Permutation pseudographs and contiguity

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

Svante Janson†
Department of Mathematics,
Uppsala University,
S-751 06 Uppsala, Sweden
svante@math.uu.se

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

30 August 2001

Abstract

The space of permutation pseudographs is a probabilistic model of 2-regular pseudographs on n vertices, where a pseudograph is produced by choosing a permutation σ of $\{1, 2, \ldots, n\}$ uniformly at random and taking the n edges $\{i, \sigma(i)\}$. We prove several contiguity results involving permutation pseudographs (contiguity is a kind of asymptotic equivalence of sequences of probability spaces). Namely, we show that a random 4-regular pseudograph is contiguous with the sum of two permutation pseudographs, the sum of a permutation pseudograph and a random Hamilton cycle, and the sum of a permutation pseudograph and a random 2-regular pseudograph. (The sum of two random pseudograph spaces is defined by choosing a pseudograph from each space independently and taking the union of the edges of the two pseudographs.) All these results are proved simultaneously by working in a general setting, where each cycle of the permutation is given a nonnegative constant multiplicative weight. A further contiguity result is proved involving the union of a weighted permutation pseudograph and a random regular graph of arbitrary degree. All corresponding results for simple graphs are obtained as corollaries.

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

1 Introduction

Recently, Kim and Wormald [7] showed that a random d-regular graph, for d even, asymptotically almost surely (a.a.s.) has an edge-decomposition into Hamilton cycles. This was done by proving the asymptotic equivalence (or *contiguity*) of two different probability models of random 4-regular graphs. The purpose of this paper is to prove similar contiguity results for various models where Hamilton cycles are replaced by permutation pseudographs, thereby confirming several conjectures of Janson [6, 5]. First we introduce some necessary notation.

Denote by $\mathcal{G}_{n,d}$ the uniform probability space of d-regular graphs on n vertices, where dn is even. As is usual in this area, we approach random d-regular graphs via the standard pairing model (see Bollobás [1]). Consider dn labelled points, with d points in each of n buckets, and take a random perfect matching of the points. We call this uniform probability space $\mathcal{P}_{n,d}$. Letting the buckets be vertices and each pair represent an edge (joining the buckets containing the two endpoints of the pair), we obtain a random regular pseudograph (which may have loops or multiple edges). Denote this probability space by $\mathcal{G}^*_{n,d}$. Graphs with no loops or multiple edges occur with equal probabilities, so the restriction of $\mathcal{G}^*_{n,d}$ to (simple) graphs gives the uniform probability space $\mathcal{G}_{n,d}$. Also let $\mathcal{G}'_{n,d}$ be the space $\mathcal{G}^*_{n,d}$ restricted to pseudographs with no loops (but which may still have multiple edges).

We can also form a random 2-regular pseudograph as follows: choose a permutation $\sigma \in \mathrm{Sym}(n)$ uniformly at random, and take the n edges $\{i, \sigma(i)\}$. The probability space of 2-regular pseudographs which results is called the *permutation pseudograph* model. Conditioning on no loops or multiple edges, we obtain the *permutation graph* model.

Suppose that $(\mathcal{B}_n)_{n\geq 1}$ and $(\hat{\mathcal{B}}_n)_{n\geq 1}$ are two sequences of probability spaces such that $\mathcal{B}_n$ and $\hat{\mathcal{B}}_n$ 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

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

In other words, an event A_n is a.a.s. true in $\mathcal{B}_n$ if and only if it is a.a.s. true in $\hat{\mathcal{B}}_n$. (Here a.a.s. stands for "asymptotically almost surely", meaning "with probability which tends to 1 as $n \to \infty$".) In this case, write

$$\mathcal{B}_n \approx \hat{\mathcal{B}}_n.$$

For more information on contiguity see [6, Sections 9.5,9.6] or [14, Section 4].

We will also need the definition of the sum of two pseudograph models. If $\mathcal{G}$ and $\hat{\mathcal{G}}$ are two probability spaces of random graphs or pseudographs on the same vertex set, then their *sum* $\mathcal{G}+\hat{\mathcal{G}}$ is the space whose elements are defined by the random pseudograph $G \cup \hat{G}$, where $G \in \mathcal{G}$ and $\hat{G} \in \hat{\mathcal{G}}$ are generated independently. The *graph-restricted sum* of $\mathcal{G}$ and $\hat{\mathcal{G}}$, denoted $\mathcal{G} \oplus \hat{\mathcal{G}}$, is the space which is the restriction of $\mathcal{G}+\hat{\mathcal{G}}$ to simple graphs.

We will prove three new contiguity results involving permutation pseudographs. Informally, we will show that the sum of two permutation pseudographs, or of a permutation pseudograph and a random Hamilton cycle, or of a permutation pseudograph and a random 2-regular pseudograph, are all contiguous with a random 4-regular pseudograph. (In the above, the term "random regular pseudograph" refers to the model $\mathcal{G}^*_{n,d}$ which arises from the pairings model.) By conditioning on having no loops or multiple edges, we obtain the corresponding results for graphs. All the pseudograph results are proved simultaneously by placing the problems in a common general setting. In fact, we will also obtain some previously-known contiguity results for free. This general setting is now described.

Let $\theta \geq 0$ be some constant. Give each permutation $\sigma \in \mathrm{Sym}(n)$ the weight

$$\theta^{\kappa(\sigma)-1},$$

where $\kappa(\sigma)$ is the number of cycles in (the disjoint cycle decomposition of) σ, and choose $\sigma \in \mathrm{Sym}(n)$ with probability proportional to its weight. Then form a 2-regular pseudograph G on n vertices, by taking the n edges $\{i, \sigma(i)\}$, as above. Let $\kappa(G)$ be the number of cycles in G, and let $\nu(G)$ be the number of cycles in G of length 1 or 2, for $G \in \mathcal{G}^*_{n,2}$. Each $G \in \mathcal{G}^*_{n,2}$ arises from exactly $2^{\kappa(G)-\nu(G)}$ permutations. If σ is chosen with probability proportional to $\theta^{\kappa(\sigma)-1}$, then G occurs with probability proportional to

$$W_\theta(G) = \theta^{\kappa(G)-1}\, 2^{\kappa(G)-\nu(G)}.$$

Let $\mathcal{F}_n(\theta)$ be the model of random 2-regular pseudographs obtained in this way. We denote this as

$$\mathcal{F}_n(\theta) = \left(\mathcal{G}^*_{n,2}\right)^{(W_\theta)}.$$

So $\mathcal{F}_n(\theta)$ is the model of 2-regular pseudographs obtained by choosing G with weight proportional to $W_\theta(G)$. When $\theta = 1$ we obtain the permutation pseudograph model, as described above. When $\theta = 1/2$ we obtain $\mathcal{G}^*_{n,2}$. To see this, note that the weight of G in $\mathcal{G}^*_{n,2}$ is $2^{n-\nu(G)}$, which is proportional to $W_{1/2}(G)$. Finally, when $\theta = 0$ the only pseudographs with nonzero weight are those with $\kappa(G) = 1$; namely, Hamilton cycles. Each Hamilton cycle gets the same weight, so we obtain $\mathcal{H}_n$, the space of uniform Hamilton cycles on n vertices.

Before stating our main result, we make a few remarks. When $\theta > 0$ it is equivalent, and customary, to use the weight $\theta^{\kappa(\sigma)}$ instead; our choice covers the case $\theta = 0$ too. It is well-known (see for example [13, (3.5.2)]) that the sum of the weights $\theta^{\kappa(\sigma)-1}$ over the $n!$ permutations in $\mathrm{Sym}(n)$ equals $(\theta+1)(\theta+2)\cdots(\theta+n-1)$. Hence the probability of choosing a given permutation σ is $\theta^{\kappa(\sigma)-1}/(\theta+1)\cdots(\theta+n-1)$.

The distribution of random permutations defined by these weights appears in several contexts, for example in the so-called *Chinese restaurant process* [11, Section 6.3]. The joint distribution of the number of cycles of various lengths for these random permutations, or equivalently for $\mathcal{F}_n(\theta)$, appears in several further contexts, and is known as the *Ewens sampling formula.* See [11, 12] for a survey.

Our main result is the following. For the proof we use the small subgraph conditioning method introduced in [9, 10] by Robinson and Wormald.

Theorem 1.1. *Let $\theta_1, \theta_2 \geq 0$ be constants. Then*

$$\mathcal{F}_n(\theta_1) + \mathcal{F}_n(\theta_2) \approx \mathcal{G}^*_{n,4}$$

except in the case $\theta_1 = \theta_2 = 0$, *when*

$$\mathcal{F}_n(\theta_1) + \mathcal{F}_n(\theta_2) \approx \mathcal{G}'_{n,4}.$$

The exception for $\theta_1 = \theta_2 = 0$ is natural, since this is the only case where $\mathcal{F}_n(\theta_1) + \mathcal{F}_n(\theta_2)$ is loopless.

As corollaries we obtain the corresponding results for (simple) graphs, by considering the graph-restricted sum: i.e. by conditioning on $\mathcal{F}_n(\theta_1) + \mathcal{F}_n(\theta_2)$ being a graph (see Corollary 1.1).

As particular cases of Theorem 1.1, we obtain several new results involving permutation pseudographs. Let $\mathcal{F}_n = \mathcal{F}_n(1)$ denote the permutation pseudograph model. By taking (θ_1, θ_2) =(0,1), $(\frac{1}{2}, 1)$ and $(1, 1)$, we show that the three models

$$\mathcal{H}_n + \mathcal{F}_n, \qquad \mathcal{G}^*_{n,2} + \mathcal{F}_n, \qquad \mathcal{F}_n + \mathcal{F}_n$$

are all contiguous with $\mathcal{G}^*_{n,4}$. We also obtain the known results

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

(see [7]) by setting $(\theta_1, \theta_2) = (0, 0)$,

$$\mathcal{H}_n + \mathcal{G}^*_{n,2} \approx \mathcal{G}^*_{n,4}$$

(see [4, 10]) by setting $(\theta_1, \theta_2) = (0, \frac{1}{2})$, and

$$\mathcal{G}^*_{n,2} + \mathcal{G}^*_{n,2} \approx \mathcal{G}^*_{n,4}$$

(see [8]) by setting $(\theta_1, \theta_2) = (\frac{1}{2}, \frac{1}{2})$.

Note that one consequence of these results is

$$\mathcal{F}_n + \mathcal{F}_n \approx \mathcal{H}_n + \mathcal{G}^*_{n,2},$$

showing that the sum of two permutation pseudographs is a.a.s. Hamiltonian. This was proved directly by Frieze [3].

In the final section we prove the following result, using the same method.

