and (2),

$$S(2k) \sim (p_{\mathrm{H}}(2k)m(2k))^4 \sim \frac{1}{16}(2^k(k-1)!)^4 \tag{6}$$

as $k \to \infty$. (This behaviour of k can be assumed, as can be seen from the following analysis applied to the case of bounded k. Note that if $k = 0$ the connection scheme is empty, and it is easy to see from the following argument that the contribution from this case is negligible.)

After deciding on one such connection scheme, there are

$$\binom{n}{2k}$$

ways to choose the labels of the vertices of type A,

$$\binom{b+k-1}{k-1}^2$$

ways to assign the numbers of B type vertices to the paths of edges of type $(1,1)$ and similarly to those of type $(2,2)$, and

$$\binom{c+k-1}{k-1}^2$$

ways to assign the numbers of C type vertices to the paths of edges of type $(1,2)$ and $(2,1)$. At this point the scheme can be regarded as a set of paths between vertices of type A, with unlabelled beads on the paths marking how many vertices of type B and C is on each. But each vertex of type B appears once on each of the two types of paths $(1,1)$ and $(2,2)$; there are

$$b!$$

ways to make the identifications of the two beads for each vertex, and similarly

$$c!$$

for the type C vertices. Finally, there are

$$(n-2k)!$$

ways to label all the type B and C vertices (noting that each labelling produces a different H-decomposition of a pairing) and

$$(4!)^n$$

ways to decide the assignments of points in the vertices, as above. The total number of ordered pairs of H-decompositions of pairings with n vertices is the product of the expressions above, summed over k and b. Noting $c = n-2k-b$, this is

$$\begin{aligned}
&\sum_k\sum_b \frac{n!(4!)^n 16^k (k+b)!^2 (n-k-b)!^2}{16(k+b)^2(n-k-b)^2(2k)!b!(n-2k-b)!} \\
&\sim \sum_k\sum_b \frac{\sqrt{2}\,\pi(4!)^n}{16(k+b)(n-k-b)}\sqrt{\frac{n}{kb(n-2k-b)}}\left(\frac{n}{e}\right)^{2n} f(\beta,\kappa)^n
\end{aligned} \tag{7}$$

where

$$f(\kappa,\beta) = \frac{16^\kappa(\beta+\kappa)^{2(\beta+\kappa)}(1-\kappa-\beta)^{2(1-\kappa-\beta)}}{(2\kappa)^{2\kappa}\beta^\beta(1-2\kappa-\beta)^{1-2\kappa-\beta}}, \quad \kappa = k/n, \quad \beta = b/n.$$

Here we are assuming $b \to \infty$ and $n - 2k - b \to \infty$, which gives negligible error by the following analysis of this expression. The region of summation has these constraints as well as $k \to \infty$.

To find the main contribution to the summation in (7), set the partial derivatives of $\log f$ with respect to κ and β equal to 0. The resulting equations are

$$\begin{aligned}(\beta+\kappa)^2(1-\beta-2\kappa) &= \beta(1-\beta-\kappa)^2,\\ 2(\beta+\kappa)(1-2\kappa-\beta) &= \kappa(1-\kappa-\beta)\end{aligned}$$

which can immediately be combined, after squaring the second, to give

$$4(1-2\kappa-\beta)\beta = \kappa^2.$$

Eliminating $\beta(1-\beta)$ between this and the earlier equation gives

$$5\kappa^2 - 2\kappa(1-\beta) = 0,$$

with solutions $\kappa = 0$ (which is on the boundary of the region of interest) or $\kappa = \frac{2}{5}(1-\beta)$. From an earlier equation this yields the only solution not on the boundary: $(\kappa, \beta) = (\frac{1}{3}, \frac{1}{6})$. Since $f(\frac{1}{3}, \frac{1}{6}) = \frac{3}{2}$ and, as a routine check shows, all boundary points have strictly lower value than this, the only significant parts of the summation in (7) are for $k \sim n/3$ and $b \sim n/6$ respectively. Putting $\kappa = \frac{1}{3} + x$ and $\beta = \frac{1}{6} + y$ and expanding $\ln f$ about $x = y = 0$ yields

$$f(\kappa, \beta) = \frac{3}{2}\exp(-11x^2 - 4xy - 2y^2 + O(x^2y + xy^2))$$

for x and y sufficiently small. From here, a standard argument shows that

$$\begin{aligned}\sum_{k=0}^{n/2}\sum_{b=0}^{n-2k} f(\beta,\kappa)^n &\sim \left(\frac{3}{2}\right)^n \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-9x^2n-2(x+y)^2n}\,dx\,dy\\ &= \frac{\pi n}{3\sqrt{2}}\left(\frac{3}{2}\right)^n.\end{aligned}$$

By the concentration of the summation, the terms $k+b$ and $n-k-b$ in (7) can be taken as $\frac{1}{2}n$ each, and so (7) is asymptotic to

$$\frac{\sqrt{3}\pi^2}{2n^2}\left(\frac{n}{e}\right)^{2n} 36^n.$$

On dividing by the number of pairings given by (3), we obtain

$$\mathbf{E}(H_2^2) \sim \frac{\sqrt{3}\pi^2}{\sqrt{8}n^2}\left(\frac{9}{4}\right)^n. \tag{8}$$

(ii) $\mathbf{E}(H_2C_k)/\mathbf{E}H_2$

To estimate this, we first count the H-decompositions of pairings G which have a given k-cycle, summed over all k-cycles, and divide by $|\mathcal{P}_{n,4}|$. As before, the counting can be done first for the possibilities for the corresponding multigraph, and then for the assignment of points to the pairs by multiplying by $(4!)^n$.

