Models of random regular graphs

N. C. Wormald

Summary This is a survey of results on properties of random regular graphs, together with an exposition of some of the main methods of obtaining these results. Related results on asymptotic enumeration are also presented, as well as various generalisations to random graphs with given degree sequence. A major feature in this area is the small subgraph conditioning method. When applicable, this establishes a relationship between random regular graphs with uniform distribution, and non-uniform models of random regular graphs in which the probability of a graph G is weighted according to the number of subgraphs G has of a certain type. Information can be obtained in this way on the probability of existence of various types of spanning subgraphs, such as Hamilton cycles and decompositions into perfect matchings. Uniformly distributed labelled random regular graphs receive most of the attention, but also included are several non-uniform models which come about in a natural way. Some of these appear as spin-offs from the small subgraph conditioning method, and some arise from algorithms which use simple approaches to generating random regular graphs. A quite separate role played by algorithms is in the derivation of random graph properties by analysing the performance of an appropriate greedy algorithm on a random regular graph. Many open problems and conjectures are given.

1 Introduction

Random graphs first appeared in clever probabilistic proofs by Erdős of the existence of graphs with special properties such as arbitrarily large girth and chromatic number. These had not at that time been found constructively. Much later, the study of random regular graphs took off, beginning with the works of Bender and Canfield [5], Bollobás [10] and Wormald [118, 117]. This has since been fuelled in part by applications in other areas such as computer science. An interesting application occurs in biogeography, where random non-negative integer matrices with given row and column sums are of interest. See Wilson [113], for example.

This paper is a survey of that part of random graph theory in which the degrees of vertices are restricted. Such work concentrates on regular graphs as the most interesting examples, and the results on regular graphs often extend easily to more general degree sequences. A *d-regular* graph is one with all vertices of degree d. For $d = 3$ these are often called *cubic* graphs.

The asymptotic enumeration of objects of a given type frequently goes hand in hand with the problem of generating the same objects uniformly at random, and with finding some of the limiting probabilities in the resulting probability space (see Jerrum and Sinclair [58]). However, the two problems are not equivalent.

The distinction between the approaches of exact and asymptotic enumeration is important. Probably the first result on short cycles in random regular graphs of degree at least 3 was the determination of the expected number of triangles in random cubic graphs [114]. This was done by using recurrence relations, and the asymptotic result $\frac{4}{3}$ was obtained. However, this method of exact enumeration followed by asymptotic analysis has not been able to reach any further for the type of problems that we are concerned with here. For instance, for the numbers of 2- and 3-connected cubic graphs there are recurrence relations [116], but no asymptotic formulae have been obtained from these. A more direct probabilistic approach, with an initially asymptotic viewpoint, can do much more and leads to asymptotic enumeration results in any case. For instance, many of the results in the book on random graphs by Bollobás [16] can be translated this way. In spite of this, exact and asymptotic enumeration turns out to be used rather heavily for some problems here, albeit indirectly (for example, consider the proofs in Section 4.1).

Results on random regular graphs can of course be used to show that there exist graphs with particular combinations of properties; for instance d-regular graphs which are d-connected, have arbitrarily large girth, are Hamiltonian and have no non-trivial automorphisms. There are other ways of getting regular graphs with interesting properties in a constructive way, some of them quite sophisticated (surveyed by Chung [26]), and these supply graphs with some of the properties that can be found in random graphs. A notable exception is the diameter, where the best known constructions are only within a factor $c > 1$ of that of a random d-regular graph.

However, the study of random regular graphs is recently blossoming, and some pretty results are newly emerging, such as the almost sure property that the edges can be partitioned into disjoint Hamilton cycles (when the degree is even). This survey attempts to cover all the results, and show some relationships between old and new.

1.1 Overview and notation

Early in this article we encounter the uniform model of random d-regular graphs and various properties of it, in Section 2, and then uniform models similar to regular graphs (such as bipartite) in Section 3. Extensions to random non-regular graphs with given degree sequence are in the main mentioned at the end of the appropriate subsection. For some properties the methods extend easily in this way (short cycles, connectivity) but for others they seem not to at all. For example, it is known that a large random 3-regular graph is almost surely Hamiltonian, and so is a random 4-regular graph. However, present methods seem not to be strong enough to show that a large random graph with say half of its vertices of degree 3 and half of degree 4 is almost surely Hamiltonian.

In all this, we postpone detailed discussion of a general method called

the small subgraph conditioning method until Section 4, because it requires considerable development and has interesting ramifications for models with non-uniform distribution. It is enough for now to know that this gives information on the distribution of random variables which count certain large subgraphs in the uniform model of random regular graphs. The problem of generating random regular graphs is considered in Section 5. We then visit some other models with non-uniform distributions, and some models which do not produce graphs with given degree sequence but are somehow similar. The final farewell is with comments on unsolved problems.

Almost all of the models discussed here assume that the graphs concerned have labelled vertices. Random unlabelled regular graphs can be accessed via the labelled model by studying the order of the automorphism group of the graphs (see Section 2.7).

For a non-negative integer j and real x, $[x]_j$ denotes $x(x-1)\cdots(x-j+1)$. We use the notation $\mathbf{P}$, $\mathbf{E}$ and $\mathbf{Var}$ for probability, expectation and variance, sometimes subscripted as in $\mathbf{P}_{\mathcal{G}}$ to specify that the probability measure is as in the probability space $\mathcal{G}$. We say that an event H_n occurs a.a.s. (asymptotically almost surely) if $\mathbf{P}H_n \to 1$ as $n \to \infty$, with any obviously necessary parity restriction on n.