Theorem 1.2. *Let* $\theta \geq 0$ *be constant, and* $d \geq 3$. *Then*

$$\mathcal{F}_n(\theta) + \mathcal{G}^*_{n,d-2} \approx \mathcal{G}^*_{n,d}$$

except when $\theta = 0$ *and* $d = 3$, *when*

$$\mathcal{F}_n(0) + \mathcal{G}^*_{n,1} \approx \mathcal{G}'_{n,3}.$$

In particular, by setting $\theta = 1$, we prove the following contiguity result for permutation pseudographs:

$$\mathcal{F}_n + \mathcal{G}^*_{n,d-2} \approx \mathcal{G}^*_{n,d}$$

for $d \geq 3$. (This result was claimed in [6, Section 9.5].) The theorem also captures several known results: for $\theta = 0$ we obtain

$$\mathcal{H}_n + \mathcal{G}^*_{n,1} \approx \mathcal{G}'_{n,3}$$

and

$$\mathcal{H}_n + \mathcal{G}^*_{n,d-2} \approx \mathcal{G}^*_{n,d}$$

for $d \geq 4$, see [4, 5]. For $\theta = \frac{1}{2}$ we obtain

$$\mathcal{G}^*_{n,2} + \mathcal{G}^*_{n,d-2} \approx \mathcal{G}^*_{n,d}$$

see [8].

Theorems 1.1 and 1.2 confirm all the conjectures of Janson relating to permutation pseudographs (see [5] and the end of [6, Section 9.5]), and imply the following extension of [6, Theorem 9.43] for sums of several random pseudographs.

Theorem 1.3. *Let $m \geq 1$ and let $d_1, \ldots, d_m \geq 1$. Let $G_1, \ldots, G_m$ be independent random pseudographs such that G_i is a copy of G^*_{n,d_i} when $d_i \neq 2$, and a copy of $\mathcal{F}_n(\theta_i)$, for arbitrary constant $\theta_i \geq 0$, when $d_i = 2$. If $d = d_1 + \cdots + d_m \geq 3$, then*

$$G_1 + \cdots + G_m \approx \mathcal{G}^*_{n,d}$$

except when the only summands are $\mathcal{G}^*_{n,1}$ and $\mathcal{F}_n(0) = \mathcal{H}_n$, in which case

$$G_1 + \cdots + G_m \approx \mathcal{G}'_{n,d}.$$

As an application, this implies that the a.a.s. result by Friedman [2] on the second eigenvalue of $\mathcal{F}_n + \cdots + \mathcal{F}_n$ applies also to $\mathcal{G}^*_{n,2d}$ (and to $\mathcal{G}_{n,2d}$).

Note however that no two of the models $\mathcal{F}_n(\theta)$ are contiguous. This follows from [6, Corollary 9.54] and the standard fact [11] that the number of cycles of length k in $\mathcal{F}_n(\theta)$ is asymptotically Poisson with mean θ/k, for $k \geq 1$, with the numbers for different lengths asymptotically independent.

By conditioning on no loops and multiple edges in Theorem 1.3, we obtain all corresponding results for simple graphs.

Corollary 1.1. *Let $m \geq 1$ and let $d_1, \ldots, d_m \geq 1$. Let $G_1, \ldots, G_m$ be independent random pseudographs such that G_i is a copy of G^*_{n,d_i} when $d_i \neq 2$, and a copy of $\mathcal{F}_n(\theta_i)$, for arbitrary constant $\theta_i \geq 0$, when $d_i = 2$. If $d = d_1 + \cdots + d_m \geq 3$, then*

$$G_1 \oplus \cdots \oplus G_m \approx \mathcal{G}_{n,d}.$$

Finally, note that Janson [5, Theorem 12] showed that $\mathcal{G}^*_{n,d}$ is contiguous with the *uniform* model of d-regular pseudographs on n vertices. Therefore all of our results hold using uniformly distributed regular pseudographs instead of the models $\mathcal{G}^*_{n,d}$ which arise from the pairing model.

1.1 Further notation and preliminary results

Our proofs use the small subgraph conditioning method, stated below. 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 [14]. A similar theorem is given in [6, Theorem 9.12].

Theorem 1.4 ([14], 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 with many contiguity results for graphs, it is most convenient to perform the calculations at the pairings level. In $\mathcal{F}_n(\theta)$, the pseudograph G is given weight $W_\theta(G)$. If we divide this weight evenly among the $2^{n-\nu(G)}$ pairings corresponding to G, then a pairing $P \in \mathcal{P}_{n,2}$ corresponding to G receives weight

$$\theta^{\kappa(G)-1}\, 2^{\kappa(G)-\nu(G)}\, 2^{\nu(G)-n} = 2^{1-n}(2\theta)^{\kappa(G)-1}.$$

That is, the pairing $P \in \mathcal{P}_{n,2}$ has weight which is proportional to

$$w_\theta(P) = (2\theta)^{\kappa(P)-1}.$$

(This demonstrates that $\mathcal{F}_n(\frac{1}{2}) = \mathcal{G}^*_{n,2}$, since then P is chosen uniformly.) Let $\hat{\mathcal{F}}_n(\theta) = \mathcal{P}_{n,2}^{(w_\theta)}$; that is, the probability space of pairings in $\mathcal{P}_{n,2}$ where pairing P receives weight $w_\theta(P) = (2\theta)^{\kappa(P)-1}$. For $\theta_1, \theta_2 \geq 0$, unless $(\theta_1, \theta_2) = (0,0)$ we show that

$$\hat{\mathcal{F}}_n(\theta_1) + \hat{\mathcal{F}}_n(\theta_2) \approx \mathcal{P}_{n,4}.$$

Here '+' denotes combining two pairings of disjoint sets of points by merging the corresponding buckets and randomly relabelling the points in each resulting bucket. In the exceptional case $\theta_1 = \theta_2 = 0$, we obtain

$$\hat{\mathcal{F}}_n(0) + \hat{\mathcal{F}}_n(0) \approx \mathcal{P}'_{n,4},$$

where $\mathcal{P}'_{n,4}$ is $\mathcal{P}_{n,4}$ conditioned on no loops; i.e., no pair has both ends in the same bucket.

We now define the random variable Y which will be analysed. Fix constants $\theta_1, \theta_2 \geq 0$. Let (F_1, F_2) be a partition of the pairs of a pairing $P \in \mathcal{P}_{n,4}$ such that exactly two points in each bucket belong to F_1 (and hence exactly two points in each bucket belong to F_2). The pseudograph on n vertices corresponding to F_i is 2-regular, for $i = 1, 2$. Call such a partition an F-decomposition of P, and write $(F_1, F_2) \vdash P$. Now F_i gives rise to a unique pairing $\tilde{F}_i \in \mathcal{P}_{n,2}$, for $i = 1, 2$, using the ordering on the points in P to order the points in each bucket in $\tilde{F}_i$. Hence we can define $\kappa(F_i) = \kappa(\tilde{F}_i)$, $G(F_i) = G(\tilde{F}_i)$, and so on. In what follows we will drop the tilde sign where there is no chance of confusion. Now define $Y(P)$ by

$$Y(P) = \sum_{(F_1,F_2)\vdash P} (2\theta_1)^{\kappa(F_1)-1} (2\theta_2)^{\kappa(F_2)-1} \tag{1}$$

where the sum is over all F-decompositions of P. Note that

$$\mathcal{P}_{n,4}^{(Y)} = \hat{\mathcal{F}}_n(\theta_1) + \hat{\mathcal{F}}_n(\theta_2),$$

since $Y(P)$ is proportional to $\mathbf{P}_{\hat{\mathcal{F}}_n(\theta_1)+\hat{\mathcal{F}}_n(\theta_2)}(P)$. Therefore it suffices to establish that Y satisfies the requirements of Theorem 1.4. The variables X_i will be the number of cycles of length i in the pairing.

The following lemma will be useful.

Lemma 1.1. *If $c \geq 0$ then*

$$\sum_{F \in \mathcal{P}_{n,2}} c^{\kappa(F)-1} \sim \frac{2^n\, n!}{2\Gamma(c/2+1)}\, n^{c/2-1}.$$

Proof. It is not difficult to see that

$$\sum_{F\in\mathcal{P}_{n,2}} c^{\kappa(F)} = 2^n \sum_{\sigma\in\mathrm{Sym}(n)} \left(\frac{c}{2}\right)^{\kappa(\sigma)}.$$

Assume for the time being that $c > 0$. Using [13, (3.5.2)] for example, we can write

$$\sum_{F\in\mathcal{P}_{n,2}} c^{\kappa(F)} = 2^n \frac{c}{2}\left(\frac{c}{2}+1\right)\cdots\left(\frac{c}{2}+n-1\right) = \frac{2^n\,\Gamma(c/2+n)}{\Gamma(c/2)}.$$

Therefore

$$\sum_{F\in\mathcal{P}_{n,2}} c^{\kappa(F)-1} = \frac{2^n\Gamma(c/2+n)}{2\Gamma(c/2+1)} \sim \frac{2^n\, n!}{2\Gamma(c/2+1)}\, n^{c/2-1},$$

and this equation holds by continuity even when $c = 0$. □

Taking $c = 1$ we recover the standard formula

$$|\mathcal{P}_{n,2}| \sim (\pi n)^{-1/2}\, 2^n n!.$$

Dividing by $|\mathcal{P}_{n,2}|$ we find:

Corollary 1.2. *If $c \geq 0$ then*

$$\mathbf{E}(c^{\kappa(F)-1}) \sim \frac{\sqrt{\pi}}{2\Gamma(c/2+1)}\, n^{(c-1)/2}$$

where the expectation is over $F \in \mathcal{P}_{n,2}$.