Only fixed $k \geq 2$ needs to consider because cycles of length 1 cannot occur in any multigraphs with H-decompositions. To make calculations apply uniformly to all $k \geq 2$ including $k = 2$, consider a labelled directed k-cycle on n vertices, which can be selected in

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

ways. The edges of the k-cycle (actually, pairs in the cycle) must be split between the two H-cycles of the H-decomposition. Thus for some $i \geq 1$, the edges of the first H-cycle induce i paths $P_1, \dots, P_i$ in the k-cycle, and those of the second H-cycle induce i paths $Q_1, \dots, Q_i$. As we shall see, the value of i will affect the number of ways the H-cycle can be chosen, but first we count k-cycles with a given value of i. This can be done by starting at an arbitrary vertex v at the start of a path induced by the first H-cycle and proceeding around the cycle in the chosen direction. The number of ways to determine the lengths of paths $P_1, Q_1, P_2, Q_2, \dots, P_i, Q_i$ is

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

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

$$\frac{k}{2i}$$

since there are k 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 the k-cycle of the multigraph and its partition into paths is decided, if the number of edges of the first H-cycle which are in the k-cycle is j, the rest of the edges in this H-cycle can be chosen in $\frac{1}{2}(n-j-1)!2^i$ ways. This comes from a cyclic ordering of the $(n-j)$-set whose elements are those i paths and the vertices not in the paths. Similarly, the rest of the edges in the second H-cycle can be chosen in $\frac{1}{2}(n-k+j-1)!2^i$ ways. The product of these, regardless of j, is asymptotic to

$$\frac{h(n)^2 2^{2i}}{n^k}.$$

The product of the last four displayed expressions must be multiplied by $4!^n$ to account for the assignment of points to the pairs, summed over $i \geq 1$. If we then divide by (4) the result is

$$\begin{aligned}
\frac{\mathbf{E}(H_2C_k)}{\mathbf{E}H_2} &\sim \frac{1}{2}[x^k]\sum_{i\geq 1}\frac{1}{i}\left(\frac{4x^2}{(1-x)^2}\right)^i \\
&= -\frac{1}{2}[x^k]\ln\left(1-\left(\frac{4x^2}{(1-x)^2}\right)\right) \\
&= \frac{-2+(-1)^k+3^k}{2k}.
\end{aligned}$$

(iii) $\mathbf{E}(H_2[C_1]_{i_1}[C_2]_{i_2}\cdots[C_j]_{i_j})/\mathbf{E}H_2$

A direct generalisation of the argument above, applied to an ordered set of i_1 cycles of length 1, i_2 2-cycles, and so on, gives

$$\frac{\mathbf{E}(H_2[C_1]_{i_1}\cdots[C_k]_{i_k})}{\mathbf{E}H_2} \sim \prod_{j=1}^{k} \rho_j^{i_j} \tag{9}$$

where

$$\rho_j = \frac{-2+(-1)^j+3^j}{2j}.$$

Note that in the case $j = 1$ the expected number will be exactly 0, and so r_1 is correctly set to 0.

(iv) Synthesis

One more crucial piece of information is needed: as is well known for this pairing model of random 4-regular multigraphs (see [1] for example), the variables C_k, $k \geq 1$, are asymptotically independent Poisson random variables with expectations

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

We write

$$\rho_j = \lambda_j(1+\delta_j) \quad \text{where} \quad \delta_j = \frac{-2+(-1)^j}{3^j}$$

and note that from (5) and (8)

$$\frac{\mathbf{E}H_2^2}{(\mathbf{E}H_2)^2} \sim \sqrt{24} = \exp\sum_{k=1}^{\infty}\lambda_k\delta_k^2. \tag{10}$$

Hence the hypotheses of [11, Theorem 4.1] are satisfied with H_2 standing for Y in that theorem, and C_k standing for X_k. To state the conclusion of that theorem in this case, we define $\mathcal{F}^{(Y)}$ to denote the probability space obtained from a space $\mathcal{F}$ by making the probability of each $G \in \mathcal{F}$ proportional to $Y(G)$. Noting that $\delta_k = -1$ only for $k = 1$, we obtain

$$\overline{\mathcal{P}}_{n,4} \approx \overline{\mathcal{P}}_{n,4}^{(H_2)} \tag{11}$$

where $\overline{\mathcal{P}}_{n,4}$ is the space $\mathcal{P}_{n,4}$ restricted to $C_1 = 0$, i.e. the uniform probability space of loopless pairings. Since $\mathcal{G}'_{n,4}$ is the random multigraph space corresponding to the space $\overline{\mathcal{P}}_{n,4}$, and $\mathcal{H}_n + \mathcal{H}_n$ is similarly the random multigraph space corresponding to $\overline{\mathcal{P}}_{n,4}^{(H_2)}$, Theorem 2 follows from (11).

One can alternatively use [6, Corollary 1] or [4, Theorem 1] in place of [11, Theorem 4.1], but then extra argument is required to make the restriction to $C_1 = 0$.

Theorem 1 follows from Theorem 2 by restricting both spaces to simple graphs (which we write as the event $C_2 = 0$ in the graph spaces as well as the pairing spaces). This observation uses only the definition of contiguity and the fact that $\mathbf{P}(C_2 = 0)$ tends to a non-zero constant in $\overline{\mathcal{P}}_{n,4}$ as well as in $\overline{\mathcal{P}}_{n,4}^{(H_2)}$ (the latter of which comes from noting that (9) implies that the variables C_k are asymptotically independent Poisson in $\overline{\mathcal{P}}_{n,4}^{(H_2)}$ with expectations ρ_k). For if A_n is a.a.s. true in $\mathcal{G}_{n,4}$ then $A_n \wedge (C_2 = 0)$ is a.a.s. true in $\mathcal{G}'_{n,4}$, so by Theorem 2 it is a.a.s. true in $\mathcal{H}_n + \mathcal{H}_n$, whence A_n is a.a.s. true in $\mathcal{H}_n \oplus \mathcal{H}_n$. The reverse is also true, and hence Theorem 1.