2 Uniform model for random regular graphs

We use $\mathcal{G}_{n,d}$ to denote the uniform probability space of d-regular graphs on the n vertices $\{1, 2, \ldots, n\}$ (where dn is even). So sampling from $\mathcal{G}_{n,d}$ is equivalent to taking such a graph uniformly at random (u.a.r.). We use $|\mathcal{G}|$ to denote the number of elements of a uniform space $\mathcal{G}$.

Another probabilistic space can be defined as follows. Suppose that dn is even, and for non-triviality take $d \geq 1$. Consider a set of dn *points* partitioned into n *cells* $v_1, v_2, \ldots, v_n$ of d points each. A perfect matching of the points into $\frac{1}{2}dn$ pairs is called a *pairing*. A pairing P corresponds to a multigraph (with loops permitted) $G(P)$ in which the cells are regarded as vertices and the pairs as edges: a pair (x, y) in P corresponds to an edge (v_i, v_j) of $G(P)$ where $x \in v_i$ and $y \in v_j$. Since each graph (which we assume is simple; i.e. has no loops or multiple edges) corresponds to precisely $(d!)^n$ pairings, a regular graph can be chosen u.a.r. by choosing a pairing u.a.r. and rejecting the result if it has loops or multiple edges. Non-simple graphs are *not* produced uniformly at random since each for each loop the number of corresponding pairings is divided by 2, and for each k-tuple edge it is divided by $k!$. This is the *pairing model* of random regular graphs, given in this form first by Bollobás, but see Section 2.1 for a description of pre-existing models. We denote the (uniform) probability space of pairings by $\mathcal{P}_{n,d}$. We can assume that the points are the elements of $\{1, \ldots, n\} \times \{1, \ldots, d\}$, so that $G(P)$ is induced by a projection.

The pairing model gives a basis for proving properties of graphs in $\mathcal{G}_{n,d}$, by doing computations in $\mathcal{P}_{n,d}$, and conditioning on the event that the cor-

responding multigraph has no loops or multiple edges. We call this event "Simple", and note that it is a function of n and d. Knowledge of the value of $\mathbf{P}_{\mathcal{P}_{n,d}}(\text{Simple})$ permits the following simple exploitation of the connection between $\mathcal{P}_{n,d}$ and $\mathcal{G}_{n,d}$.

Lemma 2.1 *Let H be an event (set of graphs) in $\mathcal{G}_{n,d}$ and H' the set of pairings in $\mathcal{P}_{n,d}$ that correspond to graphs in H. Then*

$$\mathbf{P}_{\mathcal{G}_{n,d}}(H) = \frac{\mathbf{P}_{\mathcal{P}_{n,d}}(H')}{\mathbf{P}(\text{Simple})}.$$

Proof This comes immediately from the uniformity of the two models, and the fact that each graph corresponds to the same number of pairings. ∎

A pairing can be selected u.a.r. in many different ways. In particular, the points in the pairs can be chosen sequentially. At any stage, the first point in the next random pair chosen can be selected using any rule whatsoever, as long as the second point in that pair is chosen u.a.r. from the remaining points. For example, one can insist that the next point chosen is the next one available in any pre-specified ordering of the points, or comes from a cell containing one of the points in the previous pair chosen (if any points such are still unpaired). We use this idea several times in this article and so give it a name: the *independence property* of the pairing model.

In addition, the pairing model provides a simple mechanism for enumerating d-regular graphs asymptotically. Since the number of pairings, or perfect matchings, of t points is

$$f(t) = \frac{t!}{(t/2)!2^{t/2}}, \tag{1}$$

the number of d-regular graphs on n vertices is precisely

$$|\mathcal{G}_{n,d}| = \frac{(dn)!\mathbf{P}(\text{Simple})}{(dn/2)!2^{dn/2}(d!)^n}. \tag{2}$$

Thus, an asymptotic formula for $|\mathcal{G}_{n,d}|$ can be found by estimating $\mathbf{P}(\text{Simple})$.

The pairing model for random graphs with given degree sequence is an immediate extension of the regular case: for a degree sequence $d_1, \ldots, d_n$, the cell v_i contains d_i points, and a perfect matching of all the points is selected u.a.r. Restricting to no loops or multiple edges produces u.a.r. graphs with degree sequence $\mathbf{d} = (d_1, \ldots, d_n)$. This uniform model of graphs we denote by $\mathcal{G}_{n,\mathbf{d}}$. In all asymptotic statements about this model we assume $\mathbf{d}$ is restricted to sequences with even sum.

We next discuss models for related combinatorial configurations, and then outline results on properties of $\mathcal{G}_{n,d}$.

2.1 History of uniform models

The pairing model was first given in its simple explicit form by Bollobás (beginning with [9]) and called the *configuration model*. (We prefer not to refer to pairings as configurations since the latter has a broad meaning which is useful in many contexts, for example in Section 4.2.) A little earlier, Bender and Canfield [5] used a model in their enumeration of graphs with given degrees which is implicitly a generalisation of the pairing model (see Section 4 of that paper). The difference is that they use involutions in place of perfect matchings of points, because they study 0-1 matrices with given row sums, which are the adjacency matrices of graphs with given degrees, and as a result the entries of the matrix on the main diagonal must be treated differently. (An entry 1 in the main diagonal does not correspond to a loop in the pairing model, since a loop uses up *two* points.) In addition, their analysis was made more general by allowing the possibility to specify a limited number of edges to be forbidden.