2 Calculating the Expectation

To calculate the expectation of Y we use a slightly different expression. For $G \in \mathcal{G}^*_{n,4}$ let $Y(G)$ be defined by

$$Y(G) = \sum_{\substack{P\in\mathcal{P}_{n,4}\\ G(P)=G}} Y(P).$$

Suppose that $G(\tilde{F}_1) \cup G(\tilde{F}_2) = G$ for $\tilde{F}_1, \tilde{F}_2 \in \mathcal{P}_{n,2}$. There are exactly 6^n ways to produce ordered pairs (F_1, F_2) such that F_i corresponds to $\tilde{F}_i$ for $i = 1, 2$, since all we must do is specify which 2 points correspond to F_1, for each bucket of 4 points. Each such ordered pair is an F-decomposition for a pairing P corresponding to G. Summing over all such $\tilde{F}_1$, $\tilde{F}_2$ gives all the ways which a pairing P corresponding to G can arise. Hence (dropping the tilde signs again), $Y(G)$ can also be written as

$$Y(G) = 6^n \sum_{\substack{(F_1,F_2)\in\mathcal{P}_{n,2}\times\mathcal{P}_{n,2}\\ G(F_1)\cup G(F_2)=G}} (2\theta_1)^{\kappa(F_1)-1}(2\theta_2)^{\kappa(F_2)-1}.$$

Therefore the expectation of $Y(P)$ in $\mathcal{P}_{n,4}$ is given by

$$\begin{aligned}
\mathbf{E}Y &= \frac{1}{|\mathcal{P}_{n,4}|} \sum_{G \in \mathcal{G}^*_{n,4}} Y(G) \\
&= \frac{6^n}{|\mathcal{P}_{n,4}|} \prod_{i=1}^{2} \left(\sum_{F \in \mathcal{P}_{n,2}} (2\theta_i)^{\kappa(F)-1} \right) \\
&\sim \frac{24^n \, n!^2}{4|\mathcal{P}_{n,4}|\Gamma(\theta_1+1)\Gamma(\theta_2+1)} \, n^{\theta_1+\theta_2-2} \\
&\sim \frac{\sqrt{2}\pi \, n^{\theta_1+\theta_2-1}}{4\Gamma(\theta_1+1)\Gamma(\theta_2+1)} \left(\frac{3}{2}\right)^n, \qquad (2)
\end{aligned}$$

using Lemma 1.1 and the formula

$$|\mathcal{P}_{n,4}| = |\mathcal{P}_{2n,2}| \sim (2\pi n)^{-1/2} \, 2^{2n} (2n)! \sim \frac{2^{4n} \, n!^2}{\sqrt{2}\pi n}.$$

3 Calculating the Variance

We are interested in the expected value of $Y(P)^2$. Now

$$\mathbf{E}(Y^2) = \frac{1}{|\mathcal{P}_{n,4}|} \sum_{P \in \mathcal{P}_{n,4}} \sum_{(F_1,F_2) \vdash P} (2\theta_1)^{\kappa(F_1)-1} (2\theta_2)^{\kappa(F_2)-1} \sum_{(F_3,F_4) \vdash P} (2\theta_1)^{\kappa(F_3)-1} (2\theta_2)^{\kappa(F_4)-1}.$$

We follow the same steps used by Kim and Wormald [7] in the analysis of the variance of H_2 (the number of H-decompositions of a pairing in $\mathcal{P}_{n,4}$, where an H-decomposition is a partition of a pairing into two subsets, each of which corresponds to a Hamilton cycle). In fact, it will be convenient later to compare our analysis directly with the analysis of $\mathbf{E}({H_2}^2)$ given in [7], since the most technical part of both arguments is identical.

Given $P \in \mathcal{P}_{n,4}$, let $((F_1, F_2), (F_3, F_4))$ be an ordered pair of F-decompositions of P. A pair in P is of type (i, j) if it belongs to F_i and F_{2+j}, for $1 \le i, j \le 2$. We use the same notation for the corresponding edge in $G(P)$. As in [7], a vertex in $G(P)$ is said to have

- type A if an edge of each type is incident with it,
- type B if there are two $(1, 1)$-edges and two $(2, 2)$-edges incident with it,
- type C if there are two $(1, 2)$-edges and two $(2, 1)$-edges incident with it

(note that these are the only possibilities). Consider the edges of a given type (i, j). Each such edge lies on either a closed cycle, or a path which starts and ends in type A vertices. These cycles and paths are disjoint. Each type A vertex is the endpoint of a path of type (i, j). It follows that the number of type A vertices is even. Moreover, a closed cycle of (i, j) edges contains only type B vertices if $i = j$, and it contains only type C vertices if $i \neq j$. Each type B vertex must lie on a path or cycle of type $(1, 1)$ edges, and a path or cycle of type $(2, 2)$ edges. Similarly, each type C vertex must lie on a path or cycle of type $(1, 2)$ edges, and a path or cycle of type $(2, 1)$ edges.

By ignoring the type B and type C vertices, the edges of any particular type give rise to a perfect matching on the type A vertices. The union of these four perfect matchings gives a labelled 4-regular pseudograph on the type A vertices, called a *connection scheme.* We proceed by calculating the contribution to the weight made by all possible connection schemes, and then the contribution made from all possible ways to add back the type B and type C vertices. Finally we calculate the number of pairings corresponding to each such configuration.

Suppose that there are $a = 2k$ type A vertices, b type B vertices and c type C vertices, so that $c = n - 2k - b$. There are

$$\binom{n}{2k}\binom{n-2k}{b} = \frac{n!}{(2k)!\, b!\, c!}$$

ways to assign types to vertices. The contribution from all possible connection schemes is given by

$$S(2k) = \sum_{(M_1,M_2,M_3,M_4)} (2\theta_1)^{\kappa(M_1\cup M_2)+\kappa(M_2\cup M_3)-2}\,(2\theta_2)^{\kappa(M_3\cup M_4)+\kappa(M_4\cup M_1)-2},$$

where M_ℓ is a perfect matching on $[2k]$ for $1 \le \ell \le 4$. To see why this is the correct quantity to analyse, let M_1, M_2, M_3, M_4 be the matchings induced by the edges of types $(1, 2)$, $(1, 1)$, $(2, 1)$ and $(2, 2)$, respectively. Then each $M_\ell \cup M_{\ell+1}$ forms the edges of the connection scheme induced by one of the permutation pseudographs F_i, namely $M_1 \cup M_2 = F_1$, $M_2 \cup M_3 = F_3$, $M_3 \cup M_4 = F_2$ and $M_4 \cup M_1 = F_4$. Thus for example, a cycle in $M_1 \cup M_2$ gives rise to a (possibly larger) cycle in F_1. Therefore we weight each such cycle by a factor of $2\theta_1$. Analysis of $S(2k)$ is deferred until Section 6, where we prove in Theorem 6.1 that

$$S(2k) \sim \frac{\pi^2 k^{2t-2}\,(2k)!^4}{16^{k+1}\Gamma(\theta_1+1)^2\Gamma(\theta_2+1)^2\,k!^4}$$

where $t = \theta_1 + \theta_2$.

(This assumes that $k \to \infty$. However, this behaviour of k can be assumed, as we can see by applying the following analysis to the case of bounded k.)

We now have to add in the type B and type C vertices. We can think of the four edge-labels (i, j) as colours. First consider adding the type B vertices onto type $(1, 1)$ edges. Recall that all the vertices are already labelled. Some of the type B vertices slide on to the edges between the type A vertices. If there are s vertices on an edge then there are $s!$ ways to arrange them. Hence the e.g.f. for vertices on a single edge is $1/(1 - x)$, so the e.g.f. for vertices on the k edges is $1/(1 - x)^k$. Note that sliding type B vertices onto these edges does not affect the number of cycles in the resulting pairing. Any remaining type B vertices form closed cycles of type $(1, 1)$ edges. Any such cycle of $(1,1)$ edges forms a cycle in both F_1 and F_3 (indeed any monochromatic cycle forms a cycle in two of the F_i). Hence each such cycle should contribute weight factor $4{\theta_1}^2$ to the corresponding pairing. For reasons which will become apparent later, we will instead weight an s-cycle of $(1,1)$ edges by a factor of $4{\theta_1}^2$ if $s \geq 3$, and by a factor of $2{\theta_1}^2$ otherwise. Hence the e.g.f. for these weighted cycles is

$$2{\theta_1}^2 \sum_{s \geq 1} \frac{x^s}{s} = -2{\theta_1}^2 \log(1-x). \tag{3}$$

It follows that the e.g.f. for unions of these weighted cycles is

$$\frac{1}{(1-x)^{2{\theta_1}^2}}.$$

Finally, the total contribution from adding b vertices of type B onto the $(1,1)$ edges is

$$b!\,[x^b]\,\frac{1}{(1-x)^k} \cdot \frac{1}{(1-x)^{2{\theta_1}^2}} = b! \binom{k+b+2{\theta_1}^2-1}{b} \sim \left(\frac{k+b}{k}\right)^{2{\theta_1}^2-1} \frac{(k+b)!}{k!}.$$

(Here $[x^b]$ means extraction of coefficients.) Note that the above calculation assumes that $k \to \infty$. This behaviour of k can be assumed, as above.

When adding b vertices of type B onto the $(2,2)$ edges, the contribution is

$$\left(\frac{k+b}{k}\right)^{2{\theta_2}^2-1} \frac{(k+b)!}{k!}.$$

For type C vertices, a closed cycle of either $(1,2)$ or $(2,1)$ edges should be weighted $4\theta_1\theta_2$. However, we will weight it by $4\theta_1\theta_2$ if it has length at least 3, and give it weight $2\theta_1\theta_2$ otherwise. The e.g.f. for unions of these weighted cycles is

$$\frac{1}{(1-x)^{2\theta_1\theta_2}}$$

and the total contribution from adding c vertices of type C onto the $(1,2)$ edges, say, is

$$c!\,[x^c]\,\frac{1}{(1-x)^k} \cdot \frac{1}{(1-x)^{2\theta_1\theta_2}} = c! \binom{k+c+2\theta_1\theta_2-1}{c} \sim \left(\frac{k+c}{k}\right)^{2\theta_1\theta_2-1} \frac{(k+c)!}{k!}.$$

The contribution from adding c vertices of type C onto the $(2,1)$ edges is identical.

Identification of the two sets of type B vertices (those on type $(1,1)$ edges and those on type $(2,2)$ edges) is immediate, since they are already labelled, and gives no extra factor in the weighting. Similarly, there is no extra factor from identifying the two sets of type C vertices.

We now have a 4-regular pseudograph with edge labels (i,j). It remains to calculate the number of pairings corresponding to this pseudograph. There are 4! ways to wire up a given vertex into a pairing, unless the vertex lies in a monochromatic 1-cycle or a monochromatic 2-cycle. The presence of each such cycle decreases the number of corresponding pairings by a factor of 2. However, we have already penalised each monochromatic 1-cycle and 2-cycle by a factor of 2 (for example, in (3)). Therefore we can simply multiply by $4!^n$ at this stage.

Combining all this we obtain (where as above $t = \theta_1 + \theta_2$),

$$
\begin{aligned}
&|\mathcal{P}_{n,4}|\mathbf{E}(Y^2) \\
&\sim \sum_k \sum_b \frac{n!}{(2k)!b!c!} S(2k) \left(\frac{k+b}{k}\right)^{2\theta_1{}^2+2\theta_2{}^2-2} \left(\frac{k+c}{k}\right)^{4\theta_1\theta_2-2} \frac{(k+b)!^2\,(k+c)!^2}{k!^4} 4!^n \\
&\sim \sum_k \sum_b \frac{\pi^2\, 4!^n\, (2k)!^3\, n!\, k^{2t-2}\, (k+b)^{2\theta_1{}^2+2\theta_2{}^2-2}(k+c)^{4\theta_1\theta_2-2}\, (k+b)!^2\, (k+c)!^2}{16^{k+1}\, k^{2t^2-4}\, \Gamma(\theta_1+1)^2\Gamma(\theta_2+1)^2\, b!\, c!\, k!^8}.
\end{aligned}
$$

To treat this sum, we reduce it to the special case $\theta_1 = \theta_2 = 0$, already treated in [7]. (See also [4] for the case $(\theta_1, \theta_2) = (0, \frac{1}{2})$ and [8] for the case $(\theta_1, \theta_2) = (\frac{1}{2}, \frac{1}{2})$.) We thus write the last equation as

$$
|\mathcal{P}_{n,4}|\mathbf{E}(Y^2) \sim \frac{1}{\Gamma(\theta_1+1)^2\Gamma(\theta_2+1)^2} \sum_k \sum_b k^{2t-2t^2} (k+b)^{2\theta_1{}^2+2\theta_2{}^2} (k+c)^{4\theta_1\theta_2}\, s(k,b), \quad (4)
$$

where $s(k, b)$ does not depend on θ_1 or θ_2.