∎

3 Proof of Theorem 1.3

Assuming a set S of subgraphs of G is nonempty, a *near uniform generator* for S is a randomised algorithm which on input $\epsilon > 0$ outputs a member s of S such that for any fixed $s_1 \in S$

$$\left| \Pr(s = s_1) - \frac{1}{|S|} \right| \leq \frac{\epsilon}{|S|}. \tag{12}$$

The probabilities here are with respect to the algorithm's random choices. The algorithm is called polynomial time if it runs in time polynomial in n and $1/\epsilon$. From [2] we have the following result, where HAM(G) denote the set of Hamilton cycles of G.

Theorem 5 *Let $d \geq 3$ be fixed. There is a polynomial time procedure which a.a.s. for $G \in \mathcal{G}_{n,d}$ is a near uniform generator for* HAM(G).

As with the proof of Corollary 2, we focus on d even, since for odd d the result follows by the same argument using results which were already known. So begin an induction on d with $d = 4$ and let $G \in \mathcal{G}_{n,4}$. From the theorem featuring in [9], G is a.a.s. hamiltonian. So, letting A denote event that $G \in \mathcal{G}_{n,4}$ has a Hamilton cycle and for which the procedure referred to in Theorem 5 is a near uniform generator, we have $\mathbf{P}(A) \to 1$.

Take a random $G \in \mathcal{G}_{n,4}$ and then use the procedure in Theorem 5 to get a random $H_1 \in \text{HAM}(G)$, with near uniform distribution. Let $G_1 = G \setminus E(H_1)$. Then G_1 is a 2-factor of G. Consider the resulting distribution of the pair (G_1, H_1). For $G \in A$, the probability that (G_1, H_1) arises can be computed *a posteriori* as

$$\frac{1 \pm \epsilon}{|\text{HAM}(H_1 \cup G_1)||\mathcal{G}_{n,4}|}.$$

Hence, the probability that any given 2-factor G_1 arises satisfies

$$\mathbf{P}(G_1) \geq \sum_{H: H \cup G_1 \in A} \frac{1 - \epsilon - o(1)}{|\text{HAM}(H \cup G_1)||\mathcal{G}_{n,4}|}.$$

Now take $G_0 \in \text{HAM}$. We know that the number of Hamilton cycles in $G \in \mathcal{G}_{n,4}$ is a.a.s. close to its expected value f_n, to be precise, at most $f_n \log n$; this comes from the results of [2] together with [4] (or see [11, Theorem 4.3], whose conditions are verified in [2]). From Theorem 1, for H randomly chosen, the random graph $H \cup G_0$ restricted to 4-regular graphs is contiguous to $\mathcal{G}_{n,4}$, and so it also has a.a.s. at most $f_n \log n$ Hamilton cycles. Thus, summing over H, we have

$$\mathbf{P}(G_1 = G_0) \geq |\{H : H \cup G_0 \in A\}| \cdot \frac{1 - \epsilon - o(1)}{f_n |\mathcal{G}_{n,4}| \log n}. \tag{13}$$

Let $\mathrm{HAM}(n)$ denote the set of Hamilton cycles on n vertices. It is easy to find $c > 0$ such that if we take a random $H \in \mathrm{HAM}(n)$ then with probability at least c, $E(H) \cap E(G_0) = \emptyset$. (See [11, Proof of Lemma 4.14] for example.) Since $\mathbf{P}(A) \to 1$, it follows that the first factor on the right in (13) is at least $(c - o(1))|\mathrm{HAM}(n)| \sim \frac{c_1}{\sqrt{n}}(\frac{n}{e})^n$ (where we use c_i for positive constants). Moreover, f_n is given in [8] and proved in [2] to be asymptotic to $\frac{c_2}{\sqrt{n}}(\frac{3}{2})^n$, and $|\mathcal{G}_{n,4}| \sim c_3(\frac{2n^2}{3e^2})^n$ is well known. Thus from (13),

$$\mathbf{P}(G_1 = G_0) \geq \frac{c_4 e^n}{n^n \log n}$$

and so

$$\mathbf{P}(G_1 \in \mathrm{HAM}) \geq |\mathrm{HAM}(n)| \cdot \frac{c_4 e^n}{n^n \log n} \geq \frac{c_5}{\sqrt{n} \log n}. \tag{14}$$

So with this probability, the pair (G_1, H_1) chosen above gives the desired complete hamiltonian decomposition of G. By repeating the generation of this pair n times, we find such a decomposition a.a.s. and in polynomial time.

Now consider an even $d \geq 4$. Choose a random $G \in \mathcal{G}_{n,d}$ and then using Theorem 5 to find a Hamilton cycle H_1 in G a.a.s. with almost uniform distribution. Again set $G_1 = G \setminus E(H_1)$ and repeat the arguement above. This time, we use the contiguity $H_n \oplus \mathcal{G}_{n,d-2} \approx \mathcal{G}_{n,d}$, and obtain for any $d-2$-regular graph G_0 that

$$\mathbf{P}(G_1 = G_0) \geq \frac{c_6}{|\mathcal{G}_{n,d-2}| \log n}.$$

So by induction, we can find a complete hamiltonian decomposition of G_1 in polynomial time with probability at least $(c_6 - o(1))/\log n$. Repeating $\log^2 n$ times gives success a.a.s. Including H_1 makes the required complete hamiltonian decomposition of G a.a.s. ∎

Finally, it is interesting to consider one potential near-uniform generator for the set of complete hamiltonian decompositions of a regular graph G: take a random Hamilton cycle in G and then by induction use a near-uniform generator to find a decomposition of $G - E(H)$. This is clearly not near-uniform, but one could try to equalise probabilities by counting the decompositions of $G - E(H)$. Even if these could be counted accurately, if the near-uniform generator is to work for G, it must generate *all* decompositions of G with nearly equal probabilities. The main difficulty is that the number of decompositions of $G - E(H)$ is not bounded above by a polynomial times the expected number, so a significant number of decompositions may not be found (often enough) by an algorithm running in polynomial time. Constructing an FPRAS runs into a similar difficulty.