Even earlier, Békéssy et al. [4] studied random 0-1 matrices with given row and column sums by using essentially a more general model, in which the points in the cells are arranged into classes of arbitrary specified sizes and permuted at random. The argument in [4] gave asymptotically the number of such configurations in which no two points in the same cell are of the same class. In the general case, this gives a model of random bicoloured graphs with given vertex degrees. On the other hand, in the special case that every colour class has cardinality 2, this is very similar to the pairing model, the only difference being that the edges of the graph are labelled. The results in [4] give the number of pairings which induce no loops in the multigraph. Independently of [5], the author's PhD thesis [115] studied random regular graphs from the point of view of unlabelling the edges in the model of multigraphs arising from [4], which then becomes the pairing model. However, this model was studied only indirectly in [115] because enumeration results on multigraphs with loops were sought. Once achieved, these were built on to find the numbers of simple graphs (with given degree sequence), as well as properties of such random graphs. The methods used for all these variations of the model are equivalent to those required to study the pairing model.

2.2 Related enumeration results

Of course results on probabilities in a uniform model are inextricably associated with enumeration results. Most computations in this paper aim for asymptotic results: unless otherwise specified, all limits refer to $n \to \infty$ with n restricted to even integers if d is odd.

Of major interest is an asymptotic formula for the number $|\mathcal{G}_{n,d}|$ of labelled d-regular graphs, which by (2) amounts to estimating $\mathbf{P}(\text{Simple})$. Read [91] gave an exact formula for $|\mathcal{G}_{n,d}|$ which is unfortunately too complicated to be easily amenable to asymptotic evaluation for general d. For $d = 3$ however, his analysis leads to the asymptotic expression $\frac{(6n)!}{288^n (3n)! e^2}$, which is equivalent

to (3) for $d = 3$. In addition, any number of terms in an asymptotic series expansion for the number of cubic graphs can be obtained from his result. However, the inelegance of his derivation dissuaded him from applying it to larger d.

Bender and Canfield's asymptotic formula for $|\mathcal{G}_{n,d}|$ can be stated as

$$\mathbf{P}(\text{Simple}) \sim \exp\left(\frac{1-d^2}{4}\right) \qquad \text{for fixed } d, \tag{3}$$

where here and elsewhere in this paper $a(n) \sim b(n)$ means $a(n) = (1+o(1))b(n)$ as $n \to \infty$ (with dn even, of course). Thus evaluating (2) using Stirling's formula gives the following.

Theorem 2.2 (Bender and Canfield [5]) *For fixed d*

$$|\mathcal{G}_{n,d}| \sim \sqrt{2}e^{(1-d^2)/4}\left(\frac{d^d n^d}{e^d (d!)^2}\right)^{\frac{1}{2}n}. \tag{4}$$

This was found independently in [115], and then Bollobás [9, 10] gave the pairing model proof and showed that the formula applied for $d = d(n) \leq \sqrt{2\log n} - 1$. A version of this proof of (3) is in the next section.

From (3) it follows that $\mathbf{P}(\text{Simple})$ is bounded below for fixed d, a fact which combined with Lemma 2.1 immediately yields the following. First, for an event H in $\mathcal{P}_{n,d}$, define $G(H)$ to be the event in $\mathcal{G}_{n,d}$ containing precisely all simple graphs of the form $G(P)$ for some $P \in H$.

Corollary 2.3 *Let $d \geq 1$ be fixed, and let H be an event which is a.a.s. true in $\mathcal{P}_{n,d}$. Then $G(H)$ is a.a.s. true in $\mathcal{G}_{n,d}$.*

McKay [72] used switchings (as described in Section 2.4) to extend the range of d in (3) to $d = o(n^{1/3})$. McKay and Wormald [81] then used a new sort of switching to find the formula for $d = o(\sqrt{n})$:

$$\mathbf{P}(\text{Simple}) = \exp\left(\frac{1-d^2}{4} - \frac{d^3}{12n} + O\left(\frac{d^2}{n}\right)\right). \tag{5}$$

Corollary 2.4 (McKay and Wormald [81]) *For $d = o(\sqrt{n})$ the number of d-regular graphs on n vertices is*

$$\frac{(dn)!}{(\frac{1}{2}dn)!2^{dn/2}(d!)^n}\exp\left(\frac{1-d^2}{4} - \frac{d^3}{12n} + O\left(\frac{d^2}{n}\right)\right).$$

McKay and Wormald also obtained a formula for $d \approx cn$ [80] which may have consequences for quite dense random regular graphs:

$$|\mathcal{G}_{n,d}| \sim \sqrt{2}\left(2\pi n\lambda^{d+1}(1-\lambda)^{n-d}\right)^{-n/2}\exp\left(\frac{-1+10\lambda-10\lambda^2}{12\lambda(1-\lambda)}\right)$$

where $\lambda = d/(n-1)$, provided $d = d(n)$ is an integer-valued function such that, for n sufficiently large, dn is even and $\min\{d, n-d-1\} > cn/\log n$ for some $c > \frac{2}{3}$.

The formulae mentioned in this section have their counterparts for graphs with given degree sequences, the most widely applicable being in [81] and [80]. These formulae allowed McKay and Wormald [82] to create a useful model of the degree sequence of a random graph in $\mathcal{G}(n,p)$. Since it has proved so useful, the Bender–Canfield formula for bounded degrees d_i is included here:

$$|\mathcal{G}_{n,\mathbf{d}}| \sim \frac{(2m)!e^{-\lambda-\lambda^2}}{m!2^m \prod_{i=1}^n d_i!} \tag{6}$$

where $2m = \sum d_i$ and $\lambda = \frac{1}{2m}\sum \binom{d_i}{2}$.

2.3 Short cycle distribution

The number of short cycles in random regular graphs of small degree has an asymptotically Poisson distribution, as is the usual rule with the sum of many nearly independent rare events. (This effect is called the "Poisson paradigm" by Alon and Spencer [3].) The usual method of proving this uses an asymptotic version of the fact that a Poisson variable is determined by its moments. This method is presented here, saving a much more powerful (in the case of regular graphs) switching method for Section 2.4.