In the special case $\theta_1 = \theta_2 = 0$, when $Y = H_2$, this yields

$$
|\mathcal{P}_{n,4}|\mathbf{E}(H_2{}^2) \sim \sum_k \sum_b s(k,b).
$$

This sum was analysed in [7], where it was shown that

$$
s(k,b) \sim \frac{4!^n\, \pi\sqrt{2n}}{16(k+b)(k+c)\sqrt{kbc}} \left(\frac{n}{e}\right)^{2n} f(\beta,\kappa)
$$

where here $\kappa = k/n$ (and hence has nothing to do with our κ), $\beta = b/n$ and

$$
f(\beta,\kappa) = \frac{16^\kappa\, (\beta+\kappa)^{2(\beta+\kappa)}\, (1-\kappa-\beta)^{2(1-\kappa-\beta)}}{(2\kappa)^{2\kappa}\, \beta^\beta\, (1-2\kappa-\beta)^{1-2\kappa-\beta}}.
$$

Moreover, the analysis given in [7] shows that the maximum of f occurs at $\beta = 1/6$, $\kappa = 1/3$, and thus all terms $s(k, b)$ with $(k/n, b/n)$ outside a small neighbourhood around $(1/3, 1/6)$ give an exponentially small contribution to the double sum. It follows that the sum in (4) is also dominated by terms with $k \sim n/3$, $b \sim n/6$, $c \sim n/6$. Thus

$$
\begin{aligned}
|\mathcal{P}_{n,4}|\mathbf{E}(Y^2) &\sim \frac{1}{\Gamma(\theta_1+1)^2\Gamma(\theta_2+1)^2} \sum_k \sum_b \left(\frac{n}{3}\right)^{2t-2t^2} \left(\frac{n}{2}\right)^{2\theta_1{}^2+2\theta_2{}^2} \left(\frac{n}{2}\right)^{4\theta_1\theta_2} s(k,b) \\
&= \frac{2^{-2t^2}3^{2t^2-2t}}{\Gamma(\theta_1+1)^2\Gamma(\theta_2+1)^2} n^{2t} \sum_k \sum_b s(k,b) \\
&\sim \frac{2^{-2t^2}3^{2t^2-2t}}{\Gamma(\theta_1+1)^2\Gamma(\theta_2+1)^2} n^{2t}\, |\mathcal{P}_{n,4}|\, \mathbf{E}(H_2{}^2).
\end{aligned}
$$

Moreover, by comparing with the case $\theta_1 = \theta_2 = 0$ in (2), we find

$$\mathbf{E}Y \sim \frac{n^t}{\Gamma(\theta_1+1)\Gamma(\theta_2+1)}\mathbf{E}H_2.$$

Now recall from [7] that

$$\frac{\mathbf{E}({H_2}^2)}{(\mathbf{E}H_2)^2} \sim \sqrt{24}.$$

It follows that

$$\frac{\mathbf{E}(Y^2)}{(\mathbf{E}Y)^2} \sim 2^{-2t^2}\, 3^{2t^2-2t}\, \frac{\mathbf{E}({H_2}^2)}{(\mathbf{E}H_2)^2} = 2^{3/2-2t^2}\, 3^{2t^2-2t+1/2}. \tag{5}$$

4 Effect of short cycles

Again, the argument follows the steps used by Kim and Wormald in [7]. For fixed $k \geq 1$, we must calculate

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

where C_k is the number of k-cycles in $\mathcal{P}_{n,4}$. (Here a k-cycle is a set of k pairs in P which correspond to a k-cycle in $G(P)$.) 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 k-cycles C. Let C be a labelled, directed cycle on k vertices. There are

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

choices for C. Edges of the cycle will correspond to pairs in a pairing $P \in \mathcal{P}_{n,4}$, with an F-decomposition (F_1, F_2). Nominate which 2 points in each bucket will be joined to pairs in F_1. There are 6^n ways to nominate these points. Now we only need specify F_1 and F_2 as pairings on n buckets, each with 2 points. There are three possibilities:

(i) $C \cap F_1 = C$,

(ii) $C \cap F_1 = \emptyset$,

(iii) $C \cap F_1$ is the disjoint union of i paths, for some $i \geq 1$.

First consider case (i). Here C forms a closed cycle in F_1. We must divide by 2 if C has length 3 or more, to undirect the cycle. Then wire $C \cap F_1$ into the specified points of the subpairing F_1. There are 2^k ways to do this if $k \geq 3$, and 2^{k-1} ways if $k = 1, 2$. So in either case, the net effect is that of multiplying by 2^{k-1}. Finally, we multiply by $2\theta_1$ to assign the appropriate weight to C. The contribution from extending $C \cap F_1$ to F_1 is

$$\sum_{F \in \mathcal{P}_{n-k,2}} (2\theta_1)^{\kappa(F)-1} \sim \frac{2^{n-k}\,(n-k)!\,(n-k)^{\theta_1-1}}{2\Gamma(\theta_1+1)} \sim \frac{2^{n-k}\,n!\,n^{\theta_1-1}}{2\,n^k\,\Gamma(\theta_1+1)},$$

while the contribution from F_2 is just

$$\sum_{F \in \mathcal{P}_{n,2}} (2\theta_2)^{\kappa(F)-1} \sim \frac{2^n\,n!\,n^{\theta_2-1}}{2\Gamma(\theta_2+1)},$$

using Lemma 1.1. So using (2), the total contribution from case (i) is asymptotically equal to

$$\frac{\theta_1\,24^n\,n!^2\,n^{\theta_1+\theta_2-2}}{4k\,\Gamma(\theta_1+1)\Gamma(\theta_2+1)} \sim |\mathcal{P}_{n,4}|\,\mathbf{E}Y\,\frac{\theta_1}{k}.$$

From case (ii) we similarly get a contribution of

$$|\mathcal{P}_{n,4}|\,\mathbf{E}Y\,\frac{\theta_2}{k}.$$

Now consider case (iii). Here $C \cap F_2$ is also the disjoint union of i paths. Colour pairs of F_1 red and pairs of F_2 blue. Each path (whether red or blue) consists of at least 1 edge. The number of ways to grow these $2i$ paths, starting from the least vertex in C and forming the paths in sequence red, blue, ... in the direction of the cycle, is

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

(see [7]). Shift the starting vertex around the cycle by 1 place, k times. Having done this for each directed cycle, we have counted each undirected configuration exactly $2i$ times. To account for this we multiply by $k/(2i)$, and note that now we are working with undirected cycles.

Now we must wire $F_i \cap C$ into the specified points of the subpairing F_i, and complete F_i, for $i = 1, 2$. Suppose that there are j edges in $C \cap F_1$. Then there are $i+j$ nonisolated vertices in $C \cap F_1$, and hence there are 2^{i+j} ways to wire up $F_1 \cap C$ into the specified points of F_1. Shrink each of the i red paths in the subpairing down to a single bucket. This gives $n-j$ buckets in total, each with 2 specified points. If we calculate the contribution from all possible pairings on these points, and then expand the red paths in $C \cap F_1$ back to their original size, we will have calculated the contribution from all ways of completing F_1. This is equal to

$$\sum_{F \in \mathcal{P}_{n-j,2}} (2\theta_1)^{\kappa(F)-1} \sim \frac{2^{n-j}\,(n-j)!\,(n-j)^{\theta_1-1}}{2\,\Gamma(\theta_1+1)} \sim \frac{2^{n-j}\,n!\,n^{\theta_1-1}}{2\,n^j\,\Gamma(\theta_1+1)} \qquad (6)$$

using Lemma 1.1. Similarly, there are $2^{i+(k-j)}$ ways to wire the blue paths into the specified points of the subpairing F_2, and the contribution from all ways of extending $C \cap F_2$ to F_2 is asymptotically

$$\frac{2^{n-(k-j)}\,(n-(k-j))!\,(n-(k-j))^{\theta_2-1}}{2\Gamma(\theta_2+1)} \sim \frac{2^{n-(k-j)}\,n!\,n^{\theta_2-1}}{2\,n^{k-j}\,\Gamma(\theta_2+1)}.$$

Thus each decomposition of C into $2i$ paths with j red edges contributes asymptotically

$$2^{i+j}\cdot\frac{2^{n-j}\,n!\,n^{\theta_1-1}}{2\,n^j\,\Gamma(\theta_1+1)}\cdot 2^{i+k-j}\cdot\frac{2^{n-(k-j)}\,n!\,n^{\theta_2-1}}{2\,n^{k-j}\,\Gamma(\theta_2+1)} = \frac{2^{2n+2i}\,n!^2\,n^{\theta_1+\theta_2-2}}{4n^k\,\Gamma(\theta_1+1)\Gamma(\theta_2+1)}.$$

Since this does not depend on j, and using (2) again, we find that the total contribution from case (iii) is asymptotically equal to

$$\begin{aligned}
& 6^n\cdot\frac{n^k}{k}\sum_{i\geq 1}[x^k]\left(\frac{x}{(1-x)}\right)^{2i}\cdot\frac{k}{2i}\,\frac{2^{2n+2i}n!^2n^{\theta_1+\theta_2-2}}{4n^k\Gamma(\theta_1+1)\Gamma(\theta_2+1)}\\
= \;& \frac{24^n\,n!^2\,n^{\theta_1+\theta_2-2}}{4\Gamma(\theta_1+1)\,\Gamma(\theta_2+1)}[x^k]\sum_{i\geq 1}\left(\frac{4x^2}{(1-x)^2}\right)^i\frac{1}{2i}\\
\sim \;& \mathbf{E}Y|\mathcal{P}_{n,4}|[x^k]\sum_{i\geq 1}\left(\frac{4x^2}{(1-x)^2}\right)^i\frac{1}{2i}\\
= \;& \mathbf{E}Y|\mathcal{P}_{n,4}|\cdot\frac{-2+(-1)^k+3^k}{2k}
\end{aligned}$$

as in [7].