4 Random matchings

In this section, we prove Theorem 4 and Corollary 3. Consider two independent random perfect matchings B (for blue) and R (for red) on $[2n]$. Then the graph BR formed from all edges in B and R is a disjoint union of even length cycles (with edges in $B \cap R$ regarded as cycles of length 2). Though it is well-known, for the sake of completeness and as a gentle introduction to our main arguments, we first prove a bound on the probability that the number $\kappa(BR)$ of cycles in BR is large.

It is clear that we can generate a matching uniformly at random by selecting the edges one at a time, in each case choosing an unmatched vertex u (by any rule whatsoever) and selecting its partner v uniformly at random from the remaining unmatched vertices. We call this *exposing* the edge uv.

We can clearly take the matching B as fixed, and therefore work with $\mathbf{P}_R$. We then expose red edges one by one. Take any vertex x and expose the red edge containing x (by choosing its neighbor uniformly at random among all $2n-1$ other vertices). If the red edge is identical to a blue edge, set $X_1 = 1$ and delete the two vertices. Otherwise, set $X_1 = 0$ and contract the blue-red-blue path into one blue edge consisting of the path's end-vertices. Each case produces a blue matching, but on $2n-2$ vertices. We call this *contracting* the blue matching with respect to the red edge. Now repeat the operation, exposing one more red edge for this slightly smaller blue matching, contracting to define X_2, and so on. Continue this procedure until no vertex remains. (The last red edge is necessarily identical to the last blue, and so $X_n = 1$.) It is clear that this generates a random red matching and $\kappa(BR) = X_1 + X_2 + \cdots + X_n$. Moreover, the indicator variables $\{X_i\}$ are independent with $p_i := \mathbf{P}(X_i = 1) = 1/(2n-2i+1)$.

Lemma 1 *For a positive integer $s \geq 5$,*

$$\mathbf{P}_R(\kappa(BR) \geq s \log n) = O(1/n^s).$$

Proof. Using

$$\mathbf{E}\Big(\exp(\alpha \sum_{i=1}^{n} X_i)\Big) = \prod_{i=1}^{n} \Big(1 + p_i(e^{\alpha} - 1)\Big) \leq \exp\Big((e^{\alpha} - 1) \sum_{i=1}^{n} p_i\Big),$$

and $\displaystyle\sum_{i=1}^{n} p_i \leq 1 + \frac{\log(2n)}{2}$, the Chernoff bound with $\alpha = \log 10$ gives

$$\mathbf{P}(k(BR) \geq s \log n) \leq \exp\Big(-\alpha s \log n + (e^{\alpha} - 1)\Big(1 + \frac{\log(2n)}{2}\Big)\Big) = O(1/n^s)\ . \quad \blacksquare$$

To prove Theorem 4, we consider a procedure which is similar to, but more complicated than, that above. Here there are fixed blue and red perfect matchings B and R, and a random perfect matching S of, say, silver edges. Take a vertex $x(1)$ in a shortest cycle of $BR(0) := BR$ and expose the random silver edge incident with $x(1)$. To avoid ambiguity, we assume that a fixed ordering of vertices is given and $x(1)$ is the smallest vertex in a shortest cycle. Let $y(1)$ be the other end of the silver edge. By contracting each of B and R independently with respect to the silver edge, we obtain blue and red matchings on $2n-2$ vertices which form a graph $BR(1)$. Repeating this, we obtain $x(2), y(2), \ldots, x(n-1), y(n-1)$ and graphs $BR(2), \ldots, BR(n-1)$. Clearly, the final graphs BS and RS are both Hamilton cycles if and only if this procedure creates no silver edge identical to an edge in a $BR(t)$.

The key observation is that the process $\{BR(t)\}$ tends to reduce the number of cycles unless there is only one cycle, and hence $BR(t)$ is frequently a Hamilton cycle (on $2n-2t$ vertices): if $BR(t-1)$ has more than one cycle, then the silver edge added to it is more likely to connect two (different) cycles, in which case the two cycles become one new cycle after contraction. If the other end $y(t)$ is in the cycle containing $x(t)$ and the cycle has length greater than 2, then there are two cases, depending on the distance $d(x(t),y(t))$, i.e. the length of a shortest path connecting $x(1)$ and $y(1)$. If the distance is odd but not 1, then the number of cycles increases by 1. It remains the same otherwise. (In the case that $x(t)$ and $y(t)$ are adjacent in the cycle, say by a red edge, it may appear at first sight that the number of cycles increases. However, the contraction operation throws away those two vertices and contracts the blue matching, so only one cycle is retained.) To be more precise, let $Y(t) = \kappa(BR(t-1)) - \kappa(BR(t))$. We call both $y(t)$ and the silver edge *good*, *neutral* or *bad* (with respect to $BR(t-1)$ and $x(t)$) if $Y(t) = 1, 0$ or -1, respectively. The probabilities of $y(t)$ being good, neutral and bad depend upon the length $2\sigma(t-1)$ of a shortest cycle in $BR(t-1)$ (recalling that the lengths of all cycles are even). Namely, if $\sigma(t-1) \geq 2$, then since all vertices not in the cycle containing $x(t)$ are good,

$$\mathbf{P}(Y(t)=1) = \frac{2n-2t+2-2\sigma(t-1)}{2n-2t+1}.$$

Similarly,

$$\mathbf{P}(Y(t)=0) = \frac{\sigma(t-1)+1}{2n-2t+1}$$

and

$$\mathbf{P}(Y(t)=-1) = \frac{\sigma(t-1)-2}{2n-2t+1}.$$

If the shortest cycle length is 2, i.e. a blue and a red edge coincide, then $\sigma(t-1) = 1$ and all vertices are good, or equivalently $Y(t) = 1$. Thus whenever $\kappa(BR(t-1)) > 1$, $Y(t)$ is stochastically greater than $Z(t)$, where $\mathbf{P}(Z(t) = 1) = \frac{1}{2}$ and $\mathbf{P}(Z(t) = 0) = \mathbf{P}(Z(t) = -1) = \frac{1}{4}$. Therefore, there are i.i.d. random variables $Z(1), \ldots, Z(n-1)$ such that $Y(t) \geq Z(t)$ for all t with $\kappa(BR(t-1)) > 1$. Though we do not directly use this fact, the idea will be used in the proof of Lemma 3.