We say that a set of variables $X_i = X_i^{(n)}$, for i in some finite set I, defined on a sequence of probability spaces indexed by n, are *asymptotically independent Poisson* with means λ_i if their joint distribution tends to that of independent Poisson variables whose means are fixed numbers λ_i. To be precise,

$$\lim_{n\to\infty} \mathbf{P}\left(\bigwedge_{i\in I}\{X_i = r_i\}\right) \to \prod_{i\in I} e^{-\lambda_i}\frac{\lambda_i^{r_i}}{r_i!} \tag{7}$$

for every fixed set of non-negative integers r_i, $i \in I$. (Recall that n is restricted to even numbers in all limits if d is odd.) Throughout this section, we use $Z(\lambda)$ to denote a Poisson random variable with expectation λ.

Theorem 2.5 (Bollobás [10], Wormald [115, 118]) *For d fixed, let $X_i = X_{i,n}$ ($i \geq 3$) be the number of cycles of length i in a graph in $\mathcal{G}_{n,d}$. For fixed $k \geq 3$, $X_3, \ldots, X_k$ are asymptotically independent Poisson random variables with means $\lambda_i = \frac{(d-1)^i}{2i}$.*

This was derived in [10] using the pairing model, and independently in [115, 118] essentially from (6).

Theorem 2.5 plays a major role in the results on Hamiltonicity and related properties (see Section 4.2). As mentioned above, we also need to calculate limiting probabilities in the pairing model conditional on the event Simple. For

this, a crucial step is the determination of $\mathbf{P}(\text{Simple})$ asymptotically. Since loops and multiple edges are cycles of length 1 and 2, the short cycle distribution in the multigraphs corresponding to pairings provides a convenient common generalisation of $\mathbf{P}(\text{Simple})$ and the distribution in $\mathcal{G}_{n,d}$. This leads to the following result, which was first derived in full by Bollobás.

Theorem 2.6 (Bollobás [10]) *For d fixed, let $X_i = X_{i,n}$ ($i \geq 1$) be the number of cycles of length i in the random multigraph coming from a pairing in $\mathcal{P}_{n,d}$. For $k \geq 1$, $X_1, \ldots, X_k$ are asymptotically independent Poisson random variables with means $\lambda_i = \frac{(d-1)^i}{2i}$.*

A Corollary of this is (3), obtained by considering the event $X_1 = X_2 = 0$.

Before proceeding further, the reader deserves an explanation as to why small connected subgraphs other than cycles are not examined here. It is because unless they are trees or unicyclic, they do not exist (a.a.s., unless d is permitted to grow as a function of n, in which case the analysis gets more difficult; see Section 2.4). The following result has been explicitly or implicitly given since the first examinations of random regular graphs, and is easily proved by estimating the expected number of subgraphs of the given type in $\mathcal{P}_{n,d}$.

Lemma 2.7 *For fixed d and any fixed graph F with more edges than vertices, $G \in \mathcal{G}_{n,d}$ a.a.s. contains no subgraph isomorphic to F.* ∎

Thus, the neighbourhood of a random vertex in a huge random d-regular graph looks just like part of an infinite tree. It follows from the lemma that a.a.s. no two cycles of bounded length are joined by a path of bounded length.

Proof of Theorems 2.5 and 2.6 The plan is quite simple, following the method of moments. In this, the joint factorial moments of $X_{1,n}, \ldots, X_{k,n}$ are shown to tend in the limit to those of the independent Poisson variables $Z(\lambda_1), \ldots, Z(\lambda_k)$. That is, for every sequence of fixed non-negative integers $r_1, \ldots, r_k$,

$$\lim_{n\to\infty} \mathbf{E}\left(\prod_{i=1}^{k} [X_{i,n}]_{r_i}\right) \to \prod_{i=1}^{k} \lambda_i^{r_i}. \tag{8}$$

Theorem 2.6 follows from this by a well-known result (see Lemma 2.8). Theorem 2.5 comes from Theorem 2.6 by Lemma 2.1.

To compute the left hand side of (8), for each $1 \leq i \leq k$ distinguish the pairs of points corresponding to an ordered set of r_i cycles of length i. The required value is the number of pairings with such distinguished pairs, divided by the total number $f(dn)$ of pairings. Letting $s = r_1 + 2r_2 + \cdots + kr_k$, the number of ways to choose the s distinguished pairs in the correct configuration for the desired cycles is asymptotic to

$$\frac{(d(d-1)n)^s}{\prod_{i=1}^{k} (2i)^{r_i} r_i!}$$

if exactly s different vertices are involved in the pairs, and is $O(n^{s-1})$ otherwise (since the vertices can be chosen in this many ways, and given the vertices, the distinguished pairs can form only a finite number of configurations). The cycles of each length can be ordered in $\prod_{i=1}^{k} r_i!$ ways, and the pairing can then be completed in

$$f(dn-2s) \sim f(dn)/(dn)^s$$

ways, by (1). Multiplying these together gives the right side of (8), as required. ∎

That proof required the following multivariate generalisation of Brun's sieve, which is an implication of the Bonferroni inequalities (see [3] for example). This generalisation is well known in probability theory (see Chung [27]) and is given without explicit proof in [16]. We show here how it follows from Brun's sieve in a quite elementary way. This proof mimics the structure of the proof of Theorem 2.5 given in [118]. It is included especially because it uses alterations of probability measure, which will appear again in Section 4.1. Without loss of generality we could assume that the $X_{i,n}$ are indicator variables for random events.