Putting these calculations together, we obtain

$$\frac{\mathbf{E}(YC_k)}{\mathbf{E}Y} \sim \frac{t}{k}+\frac{-2+(-1)^k+3^k}{2k} = \frac{2t-2+(-1)^k+3^k}{2k} = \rho_k$$

for $k \geq 1$, where $t = \theta_1+\theta_2$. A direct generalisation of this argument, applied to an ordered set of i_1 cycles of length 1, i_2 cycles of length 2, and so on, shows that

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

5 Synthesis

We now combine the results of Sections 2 – 4. One further piece of information is required, namely the short cycle distribution in $\mathcal{P}_{n,4}$. As is well known, in $\mathcal{P}_{n,4}$ the variables C_k are asymptotically independent Poisson random variables with expectations

$$\mathbf{E}_{\mathcal{P}_{n,4}}(C_k) \sim \lambda_k = \frac{3^k}{2k}.$$

Recall ρ_k from the previous section, and define δ_k by

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

Then

$$\begin{aligned}
\exp\left(\sum_{k=1}^{\infty} \lambda_k \, {\delta_k}^2\right) &= \exp\left(\frac{1}{2}\sum_{k=1}^{\infty} \frac{(4t^2 - 8t + 5) + (4t - 4)(-1)^k}{k\, 3^k}\right) \\
&= \exp\left(\frac{1}{2}\left(-(4t^2 - 8t + 5)\log\left(\frac{2}{3}\right) - (4t - 4)\log\left(\frac{4}{3}\right)\right)\right) \\
&= 3^{2t^2 - 2t + 1/2}\, 2^{3/2 - 2t^2}.
\end{aligned}$$

But from (5) we also know that

$$\frac{\mathbf{E}(Y^2)}{(\mathbf{E}Y)^2} \sim 3^{2t^2 - 2t + 1/2}\, 2^{3/2 - 2t^2} = \exp\left(\sum_{k=1}^{\infty} \lambda_k \, {\delta_k}^2\right).$$

Note further that for $k \geq 2$, $\rho_k > 0$ and thus $\delta_k > -1$ for all $t \geq 0$. Also $\delta_1 = 2t/3 - 1$ which equals -1 if and only if $t = 0$; that is, $\theta_1 = \theta_2 = 0$. Therefore, by Theorem 1.4, when $t > 0$ we have

$$\hat{\mathcal{F}}_n(\theta_1) + \hat{\mathcal{F}}_n(\theta_2) \approx \mathcal{P}_{n,4},$$

and when $t = 0$ we have

$$\hat{\mathcal{F}}_n(\theta_1) + \hat{\mathcal{F}}_n(\theta_2) \approx \mathcal{P}'_{n,4},$$

as claimed.

Notice that $\mathbf{E}Y^2/(\mathbf{E}Y)^2$ depends asymptotically only on $t = \theta_1 + \theta_2$, and not on θ_1, θ_2 themselves. Similarly, δ_k depends only on k and t, for $k \geq 1$.

6 Random matchings

Recall from Section 3 that we wish to analyse

$$S(2k) = \sum_{(M_1, M_2, M_3, M_4)} (2\theta_1)^{\kappa(M_1 \cup M_2) + \kappa(M_2 \cup M_3) - 2}\, (2\theta_2)^{\kappa(M_3 \cup M_4) + \kappa(M_4 \cup M_1) - 2},$$

where M_i is a perfect matching on $[2k]$ for $1 \leq i \leq 4$. We will prove the following result, which is used in Section 3. (Throughout this section we use n in place of k.)

Theorem 6.1. *With notation as above,*

$$S(2n) \sim \frac{\pi^2\, n^{2\theta_1 + 2\theta_2 - 2}\, (2n)!^4}{16^{n+1}\, \Gamma(\theta_1 + 1)^2 \Gamma(\theta_2 + 1)^2\, n!^4}.$$

Now switch language slightly and let B, R, S, S_1 and S_2 be perfect matchings of $[2n]$. (The letters stand for Blue, Red and Silver.) Write BR for $B \cup R$, and so on. Also write $\mathbf{E}_R(\cdot)$ to denote expectation over all uniformly chosen perfect matchings R of $[2n]$. Let $m(2n)$ denote the number of perfect matchings on $[2n]$, so that

$$m(2n) = \frac{(2n)!}{2^n\, n!}.$$

Now, letting $B = M_1$, $R = M_3$, $S_1 = M_2$ and $S_2 = M_4$, we obtain

$$\begin{aligned} S(2n) &= \sum_{B,R} \sum_{S_1} \sum_{S_2} (2\theta_1)^{\kappa(BS_1)+\kappa(RS_1)-2}\, (2\theta_2)^{\kappa(BS_2)+\kappa(RS_2)-2} \\ &= m(2n)^2 \sum_{B,R} X_{BR}(\theta_1)\, X_{BR}(\theta_2), \end{aligned}$$

where

$$X_{BR}(\theta) = \mathbf{E}_S((2\theta)^{\kappa(BS)+\kappa(RS)-2})$$

for any constant $\theta \geq 0$. Let $\mu(\theta) = \mathbf{E}_R(X_{BR}(\theta))$. (By symmetry, this does not depend on B.)

Lemma 6.1. *For every $\theta \geq 0$,*

$$\mu(\theta) \sim \frac{\pi\, n^{2\theta-1}}{4\,\Gamma(\theta+1)^2}.$$

Proof. We have

$$\begin{aligned} \mu(\theta) &= \mathbf{E}_R\left(\mathbf{E}_S((2\theta)^{\kappa(BS)+\kappa(RS)-2})\right) \\ &= \mathbf{E}_S((2\theta)^{\kappa(BS)-1})\mathbf{E}_R((2\theta)^{\kappa(RS)-1}) \\ &= \mathbf{E}_S((2\theta)^{\kappa(BS)-1})^2 \\ &\sim \frac{\pi\, n^{2\theta-1}}{4\,\Gamma(\theta+1)^2}. \end{aligned}$$

This follows by Corollary 1.2 with $c = 2\theta$, since adding a random silver matching to a fixed blue matching is the same as choosing a random element of $\mathcal{P}_{n,2}$, where the blue matching defines the pairs of the pairing. ☐

We will show below that $X_{BR}(\theta_1)X_{BR}(\theta_2) = \mu(\theta_1)\mu(\theta_2)(1 + o(1))$ for almost all perfect matchings B, R. For other pairs of perfect matchings we will use this result.

Lemma 6.2. *Let $\theta \geq 0$. For all perfect matchings B, R on $[2n]$, we have*

$$X_{BR}(\theta) = O\left(n^{2\theta^2-2\theta+1/2}\, \mu(\theta)\right).$$

Proof. Using the Cauchy-Schwarz inequality, we have

$$X_{BR}(\theta) \le \sqrt{X_{BB}(\theta)X_{RR}(\theta)} = \mathbf{E}_S((4\theta^2)^{\kappa(BS)-1}).$$

Using Lemma 6.1 and applying Corollary 1.2 with $c = 4\theta^2$, we find this is asymptotically equal to

$$\frac{\sqrt{\pi}\, n^{(4\theta^2-1)/2}}{2\Gamma(2\theta^2+1)} = O\left(n^{2\theta^2-2\theta+1/2}\mu(\theta)\right),$$

as claimed. □

Given B and R, we can use a stochastic process to generate a perfect matching S on $[2n]$, uniformly at random. We do this in such a way that we can detect whether a cycle has been formed in BS or RS. Let $BR(0) = BR$ and start with S empty. At time t, we choose $x(t)$ uniformly at random from the shortest cycle of $BR(t-1)$, and select $y(t)$ uniformly at random from all other vertices. The edge $\{x(t), y(t)\}$ is added to the silver matching S. If $\{x(t), y(t)\}$ is equal to a blue (respectively red) edge of $BR(t-1)$ then we have created a new cycle in the evolving graph BS (respectively RS). It is possible to create a cycle in both, if and only if $x(t)$ lies in a 2-cycle of $BR(t-1)$ and we choose $y(t)$ to be its unique neighbour.

We form the graph $BR(t)$ from $BR(t-1)$ by deleting $x(t)$ and $y(t)$; we further join the widowed B-neighbour of $x(t)$ to the B-neighbour of $y(t)$, unless $x(t)$ and $y(t)$ were neighbours of each other, and similarly for R-neighbours. This is called *contracting* the edge $\{x(t), y(t)\}$. This process is used by Kim and Wormald in [7].

The proof of Theorem 6.1 uses a *coupling* of two instances of this process. A coupling of two stochastic processes consists of a simultaneous realization of (copies of) the processes.

Given three matchings B, R_1 and R_2 on $[2n]$, we define a coupling $\{BR_1(t),\, BR_2(t)\}_{t=1}^{n}$ for the process described above. The coupling produces silver matchings (S_1, S_2) with uniform marginal probabilities. Let $BR_i(0) = BR_i$, for $i = 1, 2$ and let $S_1 = S_2 = \emptyset$. The transitions of the coupling are as follows: for $t \ge 1$ do

(i) choose $x_i(t)$ arbitrarily from the shortest cycle in $BR_i(t-1)$, for $i = 1, 2$,

(ii) choose $y_i(t)$ uniformly at random from all remaining vertices of $BR_i(t-1)$, for $i = 1, 2$, subject to one condition: if $BR_1(t-1)$ and $BR_2(t-1)$ are non-isomorphic, then $y_1(t)$ and $y_2(t)$ are chosen independently, but if $BR_1(t-1)$ and $BR_2(t-1)$ are isomorphic, then

- let $\varphi : BR_1(t-1) \to BR_2(t-1)$ be a colour preserving isomorphism such that $x_2(t) = \varphi(x_1(t))$,
- choose $y_1(t)$ uniformly at random and let $y_2(t) = \varphi(y_1(t))$.

(iii) add $\{x_i(t), y_i(t)\}$ to S_i and form $BR_i(t)$ by deleting or contracting the edge $\{x_i(t),\, y_i(t)\}$, for $i = 1, 2$, as described above.

So the marginal processes $BR_i(t)$ evolve independently until they become isomorphic, at which point they couple and evolve together. In particular, if $BR_1(t)$ and $BR_2(t)$ both are hamiltonian, they stay coupled for the rest of the process.

For convenience we define $BR_i(t) = BR_i(n)$ for $t > n$.

Lemma 6.3. *For every $A > 0$ and $b > 0$ there exists $K = K(A, b) > 0$ such that if $\kappa(BR_i) \le A \log n$ for $i = 1, 2$, then*

$$\mathbf{P}\big(\kappa(BR_1(t)) = \kappa(BR_2(t)) = 1 \textit{ for some } t \le K \log n\big) = 1 - O(n^{-b}).$$

In particular, with probability $1 - O(n^{-b})$, the processes couple before $K \log n$.

Proof. For ease of notation, write

$$\kappa_i(t) = \kappa(BR_i(t))$$

for $i = 1, 2$ and $t \ge 0$. Let

$$\kappa(t) = \max\{\kappa_1(t), \kappa_2(t)\}.$$

Since $x_i(t)$ is chosen from the smallest cycle of $BR_i(t-1)$, and κ_i decreases by 1 when $x_i(t)$ and $y_i(t)$ are in different components, it follows that

$$\mathbf{P}(\kappa_i(t+1) = \kappa_i(t) - 1) \ge \frac{\kappa_i(t) - 1}{\kappa_i(t)}$$

for $i = 1, 2$. Thus, for any given $BR_1(t)$ and $BR_2(t)$ with $\kappa(t) \ge 4$,

$$\begin{aligned} \mathbf{P}(\kappa(t+1) < \kappa(t)) &\ge \mathbf{P}(\kappa_1(t+1) = \kappa_1(t) - 1) \cdot \mathbf{P}(\kappa_2(t+1) = \kappa_2(t) - 1) \\ &\ge \frac{\kappa_1(t) - 1}{\kappa_1(t)} \cdot \frac{\kappa_2(t) - 1}{\kappa_2(t)} \\ &\ge \frac{9}{16}. \end{aligned}$$

Moreover, $\kappa(t)$ never changes by more than 1. Thus $\kappa(t)$ behaves like a random walk with negative drift as long as $\kappa(t) \ge 4$.