For the next lemma, we denote by $U(t)$ the event that the silver edge created at step t is not identical to any edge in $BR(t-1)$. Then

$$\mathbf{P}(BS, RS \in \text{HAM}) = \mathbf{P}\Big(\bigcap_{t=1}^{n-1} U(t)\Big).$$

Lemma 2 *Let l be the number of 2-cycles in BR. Then for $l \leq k \leq n-1$,*

$$\mathbf{P}\Big(\bigcap_{t=1}^{k} U(t)\Big) = O\left(\frac{n-k+1}{\sqrt{n(n-l+1)}}\right).$$

In particular, if $l \leq \frac{1}{2}n$ then

$$\mathbf{P}(BS, RS \in \text{HAM}) = O\left(\frac{1}{n}\right).$$

Proof. For any event W, define

$$\mathbf{P}_t(W) := \mathbf{P}(W|U(1), \ldots, U(t-1)),$$

and notice that

$$\mathbf{P}\Big(\bigcap_{t=1}^{k} U(t)\Big) = \prod_{t=1}^{k} \mathbf{P}_t(U(t)). \tag{15}$$

For $t \leq l$, $x(t)$ is taken from a 2-cycle and the event $U(t)$ occurs unless $y(t)$ is the other vertex of the 2-cycle. Thus, for such t,

$$\mathbf{P}_t(U(t)) = 1 - \frac{1}{2n-2t+1}.$$

Furthermore, since all vertices are good and since the number of 2-cycles decreases by 1 for each step as long as $t \leq l$, $BR(l)$ contains no 2-cycles. After this, at most two 2-cycles are created at a time and they, if any, immediately disappear within the next step or two. So, in particular, the number of 2-cycles in $BR(t)$, $t > l$, is always at most 2. Thus, conditioned on the event $W(t)$ that no 2-cycle is created in steps $t-2$ or $t-1$, which implies that $x(t)$ belongs to a cycle of length greater than 2,

$$\mathbf{P}_t(U(t)|W(t)) = 1 - \frac{2}{2n-2t+1}.$$

For $U(t)$ occurs if and only if $d(x(t), y(t)) > 1$. On the other hand,

$$\mathbf{P}_t(U(t)|\overline{W(t)}) \leq 1 - \frac{1}{2n-2t+1}.$$

Therefore, for $t > l$

$$\begin{aligned} \mathbf{P}_t(U(t)) &\leq \mathbf{P}_t(W(t))\Big(1 - \frac{2}{2n-2t+1}\Big) + \mathbf{P}_t(\overline{W(t)})\Big(1 - \frac{1}{2n-2t+1}\Big) \\ &= \Big(1 - \frac{2}{2n-2t+1}\Big) + \frac{\mathbf{P}_t(\overline{W(t)})}{2n-2t+1}. \end{aligned}$$

Clearly,

$$\mathbf{P}_t(\overline{W(t)}) \leq \frac{2}{2n-2(t-1)-1} + \frac{2}{2n-2(t-2)-1} \leq \frac{4}{2n-2t+1} \tag{16}$$

since $d(x(t), y(t))$ must be 3 to create a 2-cycle, and the case $d(x(t), y(t)) = 1$ was already excluded. Thus, using (15),

$$\begin{aligned} \mathbf{P}\Big(\bigcap_{t=1}^{k} U(t)\Big) &\leq \prod_{t=1}^{l}\Big(1 - \frac{1}{2n-2t+1}\Big) \cdot \prod_{t=l+1}^{k}\Big(1 - \frac{2}{2n-2t+1} + \frac{4}{(2n-2t+1)^2}\Big) \\ &= O\Bigg(\frac{n-k+1}{\sqrt{n(n-l+1)}}\Bigg). \quad ∎ \end{aligned}$$

We are now ready to prove the lemma which is the heart of our arguments in this section.

Lemma 3 *Let R_1 and R_2 be two perfect matchings on $[2n]$ with $\kappa(BR_1), \kappa(BR_2) \leq n/40$. Put $k = \max\{\kappa(BR_1), \kappa(BR_2), 10\log n\}$. Then for a random (silver) matching S,*

$$\mathbf{P}_S(BS, SR_1 \in \mathrm{HAM}) = \mathbf{P}_S(BS, SR_2 \in \mathrm{HAM}) + O\Big(\frac{|l_1 - l_2|}{n^2}\Big) + O\Big(\frac{k}{n^3}\Big),$$

where l_i denotes the number of 2-cycles of BR_i ($i = 1$ and 2).

Proof. Later in this proof, we will define a distribution on the set of ordered pairs (S_1, S_2) of silver matchings with the marginal distributions of S_1 and S_2 exactly equal that of S. It follows that for the corresponding probability $\mathbf{P}_{(S_1,S_2)}$ we have

$$\mathbf{P}_S(BS, SR_1 \in \mathrm{HAM}) = \mathbf{P}_{(S_1,S_2)}(BS_1, S_1R_1 \in \mathrm{HAM}),$$

$$\mathbf{P}_S(BS, SR_2 \in \mathrm{HAM}) = \mathbf{P}_{(S_1,S_2)}(BS_2, S_2R_2 \in \mathrm{HAM}).$$