Lemma 2.8 *Let $\lambda_1, \dots, \lambda_k$ be some set of fixed non-negative reals, and let $X_{1,n}, \dots, X_{k,n}$ be non-negative integer random variables defined on the same space $\mathcal{G}_n$ for each n. If (8) holds for each fixed set of non-negative integers $r_1, \dots, r_k$, then the variables $X_{1,n}, \dots, X_{k,n}$ are asymptotically independent Poisson with means λ_i.*

Proof This is by induction on k; for $k = 1$ it is Brun's sieve, so assume $k > 1$.

We take two cases: firstly $\lambda_k = 0$. Then (8) with $r_1 = \cdots = r_{k-1} = 0$ and $r_k = 1$ gives $\mathbf{E}X_k \to 0$, and so (7) holds whenever $r_k \neq 0$. On the other hand, it holds for $r_k = 0$ by the inductive hypotheses.

So now assume $\lambda_k > 0$. Let $\mathcal{G}'_n$ denote the probability space obtained from $\mathcal{G}_n$ by weighting the probability measure according to $[X_{k,n}]_{r_k}$. Let $\mathbf{P}'$ and $\mathbf{E}'$ denote probability and expectation in $\mathcal{G}'_n$, whereas $\mathbf{P}$ and $\mathbf{E}$ refer to $\mathcal{G}_n$. Then for any event R and variable X in $\mathcal{G}_n$,

$$\mathbf{P}'(R) = \frac{\mathbf{E}([X_{k,n}]_{r_k} \mid R)\mathbf{P}(R)}{\mathbf{E}[X_{k,n}]_{r_k}}, \qquad \mathbf{E}'(X) = \frac{\mathbf{E}(X[X_{k,n}]_{r_k})}{\mathbf{E}[X_{k,n}]_{r_k}}. \tag{9}$$

Thus from (8)

$$\mathbf{E}' \prod_{i=1}^{k-1} [X_{i,n}]_{r_i} \to \frac{\prod_{i=1}^{k} \lambda_i^{r_i}}{\lambda_k^{r_k}} = \prod_{i=1}^{k-1} \lambda_i^{r_i}.$$

So by the inductive hypothesis, $X_1, \dots, X_{k-1}$ are asymptotically Poisson in $\mathcal{G}'_n$ with means λ_i. The same is true in $\mathcal{G}_n$. So if R is the event that $X_1 = r_1, \dots, X_{k-1} = r_{k-1}$, we have $\mathbf{P}'(R) \to \Pi = \prod_{i=1}^{k-1} \mathbf{P}(Z(\lambda_i) = r_i)$, and also $\mathbf{P}(R) \to \Pi$. Hence from the left equation in (9),

$$\mathbf{E}([X_{k,n}]_{r_k} \mid R) - \mathbf{E}[X_{k,n}]_{r_k} \to 0.$$

But from (8) $\mathbf{E}[X_{k,n}]_{r_k} \to \lambda_k^{r_k}$, and so

$$\mathbf{E}([X_{k,n}]_{r_k} \mid R) \to \lambda_k^{r_k}.$$

That is, in the probability space obtained from $\mathcal{G}_n$ by conditioning on the event $R = R_n$, the factorial moments of $X_{k,n}$ tend to those of $Z(\lambda_k)$. Hence another application of Brun's sieve implies that $X_{k,n}$ is asymptotically Poisson with mean λ_k in this conditional space. Now

$$\begin{aligned}\mathbf{P}(X_1 = r_1, \ldots, X_k = r_k) &= \mathbf{P}(X_k = r_k \mid R)\mathbf{P}(R) \\ &\to \mathbf{P}(Z(\lambda_k) = r_k) \prod_{i=1}^{k-1} \mathbf{P}(Z(\lambda_i) = r_i)\end{aligned}$$

as required. ∎

Results on the short cycle distribution for $G \in \mathcal{G}_{n,\mathbf{d}}$ which generalise those given above were obtained in [10] and [118].

2.4 The switching method: subgraphs, eigenvalues and spanning trees

The switching method enables us to prove results about subgraphs of $G \in \mathcal{G}_{n,d}$ when d, or the size of the subgraph, grows much more quickly than is permitted in the proofs of Theorems 2.5 and 2.6 given above. It often applies for $d = o(n^c)$ where c is some "reasonable" number like $\frac{1}{2}$.

We define a *simple switching* in a pairing P to be the replacement of two pairs $\{p_1, p_2\}, \{p_3, p_4\}$ by $\{p_1, p_3\}, \{p_2, p_4\}$ or $\{p_1, p_4\}, \{p_3, p_2\}$. This induces a switching of two edges of the corresponding graph, and early switching results analysed this operation directly on graphs without reference to pairings. McKay [69] introduced switchings to the random regular graph scene to obtain general upper and lower bounds on the probability of a subgraph occurring. In [68] he used the method to obtain bounds on the probabilities of cycles of various (unbounded) lengths occurring in $G \in \mathcal{G}_{n,d}$. It follows that such G a.a.s. satisfies a certain condition which, as McKay showed in [67], implies that the distribution of eigenvalues of G tends towards a fixed function, which he determined. The argument exploits a connection between the eigenvalues and the cycle distribution via the number of closed walks of length k (noting that this is the trace of the kth power of the adjacency matrix, which is of course the sum of the kth powers of the eigenvalues). The same connection was used by Broder and Shamir [24] to obtain an asymptotic almost sure upper bound on the second-largest eigenvalue in absolute value, for d even. This bound was decreased by Friedman [39] to $2\sqrt{2d-1} + 2\log d + c'$ holding with probability $1 - O(n^{-c})$, where c' depends on c.