Let $W(t)$ be defined by $W(0) = \kappa(0)$ and $W(t) = W(t-1) + Q(t)$, for $t \ge 1$, where the $Q(t)$ are independent variables each with the distribution

$$Q(t) = \begin{cases} -1 & \text{with probability } 9/16, \\ 1 & \text{with probability } 7/16. \end{cases}$$

Moreover, let $N(t)$ be the number of times $j = 0, \ldots, t$ such that $\kappa(j) \le 3$, with $N(-1) = 0$. Then the process $\{\kappa(t) - 2N(t-1)\}$ is stochastically dominated by the process $\{W(t)\}$. (When $\kappa(t) \le 3$ this follows since both $\kappa(t)$ and $W(t)$ change by at most 1 per step.)

Note that $\mathbf{E}Q(t) < 0$, and that we thus can find $\alpha > 0$ sufficiently small such that $\mathbf{E}e^{\alpha Q(t)} < 1$. We write $\mathbf{E}e^{\alpha Q(t)} = e^{-\beta}$, where $\beta > 0$. Then

$$\mathbf{E}e^{-2\alpha N(t)} \le \mathbf{E}e^{\alpha(\kappa(t) - 2N(t))} \le \mathbf{E}e^{\alpha W(t)} = e^{\alpha\kappa(0)}\Big(\mathbf{E}e^{\alpha Q(1)}\Big)^t = e^{\alpha\kappa(0) - t\beta}. \tag{7}$$

Let $D = 2b/\log(16/15)$ and $K = (\alpha A + 2\alpha D + b)/\beta$. Note that D and K are both positive. Then, by (7) and our assumption $\kappa(0) \le A \log n$, for $t \ge K \log n$, Markov's inequality gives

$$\mathbf{P}(N(t) \le D \log n) \le e^{2\alpha D \log n} \mathbf{E} e^{-2\alpha N(t)} \le e^{2\alpha D \log n + \alpha\kappa(0) - t\beta} \le n^{2\alpha D + \alpha A - \beta K} = n^{-b}. \quad (8)$$

Let $0 \le T_1 < T_2 < \dots$ be an enumeration of the (random) times $\{t : \kappa(t) \le 3\}$. Each time we have $\kappa(t) \le 3$, the probability that $\kappa(t+2) = 1$ is at least $1/16$. This follows by direct calculations showing that $\mathbf{P}(\kappa_i(t+1) = \kappa_i(t) - 1) \ge \frac{1}{2}$. Thus

$$\mathbf{P}(\kappa(t+1) = \kappa(t) - 1) \ge \frac{1}{4}$$

for any $BR_1(t)$, $BR_2(t)$. Consequently, the events

$$\mathcal{E}_k = \{\kappa(t) = 1 \text{ for some } t \text{ with } T_{2k-1} \le t \le T_{2k+1}\}, \qquad k = 1, 2, \dots$$

all have probabilities at least $1/4$, also conditioned on the history up to T_{2k-1}, and consequently, for any $k \ge 0$

$$\mathbf{P}\big(\kappa(t) > 1 \text{ for all } t \le T_{2k+1}\big) = \mathbf{P}\Big(\bigcap_{j=1}^{k} \overline{\mathcal{E}_j}\Big) \le \Big(\frac{15}{16}\Big)^k.$$

Choosing $k = \frac{1}{2}\lceil D \log n \rceil - 1$, we obtain by our choice of D

$$\mathbf{P}\big(\kappa(t) > 1 \text{ for all } t \le T_{\lceil D \log n \rceil}\big) \le \Big(\frac{15}{16}\Big)^{\frac{1}{2}} \lceil D \log n \rceil - 1 < 2n^{-b}. \quad (9)$$

Moreover, either $T_{\lceil D \log n \rceil} \le K \log n$ or $N(\lceil K \log n \rceil) \le D \log n$, so combining (8) and (9) we obtain

$$\begin{aligned} \mathbf{P}\big(\kappa(t) > 1 \text{ for all } t \le K \log n\big) &\le \mathbf{P}\big(\kappa(t) > 1 \text{ for all } t \le T_{\lceil D \log n \rceil}\big) \\ &\qquad + \mathbf{P}\big(N(\lceil K \log n \rceil) \le D \log n\big) \\ &< 3n^{-b}. \end{aligned}$$

as claimed. $\square$

Lemma 6.4. *Let $\theta \ge 0$ and $A \ge \max\{2(\theta^2+1), 5\}$ be given. If B, R are perfect matchings on $[2n]$ such that $\kappa(BR) \le A \log n$, then*

$$X_{BR}(\theta) = \mu(\theta) + O(n^{-1/4}\mu(\theta)).$$

(The choice of $1/4$ is arbitrary, and could be replaced by $1 - \varepsilon$ for any positive constant ε. We make no attempt to optimize the argument.)

Proof. This is the case $\alpha = -1/4$ of the following property, which we prove for all $\alpha \ge -3/4$ by an induction, successively improving α.

$H(\alpha):$ For each $A \geq \max\{2(\theta^2+1), 5\}$, there exists $C = C(A, \alpha, \theta)$ such that for every $n \geq 1$ and perfect matchings B and R on $[2n]$ with $\kappa(BR) \leq A \log n$,

$$|X_{BR}(\theta) - \mu(\theta)| \leq Cn^{\alpha}\mu(\theta).$$

Note first that Lemma 6.2 shows that $H(\alpha)$ holds if $\alpha \geq 2\theta^2 - 2\theta + 1/2$. (This value is always nonnegative.) Now suppose that $H(\alpha)$ holds for some $\alpha \geq 0$; we will show that then $H(\alpha - 3/4)$ holds too. From this it follows that $H(\alpha')$ holds for all $\alpha' \geq \alpha - 3/4$. This will prove the result.

Fix $A \geq \max\{2(\theta^2+1), 5\}$ and consider three matchings B, R_1, R_2 such that $\kappa(BR_i) \leq A \log n$ for $i = 1, 2$. Let $T = \lfloor K \log n \rfloor$, where K is as in Lemma 6.3 with $b = 8(\theta^4+1)$. We may assume that n is so large that $T \leq n/2$ and $\log(n-T) \geq \frac{1}{2}\log n$, since $H(\alpha - 3/4)$ trivially is satisfied for smaller n if we choose C large enough. Consider the coupling $\{BR_1(t),\ BR_2(t)\}_{t=1}^{n}$ described above. Define two events for the coupling:

$$\begin{aligned} H_1 &= \{\kappa(BR_1(t)) = \kappa(BR_2(t)) = 1 \text{ for some } t \leq T\}, \\ H_2 &= \{\text{no cycle is created in steps } 1, \ldots, T\}. \end{aligned}$$

By Lemma 6.3, $\mathbf{P}(H_1) = 1 - O(n^{-b})$. Applying the Cauchy-Schwarz inequality gives

$$\mathbf{E}_S((2\theta)^{\kappa(BS)+\kappa(R_iS)-2} \mid \overline{H_1}) \cdot \mathbf{P}(\overline{H_1}) \leq \sqrt{\mathbf{E}_S((4\theta^2)^{\kappa(BS)+\kappa(R_iS)-2})\mathbf{P}(\overline{H_1})}. \tag{10}$$

Note that

$$\mathbf{E}_S((4\theta^2)^{\kappa(BS)+\kappa(R_iS)-2}) = X_{BR_i}(\varphi)$$

where $\varphi = 2\theta^2$. Hence using Lemmas 6.1 and 6.2, we have

$$\mathbf{E}_S((4\theta^2)^{\kappa(BS)+\kappa(R_iS)-2}) = O\left(n^{2\varphi^2-2\varphi+1/2}\mu(\varphi)\right) = O\left(n^{2\varphi^2}\right) = O\left(n^{8\theta^4}\right).$$

Substituting this into (10) and applying Lemma 6.1 again, we find that

$$\begin{aligned} \mathbf{E}_S((2\theta)^{\kappa(BS)+\kappa(R_iS)-2} \mid \overline{H_1}) \cdot \mathbf{P}(\overline{H_1}) &= O\left(n^{4\theta^4-b/2}\right) \\ &= O\left(n^{4\theta^4-2\theta+1-b/2}\mu(\theta)\right) \\ &= O\left(n^{-3/4}\mu(\theta)\right) \end{aligned} \tag{11}$$

for $i = 1, 2$. Here the final equality follows since $4\theta^4 - 2\theta + 7/4 < b/2$, by choice of b.

We turn to $\overline{H_2}$. Recall that a cycle is created at step t if and only if the silver edge $\{x_i(t), y_i(t)\}$ is identical to an edge of $BR_i(t-1)$, for some $i \in \{1, 2\}$; we call this a "bad event". Write

$$\overline{H_2} = \bigcup_{t=1}^{T} F_t$$

where

$$F_t = \{\text{ the first bad event occurs in step } t\}.$$

If F_t happens, we use the Markov property of the process. Namely, the final $n-t$ steps of the process are exactly equivalent to *starting* the process with initial matchings $B(t)$ and $R_i(t)$, each on $2(n-t)$ points. Since

$$\kappa(BR_i(t)) \leq \kappa(BR_i) + t \leq A\log n + K\log n \leq 2(A+K)\log(n-t)$$

and $2(A+K) > A$, we may apply the induction hypothesis $H(\alpha)$ and Lemma 6.1 to $BR_i(t)$, with n replaced by $n-t$. This gives

$$X_{BR_i(t)}(\theta) \leq O\left((n-t)^{2\theta-1}\right)\,(1+C(n-t)^{\alpha}) = O(n^{\alpha+2\theta-1}).$$

Thus, since we have created at most two cycles in step t, and none before,

$$\begin{aligned}
\mathbf{E}((2\theta)^{\kappa(BS_1)+\kappa(R_1S_1)-2} \mid F_t) &\leq \max\left\{1,(2\theta)^2\right\}\mathbf{E}(X_{BR_1(t)}(\theta) \mid F_t)\\
&= O(n^{\alpha+2\theta-1})\\
&= O(n^{\alpha}\mu(\theta)).
\end{aligned}$$

Moreover,

$$\mathbf{P}(F_t) \leq \frac{4}{2n-2t+1} \leq \frac{4}{n}$$

and so

$$\sum_{t=1}^{T}\mathbf{E}((2\theta)^{\kappa(BS_1)+\kappa(R_1S_1)-2} \mid F_t)\mathbf{P}(F_t) = O(Tn^{\alpha-1}\mu(\theta)) = O(n^{\alpha-3/4}\mu(\theta)) \qquad (12)$$

and similarly for B, R_2 and S_2.