Let $\mathcal{A}_1, \mathcal{A}_2$ be the events $\{BS_1, S_1R_1 \in \mathrm{HAM}\}$ and $\{BS_2, S_2R_2 \in \mathrm{HAM}\}$ respectively. Then

$$\mathbf{P}_S(BS, SR_1 \in \mathrm{HAM}) - \mathbf{P}_S(BS, SR_2 \in \mathrm{HAM}) = \mathbf{P}_{(S_1,S_2)}\Big(\mathcal{A}_1 \backslash \mathcal{A}_2\Big) - \mathbf{P}_{(S_1,S_2)}\Big(\mathcal{A}_2 \backslash \mathcal{A}_1\Big).$$

In the rest of the proof, $\mathbf{P}$ denotes $\mathbf{P}_{(S_1,S_2)}$. Without loss of generality we may assume $l_2 \leq l_1$. From above,s it is enough to show that

$$\mathbf{P}\Big(\mathcal{A}_1 \setminus \mathcal{A}_2\Big) = O\Big(\frac{l_1 - l_2}{n^2}\Big) + O\Big(\frac{k}{n^3}\Big) \quad \text{and} \quad \mathbf{P}\Big(\mathcal{A}_2 \setminus \mathcal{A}_1\Big) = O\Big(\frac{k}{n^3}\Big). \qquad (17)$$

For $i = 1$ and 2, let $x_i(1)$ be the smallest vertex in a shortest cycle of BR_i, and denote that cycle by $C_i(0)$. Call a vertex of BR_i good, neutral or bad if it is such with respect to BR_i and $x_i(1)$. Also define $BR_i(0) = BR_i$.

If the length $|C_1(0)|$ of $C_1(0)$ is at least $|C_2(0)|$, then take a bijection on $[2n]$ with the following properties:

(i) the vertices (or vertex) adjacent to $x_2(1)$ in $BR_2(0)$ are (or is) mapped to vertices adjacent to $x_1(1)$ in $BR_1(0)$;

(ii) all bad vertices of $BR_2(0)$ are mapped to bad vertices of $BR_1(0)$;

(iii) and all neutral vertices which are not distance 1 from $x_2(1)$ are mapped, if possible, to bad vertices of $BR_1(0)$ and otherwise to neutral vertices of $BR_1(0)$;

(iv) good vertices of $BR_2(0)$ are mapped to the remaining vertices of $BR_1(0)$ (whether they are good or not);

(v) If $BR_1(0)$ and $BR_2(0)$ are isomorphic graphs then the mapping is an isomorphism.

This choice of bijection is possible since $|C_1(0)| \geq |C_2(0)|$. Choose $y_1(1)$ uniformly at random from $[2n] \setminus \{x_1(1)\}$. We couple $y_1(1)$ and $y_2(1)$ by selecting $y_2(1)$ so that it corresponds to $y_1(1)$ under the bijection. If $|C_1(0)| < |C_2(0)|$, reverse the roles of BR_1 and BR_2 in the above discussion.

The first (silver) edge in S_i is $x_i(1)y_i(1)$, thus determining $BR_i(1)$ for $i = 1$ and 2. Repeating this procedure, we obtain a random process $\{x_i(t), y_i(t)\}_{t=1,\ldots,n-1}$ for

$i = 1$ and 2. Define S_i to contain the set of edges $\{x_i(t), y_i(t)\}$ together with the remaining pair of vertices (for $i = 1$ and 2). This induces the distribution on ordered pairs (S_1, S_2) mentioned at the start of this proof.

For any $t \geq 1$ and $i = 1$ and 2, from the above process to time t define the events $U_i(t)$ and $W_i(t)$ determined by $BR_i(t-1)$ and $x_i(t)y_i(t)$ as $U(t)$ and $W(t)$ were determined by $BR(t-1)$ and $x(t)y(t)$. Also let

$$\begin{aligned}\kappa(t) &= \max\{\kappa(BR_1(t)), \kappa(BR_2(t))\} \quad (t \geq 0),\\ Y(t) &= \kappa(t-1) - \kappa(t) \quad (t \geq 1).\end{aligned}$$

It is clear that the symmetric difference of $U_1(t)$ and $U_2(t)$ can only contain those processes where $|C_1(t-1)| = 2$ or $|C_2(t-1)| = 2$ (but not both). If a process is in one but not the other and say $|C_1(t-1)| \geq |C_2(t-1)|$, then it must be in $U_2(t) \setminus U_1(t)$, and we have $|C_2(t-1)| = 2$ and $|C_1(t-1)| > 2$. Hence

$$U_2(t) \setminus U_1(t) \subseteq \{|C_2(t-1)| = 2\} \cap \{|C_1(t-1)| > 2\}, \tag{18}$$

and similarly

$$U_1(t) \setminus U_2(t) \subseteq \{|C_1(t-1)| = 2\} \cap \{|C_2(t-1)| > 2\}. \tag{19}$$

For any $t \geq 1$, condition on some trajectory of the process up to time $t-1$, i.e. on the sequence $\{x_i(j), y_i(j)\}_{j=0,\dots,t-1}$, for which $|C_1(t-1)| \geq |C_2(t-1)|$. Suppose on the one hand that $\kappa(BR_1(t-1)) > 1$. If $y_1(t)$ is good, which occurs with probability $p \geq \frac{1}{2}$, then so is $y_2(t)$, and hence $\kappa(t) = \kappa(t-1) - 1$ and $Y(t) = 1$. Similarly, if $y_1(t)$ is neutral, which occurs with probability at least $(1-p)/2$, then $y_2(t)$ is neutral or good and $Y(t) \geq 0$. As before, $Y(t)$ is stochastically greater than $Z(t)$ where $\mathbf{P}(Z(t) = 1) = \frac{1}{2}, \mathbf{P}(Z(t) = 0) = \mathbf{P}(Z(t) = -1) = \frac{1}{4}$.