In [70] McKay used another eigenvalue connection, the matrix tree theorem, in a similar way to show that if a sequence of regular graphs G_i on $n_i \to \infty$

vertices satisfies a certain asymptotic condition on numbers of cycles then the number of spanning trees in G_i is $(c_d + o(1))^{n_i}$ where

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

The required condition is shown, using the switching results, to be true a.a.s. for $G \in \mathcal{G}_{n,d}$, implying that the nth root of the number of spanning trees in G is a.a.s. asymptotic to c_d (by which we mean it is a.a.s. in the interval $(c_d - \epsilon, c_d + \epsilon)$ for all $\epsilon > 0$). McKay also proved a variety of stronger versions and other variations of these statements.

The main power of switchings is the ability to obtain accurate estimates of very small probabilities. As an example, we illustrate with a use of the more modern switchings introduced in [79] and [81] which turn out to give results more easily and stronger than using simple switchings as in McKay [71, 72]. The following result is part of the proof of (5) (giving an asymptotic formula for $|\mathcal{G}_{n,d}|$) for $d = o(n^{1/3})$ in [79].

Lemma 2.9 (McKay and Wormald [79]) *Let $S_{a,b}$ denote the set of pairings $P \in \mathcal{P}_{n,d}$ such that $G(P)$ has precisely a loops, b double edges, and no edges of multiplicity greater than 2, nor double loops. Then for $1 \le a < d + \omega(n)$ and $b < d^2 + \omega(n)$, where $\omega(n) \to \infty$ arbitrarily slowly, $|S_{a,b}|/|S_{a-1,b}| = \frac{d-1}{2a}(1 + O(\frac{d+\omega(n)}{n}))$.*

Proof Denote a pair of points $\{p_i, p_j\}$ by p_ip_j. Given any pairing in $S_{a,b}$, choose a pair p_1p_2 which projects onto a loop, choose two other pairs p_3p_4 and p_5p_6, and replace all three pairs by the pairs p_1p_3, p_2p_5 and p_4p_6 to produce a pairing P'. This switching operation can be written as the composition of two simple switchings.

We count how many ways this can be done so that P' lands in $S_{a-1,b}$. There are $2a$ ways to choose p_1 and p_2 (as an ordered pair), and roughly dn ways to choose each of p_3 and p_5, which determine p_4 and p_6. After this there is only one way to perform the switching described. However, some choices of p_3 and p_5 lead to unwanted multiple edges created by the switching ($O(d^3n)$ choices do this) or to triple edges ($O(bdn)$ do this) or unwanted loops ($O(adn)$ do this) or destroy loops or double edges unintentionally ($O(adn + bdn)$ do this). So the number of P' in $S_{a-1,b}$ corresponding to P is

$$2ad^2n^2\left(1 - O\left(\frac{d^2+a+b}{dn}\right)\right). \tag{10}$$

On the other hand, in how many ways can $P' \in S_{a-1,b}$ be produced? To reverse the procedure, choose p_1 and p_2 in the same cell (in $d(d-1)n$ ways), and choose p_4 to be any other point (in $dn-2$ ways). Then p_3, p_5 and p_6 are determined, in that order, as the points paired in P' with the three chosen points. Call this

choice "bad" if it does not determine the reverse of the procedure as described. There are only a few causes of badness, but counting as before shows that the number of bad choices cannot exceed $O(d^3an+d^2bn+d^4n)$. Hence the number of P in $S_{a,b}$ corresponding to P' is

$$d^2(d-1)n^2\left(1-O\left(\frac{d^2+da+b}{dn}\right)\right),$$

and the result follows on division of (10) by this expression. ∎

This lemma can be used to obtain the asymptotic probability that $a=0$, conditional upon a given value of b, by estimating $|S_{a,b}|/|S_{0,b}|$ (in the form of a telescoping product of the ratios in the lemma) and summing over all a which contribute significantly. (Large values of a can be ruled out by a simple expectation argument.) This gives the reciprocal of the desired probability for $d=o(n^{1/3})$. For d in this range, the restrictions in the hypotheses of the lemma are easily justified in almost all pairings by expectation arguments. To complete an estimation of $\mathbf{P}$(Simple), we can use other types of switchings for estimation of $|S_{0,b}|/|S_{0,0}|$, and combining the two, we can estimate $|S_{a,b}|/|S_{0,0}|$ for all the significant terms. Summing over a and b gives an estimate of $|\mathcal{P}_{n,d}|/|S_{0,0}|$, which gives the estimate of $1/\mathbf{P}$(Simple).

The superiority of the switching used in this proof over a simple switching (switch a loop-inducing pair with any other pair) is that the number of ways to reverse the procedure is almost independent of P. For the single switching, the number of valid ways to reverse the procedure, given P', depends heavily on the number of triangles in P'. McKay [72] argued about the average number of reversals, in order to reach $d=o(n^{1/3})$; similar averaging arguments in [81] for the "modern" switchings gave (5) (reaching $d=o(\sqrt{n})$). Possibly this averaging can be pushed further, but at the expense of a lot of work, and the next step would only be $n^{3/5}$.

Further uses of switchings are mentioned in Sections 2.5, 2.6, 2.7, 2.8, 3.2 and 5. Switchings also work fine on graphs in the model $\mathcal{G}_{n,\mathbf{d}}$. In addition, an argument using one simple switching implies that many variables defined on random regular graphs are sharply concentrated (Theorem 2.19). A general description of the switching argument and an application to $\mathcal{G}(n,p)$ can also be found in [123]. There the switchings give a relationship between probabilities of events rather than cardinalities of sets, but they still taste of counting.