Conditional on $H_1 \cap H_2$, we know that no cycles were created before step T, and that coupling has occurred by step T. Thus

$$\mathbf{E}((2\theta)^{\kappa(BS_1)+\kappa(R_1S_1)-2} \mid H_1\cap H_2) = \mathbf{E}((2\theta)^{\kappa(BS_2)+\kappa(R_2S_2)-2} \mid H_1\cap H_2) \qquad (13)$$

(where the expectation is over all choices made in the coupled process).

Combining (11), (12) we find that

$$X_{BR_i}(\theta) = \mathbf{E}((2\theta)^{\kappa(BS_i)+\kappa(R_iS_i)-2} \mid H_1\cap H_2)\cdot\mathbf{P}(H_1\cap H_2) + O(n^{\alpha-3/4}\,\mu(\theta))$$

for $i=1,2$. Moreover, the first term is the same for $i=1,2$, by (13). Therefore

$$X_{BR_1}(\theta) = X_{BR_2}(\theta) + O(n^{\alpha-3/4}\mu(\theta)), \qquad (14)$$

whenever $\kappa(BR_i) \leq A\log n$, for $i=1,2$.

Now fix perfect matchings B, R on $[2n]$ such that $\kappa(BR) \leq A\log n$. It is well-known that $\mathbf{P}(\kappa(BR^*) > A\log n) = O(n^{-A})$, where R^* is a perfect matching on $2n$ chosen uniformly at random. (This follows from a lemma of Bernstein's on tail estimates for sums of random variables: see [6, Remark 2.9]. For a proof when $A \geq 5$ see [7, Lemma

1]: some positive constant lower bound on A is certainly required.) Thus, by Lemma 6.2 and (14),

$$\begin{aligned}
X_{BR}(\theta) - \mu(\theta) &= \mathbf{E}_{R^*}\big(X_{BR}(\theta) - X_{BR^*}(\theta)\big) \\
&= \mathbf{P}\big(\kappa(BR^*) > A\log n\big)\cdot O(n^{2\theta^2-2\theta+1/2}\mu(\theta)) \\
&\qquad + \mathbf{P}\big(\kappa(BR^*) \le A\log n\big)\cdot O(n^{\alpha-3/4}\mu(\theta)) \\
&= O\big(n^{2\theta^2-2\theta+1/2-A}\mu(\theta) + n^{\alpha-3/4}\mu(\theta)\big) \\
&= O(n^{\alpha-3/4}\mu(\theta)),
\end{aligned}$$

by choice of A. This shows that $H(\alpha - 3/4)$ holds, completing the proof. □

Before continuing, we make a note about the previous proof. We are particularly interested in the values $\theta \in \{0, \frac{1}{2}, 1\}$, corresponding to random Hamilton cycles, random pseudographs arising from the pairings model, and permutation pseudographs, respectively. For these values of θ, we have $0 \le 2\theta^2 - 2\theta + 1/2 \le 1/2$. So applying the inductive step of the proof of Lemma 6.4 gives the conclusion immediately, which avoids the inductive framework.

The next result looks at the expected value of $X_{BR}(\theta_1)X_{BR}(\theta_2)$ over uniform choice of the perfect matching R, with B fixed.

Lemma 6.5. *Let B be a fixed perfect matching on $[2n]$. With notation as above,*

$$\mathbf{E}_R(X_{BR}(\theta_1)X_{BR}(\theta_2)) = \mu(\theta_1)\mu(\theta_2)\,(1+o(1)).$$

Proof. Fix $A = \max\{2({\theta_1}^2 + {\theta_2}^2 + 1), 5\}$. Letting $\mathcal{E}_i$ be the event $\kappa(BR_i) > A\log(n)$, we know that $\mathbf{P}(\mathcal{E}_i) = O(n^{-A})$. Consequently, using Lemma 6.2 for $\mathcal{E}_1 \cup \mathcal{E}_2$ and Lemma 6.4 for its complement $\overline{\mathcal{E}_1} \cap \overline{\mathcal{E}_2}$,

$$\begin{aligned}
\mathbf{E}_R(X_{BR}(\theta_1)X_{BR}(\theta_2)) &= \mathbf{E}_R(X_{BR}(\theta_1)X_{BR}(\theta_2) \mid \mathcal{E}_1 \cup \mathcal{E}_2)\mathbf{P}(\mathcal{E}_1 \cup \mathcal{E}_2) \\
&\qquad + \mathbf{E}_R(X_{BR}(\theta_1)X_{BR}(\theta_2) \mid \overline{\mathcal{E}_1} \cap \overline{\mathcal{E}_2})\mathbf{P}(\overline{\mathcal{E}_1} \cap \overline{\mathcal{E}_2}) \\
&= O(n^{2{\theta_1}^2-2\theta_1+1/2}\mu(\theta_1)n^{2{\theta_2}^2-2\theta_2+1/2}\mu(\theta_2)n^{-A}) \\
&\qquad + (1 - O(n^{-A}))\big(1 + O(n^{-1/4})\big)\mu(\theta_1)\mu(\theta_2) \\
&= \mu(\theta_1)\mu(\theta_2)\,(1+o(1))
\end{aligned}$$

as required. □

Proof of Theorem 6.1. The desired result

$$S(2n) \sim \frac{\pi^2\, n^{2\theta_1+2\theta_2-2}\,(2n)!^4}{16^{n+1}\,\Gamma(\theta_1+1)^2\Gamma(\theta_2+1)^2\,n!^4}$$

follows immediately from the equation

$$S(2n) = m(2n)^3 \sum_B \mathbf{E}_R(X_{BR}(\theta_1)\,X_{BR}(\theta_2)),$$

using Lemmas 6.1 and 6.5. □

7 Further results

In this section we prove Theorem 1.2, again using the small subgraph conditioning method. The argument follows that of Sections 2 – 5 but does not require any coupling. At the end of the section we prove Theorem 1.3.

Proof of Theorem 1.2. Let $\theta \geq 0$ and $d \geq 3$ be fixed constants. Given $P \in \mathcal{P}_{n,d}$, we are interested in partitions (F, Q) of the pairs of P such that $G(F)$ is a 2-regular pseudograph and $G(Q)$ is a $(d-2)$-regular pseudograph. Let

$$Y(P) = \sum_{(F,Q)\vdash P} (2\theta)^{\kappa(F)-1}.$$

We will compare the expectation and variance of Y with the expectation and variance of Z_H, the number of H-cycles in an element of $\mathcal{P}_{n,d}$, which equals Y when $\theta = 0$. The random variable Z_H was investigated by Frieze et al. in [4]. (An H-cycle is a set of n pairs which correspond to a Hamilton cycle.)

It is straightforward to show, using Lemma 1.1, that

$$\begin{aligned} |\mathcal{P}_{n,d}|\mathbf{E}Y &= \binom{d}{2}^n \sum_{F\in\mathcal{P}_{n,2}} (2\theta)^{\kappa(F)-1}\, |\mathcal{P}_{n,d-2}| \\ &\sim d^n(d-1)^n\, n! \frac{n^{\theta-1}}{2\Gamma(\theta+1)}\, |\mathcal{P}_{n,d-2}|. \end{aligned}$$

Putting $\theta = 0$ here gives an equation for $|\mathcal{P}_{n,d}|\mathbf{E}Z_H$, from which we conclude that

$$\mathbf{E}Y \sim \frac{n^\theta}{\Gamma(\theta+1)}\, \mathbf{E}Z_H. \tag{15}$$

The variance calculation is simpler than that of Section 3, since no coupling argument is required. We calculate

$$|\mathcal{P}_{n,d}|\mathbf{E}(Y^2) = \sum_{P\in\mathcal{P}_{n,d}} \sum_{(F_1,Q_1)\vdash P} \sum_{(F_2,Q_2)\vdash P} (2\theta)^{\kappa(F_1)+\kappa(F_2)-2}.$$

First consider $F_1 \cap F_2$. Colour the corresponding edges in $G(P)$ *gold*. We imitate the argument of Section 3 for the golden edges, but argue more directly for the remaining edges. As in Section 3, say that a vertex in $G(P)$ has type A, B or C if it is incident with 1, 2 or 0 golden edges, respectively. The golden edges form a union of some paths and some closed cycles. The endpoints of these paths have type A, while the interior vertices of the paths and cycles have type B. Moreover, each vertex of type A is the endpoint of a golden path, and each vertex of type B lies on a golden path or cycle. Suppose that there are $a = 2k$ type A vertices, b type B vertices and $c = n - 2k - b$ type C vertices. Then again there are

$$\frac{n!}{(2k)!\, b!\, c!}$$

ways to assign types to these vertices. Instead of the connection scheme used in Section 3, we just choose a perfect matching of the type A vertices, to determine the endpoints of each golden path. There are

$$\frac{(2k)!}{2^k\, k!}$$

such matchings. Now add the type B vertices onto the golden paths and cycles. Each golden cycle is a cycle in F_1 and in F_2. Therefore it should contribute weight $4\theta^2$. Instead, as in Section 3, we give each golden cycle weight $4\theta^2$ if it has length greater than 2, and weight $2\theta^2$ otherwise. This underweights each cycle of length 1 or 2 by a factor of 2, but this will be corrected below. Using these weights, the contribution from adding all the type B vertices is

$$b!\binom{k+b+2\theta^2-1}{b} \sim \left(\frac{k+b}{k}\right)^{2\theta^2-1}\,\frac{(k+b)!}{k!}.$$

(This assumes that $k\to\infty$, as in Section 3.)