On the other hand, suppose that $\kappa(BR_1(t-1)) = 1$, and $\kappa(BR_2(t-1)) \geq 2$. Then all bad and neutral vertices of $BR_2(t-1)$, except those adjacent to $x_2(t-1)$, are mapped to bad vertices of $BR_1(t-1)$. Hence if $y_1(t-1)$ is neutral and not adjacent to $x_1(t-1)$, which occurs with probability $p = \frac{n-t}{2n-2t+1}$, then $y_2(t-1)$ is good and $Y(t) = 1$. Alternatively, conditioning on $y_1(t-1)$ being bad, $y_2(t-1)$ is bad with probability at most $\frac{n-t-3}{2n-2t-2}$, so that $Y(t) = -1$ with probability at most $(1-p)\frac{n-t-3}{2n-2t-2} < \frac{1}{4}$. Thus we have a conclusion similar to that in the above case: $Y(t)$ is stochastically greater than $Z_\epsilon(t)$ where $\mathbf{P}(Z_\epsilon(t) = 1) = \frac{1}{2} - \epsilon$, $\mathbf{P}(Z_\epsilon(t) = 0) = \frac{1}{4} + \epsilon$, and $\mathbf{P}(Z_\epsilon(t) = -1) = \frac{1}{4}$, provided $\epsilon > \frac{1}{4n-4t+2}$.

The same is also true by symmetry in the case $|C_1(t-1)| < |C_2(t-1)|$. Approximating $Z(t)$ from below by a coupled copy of $Z_\epsilon(t)$, we can obtain i.i.d. random variables $\{Z_\epsilon(t)\}_{t=1,\dots,n-1}$ with $Y(t) - Z_\epsilon(t) \geq 0$ whenever $\kappa(t-1) \geq 2$.

Define the stopping time T to be the smallest t such that $\kappa(t) = 1$. Since $\kappa(BR_1(n-1)) = \kappa(BR_2(n-1)) = 1$, T is well-defined. The event $\mathcal{A}_1 \setminus \mathcal{A}_2$ implies that $U_1(t)$ occurs for all t and $U_2(t)$ does not occur at some t, say T_2. If $t > T$, then the events $U_1(t)$ and $U_2(t)$ occur at the same time. Thus provided $\mathcal{A}_1 \setminus \mathcal{A}_2$ occurs and $T \leq 20k$, we have $T_2 \leq 20k \leq \frac{1}{2}n$ and the two events $\bigcap_{t=1}^{20k} U_1(t) \setminus \bigcap_{t=1}^{20k} U_2(t)$, and $\bigcap_{t=20k+1}^{n-1} U_1(t)$, must occur. It follows that

$$\begin{aligned}\mathbf{P}(\mathcal{A}_1 \setminus \mathcal{A}_2) &\leq \mathbf{P}(T \geq 20k) + \mathbf{P}(\{T \leq 20k\} \cap (\mathcal{A}_1 \setminus \mathcal{A}_2))\\ &\leq \mathbf{P}(T \geq 20k) + \mathbf{P}\Big(\mathcal{B} \cap \Big(\bigcap_{t=1}^{20k} U_1(t) \setminus \bigcap_{t=1}^{20k} U_2(t)\Big)\Big) \qquad (20)\end{aligned}$$

where $\mathcal{B} = \bigcap_{t=20k+1}^{n-1} U_1(t)$. If $T \geq 20k$, then $Y(t) \geq Z_\epsilon(t)$ for all $t \leq 20k-1$, and

$$2 \leq \kappa(20k-1) = \kappa(0) - (Y(1) + \cdots + Y(20k-1)) \leq k - (Z_\epsilon(1) + \cdots + Z_\epsilon(20k-1)).$$

Therefore,

$$\mathbf{P}(T \geq 20k) \leq \mathbf{P}(Z_\epsilon(1) + \cdots + Z_\epsilon(20k-1) \leq k) = O(1/n^4), \tag{21}$$

where in the equality we use the Chernoff bound for i.i.d. random variables $Z_\epsilon(t)$, choosing $\epsilon > 0$ sufficiently small but the same for each t.

For the second term in (20), since

$$\bigcap_{t=1}^{20k} U_1(t) \setminus \bigcap_{t=1}^{20k} U_2(t) \subseteq \bigcup_{t=1}^{20k} (U_1(t) \setminus U_2(t)),$$

we have

$$\begin{aligned} \mathbf{P}\Big(\mathcal{B} \cap \Big(\bigcap_{t=1}^{20k} U_1(t) \setminus \bigcap_{t=1}^{20k} U_2(t)\Big)\Big) &\leq \mathbf{P}\Big(\mathcal{B} \cap \Big(\bigcup_{t=1}^{20k} U_1(t) \setminus U_2(t)\Big)\Big) \\ &\leq \sum_{t=1}^{20k} \mathbf{P}\Big((U_1(t) \setminus U_2(t)) \cap \mathcal{B}\Big). \end{aligned} \tag{22}$$

For $t' \leq 20k$, $U_1(t') \backslash U_2(t')$ is completely determined by $\{x_1(t), y_1(t), x_2(t), y_2(t)\}_{t=1,\ldots,20k}$, and so we have by [3, Section 3.7] for example

$$\begin{aligned} \mathbf{P}\Big((U_1(t') \setminus U_2(t')) \cap \mathcal{B}\Big) &= \mathbf{E}\Big(I(U_1(t') \setminus U_2(t'))I(\mathcal{B})\Big) \\ &= \mathbf{E}\Big(I\Big(U_1(t') \setminus U_2(t')\Big)\mathbf{E}\Big(I(\mathcal{B})\Big|\{x_1(t), y_1(t), x_2(t), y_2(t)\}_{t=1,\ldots,20k}\Big)\Big). \end{aligned} \tag{23}$$

As $l_1 \leq k$, it follows that $BR_1(20k)$ has at most two 2-cycles as argued in the proof of Lemma 2. That lemma applied to the process which begins from $t = 20k$ now yields

$$\mathbf{E}\Big(I(\mathcal{B})\Big|\{x_1(t), y_1(t), x_2(t), y_2(t)\}_{t=1,\ldots,20k}\Big) = O\Big(\frac{1}{n-20k}\Big) = O(1/n). \tag{24}$$