2.5 Longer cycles

Theorem 2.5 implies an asymptotic formula for the probability that $G\in\mathcal{G}_{n,d}$ has girth k when k and d are fixed (see Theorem 2.12). By the switching method we can show that this asymptotic formula also applies when $k\to\infty$, simultaneously with $d\to\infty$. This is done for k close to $\frac{1}{2}\log_d n$ by McKay et al. [83]. The results on distribution of short cycles can also be extended in this way to cycles of unbounded length.

Garmo [50] used the small subgraph conditioning method to examine long cycles. He showed that provided $k = k(n) \to \infty$ faster than $c \log n$ but $k = o(n)$, the number of k-cycles in $G \in \mathcal{G}_{n,d}$ is concentrated near its expectation, and he found its limiting distribution on the appropriate scale. If $k/n \to q$, $0 < q < 1$, the number is not so concentrated, the variance is a non-zero constant times the square of the expectation, and the limiting distribution was found.

2.6 Connectivity and diameter

For fixed $d \geq 3$, $G \in \mathcal{G}_{n,d}$ is d-connected a.a.s., as was shown independently by Bollobás [11] and by Wormald [117] (originally in [115]). This result can readily be extended using switchings (Section 2.4) to d growing rather modestly with n. The following extension seems to be the strongest in this sense which has appeared in print, but the upper bound was just what was needed at the time and by no means represents the limit imposed by the method.

Theorem 2.10 (Łuczak [63]) *For $3 \leq d(n) \leq n^{0.02}$, $G \in \mathcal{G}_{n,d}$ is d-connected a.a.s.*

This is proved below in the more restricted range $3 \leq d(n) \leq 3 \log n$, which is about as far as one can easily go without switchings or something similar. The proof below actually gives the bound $O(n^{2-d})$ on the probability that $G \in \mathcal{G}_{n,d}$ is not d-connected for d fixed. It is easy to see that this bound cannot be improved for fixed d (apart from the value of the constant implicit in the O). In [117] and [63], non-regular versions of this theorem were given. (See Theorem 2.16 below.)

Curiously, larger values of d seem to be more difficult to handle, but we can confidently conjecture that the same basic result holds even for d of the order of n.

Conjecture 2.11 *For $3 \leq d = d(n) < n - 3$, $G \in \mathcal{G}_{n,d}$ is d-connected a.a.s.*

This conjecture is trivially true also for $d = n - 1$ and $n - 2$. The exclusion of $n - 3$ is necessary, since the complement of G is then a 2-regular graph which has non-zero probability of having a 4-cycle. Deleting the other $n - 4$ vertices from G gives a disconnected graph. To prove the conjecture for d bigger than $n^{0.02}$ would not be too hard using switchings; something like $\sqrt{n}$ should be possible. However, to reach past $d = n^{1-\epsilon}$ for arbitrarily small ϵ would probably require a really new idea. On the other hand, for d very close to n, it would become easy to verify by analysing the complement.

For $d = 2$ the story is different: a random graph in $\mathcal{G}_{n,2}$ is a.a.s. disconnected. In fact, it is connected if and only if it is Hamiltonian, the probability of which is easy to compute asymptotically. The number of Hamiltonian graphs is $\frac{n!}{2n}$, and dividing by $|\mathcal{G}_{n,2}|$ (say from Theorem 2.2) gives

$$\mathbf{P}(G \in \mathcal{G}_{n,2} \text{ is Hamiltonian}) \sim \tfrac{1}{2} e^{3/4} \sqrt{\pi} n^{-1/2} \approx 1.876 n^{-1/2}. \qquad (11)$$

The cyclic connectivity of regular graphs, especially 3-regular, has also been of considerable interest. A set S of vertices or edges in a connected graph G is *cycle-separating* if at least two components of $G - S$ contain cycles. G is *cyclically k-vertex-connected* (*cyclically k-edge-connected*) if it has no cycle-separating vertex (respectively, edge) set of cardinality less than k. For $d \geq 3$, a large d-regular graph with a j-cycle immediately has cycle-separating vertex and edge sets of cardinality $(d-2)j$. By Theorem 2.5, the probability that the shortest cycle has length j tends to a non-zero constant. This shortest cycle length, or girth, is asymptotically the determining factor for cyclic connectivity, as shown in [117].

Theorem 2.12 *Let $d \geq 3$ and $k \geq 3$ be fixed. The probabilities that $G \in \mathcal{G}_{n,d}$ is*

(i) cyclically k-vertex-connected

(ii) cyclically k-edge-connected

(iii) of girth at least $k/(d-2)$

are all asymptotic to

$$\prod_{3 \leq i < k/(d-2)} \exp\left(\frac{-(d-1)^i}{2i}\right)$$

as $n \to \infty$.

Theorem 2.5 determines this limiting probability from (iii).

Theorems 2.10 and 2.12 are both proved by showing that deleting a small subset of the vertices of $G \in \mathcal{G}_{n,d}$ a.a.s. does not produce two reasonably large components. This uses a simple expectation argument (and, essentially and surreptitiously, Markov's inequality, which can be found in many probability texts such as Grimmett and Stirzaker [53]). It can employ either the formula for the number of graphs with given degree sequence (as in [117]), or calculations directly in the pairing model and then making use of Lemma 2.1 (as by Bollobás [11] and [16, Thm VII.32], although the proof in the latter is over-simplified: the analysis is false for large values of the variable a). Use of the pairing model approach is demonstrated in the proofs given below.