Now choose which points in each bucket belong to pairs in F_1 and F_2. For each vertex of type A there are d choices for the point in $F_1\cap F_2$, $d-1$ for the other point in F_1 and $d-2$ for the other point in F_2. For each vertex of type C we choose the points in F_1 in $\binom{d}{2}$ ways and then the points in F_2 in $\binom{d-2}{2}$ ways. For each vertex of type B there are $\binom{d}{2}$ ways to choose the two points in $F_1\cap F_2$. The position of this vertex on a golden path or cycle of $G(P)$ has already been determined. Multiply by 2 to count the number of ways to wire these points to their neighbours in the path or cycle. This factor $d(d-1)$ for each type B vertex gives a factor 2 too much for each golden cycle of length 1 or 2. We have already compensated for this factor in the preceding step. Thus the total factor is

$$(d(d-1)(d-2))^{2k}\,(d(d-1))^b\,\left(\binom{d}{2}\binom{d-2}{2}\right)^c = 2^{-2c}\,d^n(d-1)^n(d-2)^{2k+c}(d-3)^c.$$

At this stage, $F_1\cap F_2$ is complete determined. We now specify the remaining pairs in F_1, F_2, as well as the pairs in $Q_1\cap Q_2$. To specify the remaining pairs of F_1, delete each golden cycle and contract each golden path to a single vertex. The number of vertices left is $n-(2k+b)+k=k+c$, and each remaining vertex corresponds to a bucket with two points from $F_1\cap Q_2$ specified. (For vertices which correspond to shrunken paths, each endpoint of the path contributes one of these points.) Choose a perfect matching on these $2(k+c)$ points with the usual weighting: these determine the remaining pairs of F_1. Therefore the contribution from completing F_1 is

$$\sum_{F\in\mathcal{P}_{k+c,2}} (2\theta)^{\kappa(F)-1} \sim \frac{2^{k+c}\,(k+c)!}{2\Gamma(\theta+1)}(k+c)^{\theta-1},$$

using Lemma 1.1 and assuming that $k\to\infty$. Completing F_2 contributes the same factor. Finally, to specify the pairs of $Q_1\cap Q_2$ we need only form a perfect matching of the remaining the remaining $dn-6k-2b-4c=(d-4)n+2k+2b$ points. The number of ways to do this is

$$|\mathcal{P}_{(d-4)n+2k+2b,1}| = \frac{((d-4)n+2k+2b)!}{2^{(d/2-2)n+k+b}\,((d/2-2)n+k+b)!}.$$

Combining all these factors gives

$$|\mathcal{P}_{n,d}|\mathbf{E}(Y^2) \sim \sum_k \sum_b \left(\frac{k+b}{k}\right)^{2\theta^2} \frac{(k+c)^{2\theta}}{\Gamma(\theta+1)^2}\, t(k,b), \tag{16}$$

where

$$\begin{aligned} t(k,b) &= \frac{n!\,k\,(k+b)!\,(k+c)!^2\,d^n(d-1)^n(d-2)^{2k+c}\,(d-3)^c}{4\,k!^2\,b!\,c!\,(k+b)\,(k+c)^2} \\ &\quad \times \frac{((d-4)n+2k+2b)!}{2^{(d/2-2)n+b}\,((d/2-2)n+k+b)!} \end{aligned}$$

does not depend on θ. The case $\theta=0$ was analysed in [4] (for $d=3$, which is simpler since then $c=0$, see [9]). The method of [4] is similar but the notation differs: denoting their k and a by k' and a', we have $a'=k$ (the number of paths in $F_1\cap F_2$) and $k'=k+b$ (the number of edges in $F_1\cap F_2$). With these substitutions, (16) with $\theta=0$ is equivalent to [4, (14)]. It is shown in [4] that only terms with $k'/n\sim\kappa_0=2/d$ and $a'/n\sim\alpha_0=2(d-2)/d(d-1)$ contribute significantly to the sum. It follows that only terms with $k\sim\alpha_0 n$ and $b\sim(\kappa_0-\alpha_0)n$ contribute significantly to the sum in (16). Therefore

$$\begin{aligned} |\mathcal{P}_{n,d}|\mathbf{E}(Y^2) &\sim \left(\frac{\kappa_0}{\alpha_0}\right)^{2\theta^2} \frac{((1-\kappa_0)n)^{2\theta}}{\Gamma(\theta+1)^2} \sum_k\sum_b t(k,b) \\ &\sim \left(\frac{\kappa_0}{\alpha_0}\right)^{2\theta^2} \frac{((1-\kappa_0)n)^{2\theta}}{\Gamma(\theta+1)^2} |\mathcal{P}_{n,d}|\mathbf{E}({Z_H}^2) \\ &= \left(\frac{d-1}{d-2}\right)^{2\theta^2} \left(\frac{d-2}{d}\right)^{2\theta} \frac{n^{2\theta}}{\Gamma(\theta+1)^2} |\mathcal{P}_{n,d}|\mathbf{E}({Z_H}^2). \end{aligned}$$

Combining these with (15) we finally find that

$$\frac{\mathbf{E}(Y^2)}{(\mathbf{E}Y)^2} \sim \left(\frac{d-1}{d-2}\right)^{2\theta^2} \left(\frac{d-2}{d}\right)^{2\theta} \frac{\mathbf{E}({Z_H}^2)}{(\mathbf{E}Z_H)^2} \sim \left(\frac{d-1}{d-2}\right)^{2\theta^2} \left(\frac{d-2}{d}\right)^{2\theta-1}.$$

The second asymptotic equality follows since $\mathbf{E}({Z_H}^2)/(\mathbf{E}Z_H)^2 \sim d/(d-2)$, as shown in [4].

Finally, we must calculate the effect of short cycles. The approach is very similar to that used by Robalewska in [8, Lemma 1]. The case $C\cap F=C$ is handled as in Section 4, and gives a contribution of

$$\frac{\theta}{k}\,|\mathcal{P}_{n,d}|\mathbf{E}Y.$$

Next suppose that $C \cap F = \emptyset$. Then divide by 2 to undirect the cycle, if $k \geq 3$. There are $((d-2)(d-3))^k$ ways to wire C into the specified points of the subpairing Q, if $k \geq 3$, and $((d-2)(d-3))^k/2$ ways if $k = 1, 2$. (The net effect of this is to multiply by $((d-2)(d-3))^k/2$ for all k.) There are asymptotically

$$\frac{|\mathcal{P}_{n,d-2}|}{(n(d-2))^k}$$

ways to complete the pairing Q, since k pairs are already specified. Finally the contribution from all possible choices of F is $2^n\, n!\, n^{\theta-1}/2\Gamma(\theta+1)$, by Lemma 1.1. Hence the contribution from case (ii) is

$$\binom{d}{2}^n \cdot \frac{n^k}{k} \cdot \frac{((d-2)(d-3))^k}{2} \cdot \frac{|\mathcal{P}_{n,d-2}|}{(n(d-2))^k} \cdot \frac{2^n\, n!\, n^{\theta-1}}{2\Gamma(\theta+1)} = \frac{(d-3)^k}{2k}\, |\mathcal{P}_{n,d}| \mathbf{E} Y.$$

Finally, suppose that $C \cap F$ is the union of i disjoint paths. Suppose that these paths contain $k - r$ vertices, so that there are r isolated vertices in $C \cap F$. Let N_{ir} be the number of ways to ensure that $C \cap F$ has this form, starting without the vertices of C specified. Then Robalewska [8] shows that

$$N_{ir} \sim n^k [x^k y^i z^r] \left(-\frac{1}{2} \log((1-x)(1-xz) - x^2 y) \right).$$

There are 2^{k-r} ways to wire $C \cap F$ into the specified points of the subpairing F. The contribution from extending $C \cap F$ to F in all possible ways is

$$\sum_{F \in \mathcal{P}_{n-(k-(i+r)),2}} (2\theta)^{\kappa(F)-1} \sim \frac{2^{n-(k-(i+r))} n! n^{\theta-1}}{2\, n^{k-(i+r)}\, \Gamma(\theta+1)},$$

as in (6). Next, there are

$$(d-2)^{2i+r}(d-3)^r$$

ways to wire $C \cap Q$ into the specified points of the subpairing Q. Then the number of ways to complete $C \cap Q$ to Q is asymptotically equal to

$$\frac{|\mathcal{P}_{n,d-2}|}{(n(d-2))^{r+i}}.$$

Hence the total contribution from case (iii) is

$$|\mathcal{P}_{n,d}| \mathbf{E} Y \sum_{i \geq 1} \sum_{r \geq 0} [x^k y^i z^r] \left(-\frac{1}{2} \log((1-x)(1-xz) - x^2 y) \right) \cdot (2(d-2))^i (d-3)^r.$$

Putting this together (and mirroring the calculations in [8, Lemma 1]), noting that only the first term depends on θ, we obtain

$$\begin{aligned}
\frac{\mathbf{E}(YC_k)}{\mathbf{E}Y} &= \frac{\theta}{k} + \frac{(d-3)^k}{2k} \\
&\quad + \sum_{i\geq 1}\sum_{r\geq 0}[x^k y^i z^r]\left(-\frac{1}{2}\log((1-x)(1-xz)-x^2y)\right)\cdot(2(d-2))^i(d-3)^r \\
&= \frac{2\theta - 1 + (-1)^k + (d-1)^k}{2k} \\
&= \rho_k.
\end{aligned}$$

It is routine to verify that Y behaves well under joint short cycle distributions.

As is well-known, the distribution of k-cycles in $\mathcal{P}_{n,d}$ is asymptotically Poisson with mean

$$\lambda_k = \frac{(d-1)^k}{2k}.$$

Define δ_k for $k \geq 1$ by

$$\delta_k = \frac{\rho_k}{\lambda_k} - 1 = \frac{2\theta - 1 + (-1)^k}{(d-1)^k}.$$

Then

$$\begin{aligned}
\exp\left(\sum_{k\geq 1}\lambda_k {\delta_k}^2\right) &= \exp\left(\sum_{k\geq 1}\frac{2\theta^2 - 2\theta + 1 + (2\theta-1)(-1)^k}{k(d-1)^k}\right) \\
&= \left(\frac{d-1}{d-2}\right)^{2\theta^2-2\theta+1}\left(\frac{d-1}{d}\right)^{2\theta-1} \\
&= \frac{\mathbf{E}(Y^2)}{(\mathbf{E}Y)^2}.
\end{aligned}$$

Therefore, using Theorem 1.4 we have

$$\mathcal{P}_{n,d-2} + \hat{\mathcal{F}}_n(\theta) \approx \mathcal{P}_{n,d}$$

when $\theta > 0$ or $d > 3$, and

$$\mathcal{P}_{n,1} + \hat{\mathcal{F}}_n(0) \approx \mathcal{P}'_{n,3}$$

(since here $\rho_1 = 0$ and $\delta_1 = -1$.) This completes the proof. □

Proof of Theorem 1.3. Combine the graphs in a suitable order using Theorems 1.1 and 1.2, together with known results for sums of $\mathcal{G}^*_{n,d}$ as in the proof of [5, Theorem 9.43] or [14, Corollary 4.17]. □

References

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

[2] Friedman, J. (1991) On the second eigenvalue and random walks in random d-regular graphs, *Combinatorica* **11** 331–362.

[3] Frieze, A. (2001) Hamilton cycles in the union of random permutations, *Random Structures Algorithms* **18** 83–94.

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

[5] Janson, J. (1995) Random regular graphs: asymptotic distributions and contiguity, *Comb. Prob. Comp.* **4** 369–405.

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

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

[8] Robalewska, H. 2-factors in random regular graphs, *J. Graph Theory* **23** 215–224.

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

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

[11] Tavaré, S. and Ewens, W. J. (1995) *The Ewens Sampling Formula*, Manuscript, available from `http://www-hto.usc.edu/papers/abstracts/esfrep.html`.

[12] Tavaré, S. and Ewens, W. J. (1997) Multivariate Ewens Distribution, appeared as Chapter 41 of Johnson, N. L., Kotz, S. and Balakrishnan, N., *Discrete Multivariate Distributions*, Wiley.

[13] Wilf, H. S. (1990) *generatingfunctionology*, Academic Press.

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