By (19), we know that $U_1(t) \setminus U_2(t) = \emptyset$ unless $t > l_2$. If $l_2 < t \leq l_1$, then

$$\mathbf{E}\Big(I(U_1(t) \setminus U_2(t))\Big) \leq \frac{1}{2n-2t+1} \leq \frac{1}{n}$$

since $l_1 \leq k \leq n/40$. If $l_1 < t \leq 20k$ then (19) implies that $U_1(t) \backslash U_2(t) \subseteq \overline{W_1(t)} \backslash U_2(t)$ and hence

$$\begin{aligned} \mathbf{E}\Big(I(U_1(t) \setminus U_2(t))\Big) &\leq \mathbf{E}\Big(I\Big(\overline{W_1(t)} \setminus U_2(t)\Big)\Big) = \mathbf{P}(\overline{W_1(t)})\mathbf{P}\Big(\overline{U_2(t)}|\overline{W_1(t)}\Big) \\ &\leq \frac{4}{2n-2t+1} \cdot \frac{2}{2n-2t+1} = O(1/n^2) \end{aligned}$$

using (16). Therefore by (21), (22), (23) and (24),

$$\mathbf{P}(\mathcal{A}_1 \setminus \mathcal{A}_2) = O\Big(\frac{l_1 - l_2}{n^2}\Big) + O\Big(\frac{k}{n^3}\Big).$$

The same argument yields

$$\mathbf{P}(\mathcal{A}_2 \setminus \mathcal{A}_1) = O\Big(\frac{k}{n^3}\Big),$$

as required for (17). ∎

We are now in a position to prove Theorem 4. For fixed matchings B, R_1 and R_2 and a random matching S,

$$\mathbf{P}_S(BS, SR_1 \in \mathrm{HAM}) = \mathbf{P}_S(BS, SR_2 \in \mathrm{HAM}) + O\Big(\frac{|l_1 - l_2|}{n^2}\Big) + O\Big(\frac{k}{n^3}\Big).$$

Now for a random matching R_2, we have $\kappa(BR_2) < 5 \log n$ with probability $1 - O(1/n^5)$ by Lemma 1. Also it is easily seen that $\mathbf{E}_{R_2} l_2$ is bounded. So by taking the expectation of both sides of the equation with respect to R_2, we obtain

$$\mathbf{P}_S(BS, SR_1 \in \mathrm{HAM}) = \mathbf{E}_{R_2}\mathbf{P}_S(BS, SR_2 \in \mathrm{HAM}) + O\Big(\frac{l_1 + 1}{n^2}\Big) + O\Big(\frac{\kappa(SR_1) + \log n}{n^3}\Big).$$

Theorem 4 follows since

$$\begin{aligned}\mathbf{E}_{R_2}\mathbf{P}_S(BS, SR_2 \in \mathrm{HAM}) &= \mathbf{E}_{R_2}\mathbf{E}_S(I(BS \in \mathrm{HAM})I(SR_2 \in \mathrm{HAM})) \\ &= \mathbf{E}_S\mathbf{E}_{R_2}(I(BS \in \mathrm{HAM})I(SR_2 \in \mathrm{HAM})) \\ &= (\mathbf{P}_S(BS \in \mathrm{HAM}))^2 \\ &= p_{\mathrm{H}}(2n)^2,\end{aligned}$$

which is of order $1/n$.

Corollary 3 also easily follows using Lemma 1. We can clearly assume $r \geq 3$. Take an ordering $v_1, \ldots, v_r$ of the vertices of G so that the number $d(j)$ of $i > j$ with $v_i v_j \in E(G)$ is at most 2. Let r^* denote the number of j with $d(j) = 2$. For each such j, if v_{i_1} and v_{i_2} are the two neighbours of v_j with $i_1, i_2 > j$ then by Lemma 1, with probability at least $1 - O(1/n^{r+2})$, we have $\kappa(M_{v_i} M_{v_j}) \leq (r + 2) \log n$.

We can now estimate the probability that $M_{v_i} M_{v_j} \in \mathrm{HAM}$ for all i and j such that $v_i v_j \in E(G)$. The strategy is to treat the matchings in the order $M_{v_r}, M_{v_{r-1}}, \ldots, M_{v_1}$ and check H-compatibility of each with those of the others which are required and have already appeared, conditioning progressively on each step being successful. For those r^* where two others are involved, we use Theorem 4. (Note that κ is an upper bound for the number l of 2-cycles.) Otherwise, there is only one other matching involved, and the probability that the new matching is H-compatible with it is just $p_{\mathrm{H}}(2n)$. So the required final probability is

$$\begin{aligned}&\prod_{j=1}^{r}(1 + O(rI(d(j) = 2)\log n/n))(p_{\mathrm{H}}(2n))^{d(j)} + O\left(\frac{r}{n^{r+2}}\right) \\ &= (1 + O(rr^* \log n/n))(p_{\mathrm{H}}(2n))^m.\end{aligned}$$

Trivially, $r^* \leq r$. For a cycle of length r, the natural ordering yields $r^* = 1$.

References

[1] B. Bollobás, *Random graphs*, Academic Press, London, 1985.

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

[3] G.R. Grimmett and D.R. Stirzaker, *Probability and Random Processes*, 2nd edn, Clarendon (1992).

[4] S. Janson, Random regular graphs: Asymptotic distributions and contiguity, *Combinatorics, Probability and Computing* **4** (1995), 1–37.

[5] C. McDiarmid, On the method of bounded differences, *Surveys in Combinatorics, 1989*, J. Siemons ed. (1989), 148-188.

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

[7] H. Robalewska, 2-factors in random regular graphs, *Journal of Graph Theory* **23** (1996), 215–224.

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

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

[10] A. Steger and N.C. Wormald, Generating random regular graphs quickly, *Combinatorics, Probability and Computing* **8** (1999), 377–396.

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