A notion closely related to the usual proofs of the connectivity results (for example Lemma 2.14 below) is the *isoperimetric number* of a graph G, which is the minimum value of the ratio $|\partial U|/|U|$ over all $U \subseteq V(G)$, $|U| \leq |V(G)|/2$. Here ∂U is the set of edges with exactly one endpoint in U. Bollobás [17] gave an asymptotically almost sure lower bound on this number for $G \in \mathcal{G}_{n,d}$. It follows that the number is a.a.s. between $\frac{1}{2}d - \epsilon(d)$ and $\frac{1}{2}d + \epsilon(d)$ where $\epsilon(d) \to 0$ as $d \to \infty$.

The *diameter*, $\mathrm{diam}(G)$, of a graph G is the maximum, over all $u, v \in V(G)$, of the distance from u to v. For a d-regular graph G this clearly satisfies

$|V(G)| \le 1 + \sum_{i=0}^{\mathrm{diam}(G)-1} d(d-1)^i$. Graphs realising this bound are called Moore graphs and are extremely rare. Roughly speaking, given n and d the inequality gives a lower bound on the diameter of roughly $\log_{d-1} n$. Using methods somewhat related to those used for connectivity, Bollobás and de la Vega [21] showed the following.

Theorem 2.13 *The diameter diam(G) of $G \in \mathcal{G}_{n,d}$ a.a.s. satisfies* $1+\lfloor \log_{d-1} n \rfloor + \left\lfloor \log_{d-1}\left(\frac{(d-2)}{6d}\log n\right)\right\rfloor \le \mathrm{diam}(G) \le 1 + \left\lceil \log_{d-1}((2+\epsilon)dn\log n)\right\rceil$.

Another result involving distance was obtained by Bollobás [13], on the probability that for $G \in \mathcal{G}_{n,d}$, each vertex v is determined uniquely by the sequence $n_1(v), \dots, n_k(v)$ where $n_i(v)$ is the number of vertices of distance i from v, and k is a function of n. This event implies that the automorphism group of G is trivial.

We turn now to the proofs of Theorems 2.10 and 2.12. These are included not only because the latter has not been proved directly using the pairing model before, but also as an example of using the independence property of the pairing model. This is in the following lemma. First a technical definition: let us define an (s,j)-*separating set* of a graph G to be a set $S \subseteq V(G)$ with $|S| = s$ such that $G - S$ has a component with exactly j vertices.

Lemma 2.14 *Let d and s be arbitrary. For $P \in \mathcal{P}_{n,d}$,*

$$\mathbf{P}(G(P) \text{ has an } (s,j)\text{-separating set}) < 3^{2+s/d}\left(\frac{j+s}{n}\right)^{j(\frac{1}{2}d-1)} n^{\frac{1}{2}s}(j+s)^{\frac{3}{2}s}$$

for all $j < \frac{2}{3}n - s$.

Proof Let X denote the expected number of (s,j)-separating sets. To show $\mathbf{P}(X \ge 1) \to 0$ it is enough, by Markov's inequality, to show $\mathbf{E}X \to 0$. If such a set S exists, let F be a component of $G - S$ with $|V(F)| = j$ and let $T = S \cap N(F)$. Put $t = |T|$. The j cells of P corresponding to $V(F)$ and the t cells for T can be chosen in at most $\binom{n}{j} n^t$ ways.

Since each vertex in T has degree at least 1 in the subgraph of G induced by $F \cup T$, the number r of edges with at least one end in F satisfies $r \ge (jd+t)/2$. Given r and the cells in $V(F) \cup T$, what is the probability that a random pairing has r pairs of the required type? By the independence property of the pairing model, such pairs can be selected by repeatedly selecting the pair containing the next unpaired point in a cell in $V(F)$. The probability of succeeding in the ith step of this random process, conditional upon succeeding in all earlier steps, is $\frac{(j+t)d-2i+1}{nd-2i+1}$. So the probability of selecting r such pairs is at most

$$\prod_{i=1}^{r} \frac{j+t-(2i-1)/d}{n-(2i-1)/d} \le \left(\frac{[j+t]_{2r/d}}{[n]_{2r/d}}\right)^{\frac{1}{2}d} \tag{12}$$

where we define $[j]_{i+\epsilon}$ for $i \in \mathbf{Z}$ and $0 < \epsilon < 1$ to be $[j]_i(j-i)^\epsilon$. Thus

$$\mathbf{E}X \le \sum_{t=0}^{s} \sum_{r} \left(\frac{[j+t]_{2r/d}}{[n]_{2r/d}} \right)^{\frac{1}{2}d} \binom{n}{j} n^t < 3 \sum_{t=0}^{s} \left(\frac{[j+s]_{j+t/d}}{[n]_{j+t/d}} \right)^{\frac{1}{2}d} \frac{[n]_j}{j!} n^t \quad (13)$$

since $j + t \le j + s < \frac{2}{3}n$ and $2r/d \ge j + t/d$. The sum over $t \le s$ can be conveniently bounded by 3 times the term with $t = s$. We have

$$\frac{[j+s]_{j+s/d}}{[n]_{j+s/d}} < \min\left(\frac{(j+s)^s j!}{[n]_j (n-j-s/d)^{s/d}}, \frac{(j+s)^{j+s/d}}{n^{j+s/d}} \right).$$

Apply the first term in the minimum to one factor in (13) along with $n - j - \frac{s}{d} > \frac{1}{3}n$, and use the second term on the other factors, to obtain the lemma. ∎

In the case of graphs, which by definition have no loops or multiple edges, the result of Lemma 2.14 can be improved for small j as follows.

Lemma 2.15 *Let $d = d(n) < n^{1/3}$. For $P \in \mathcal{P}_{n,d}$,*

$$\mathbf{P}(G(P) \in \mathcal{G}_{n,d} \text{ and has an } (s,j)\text{-separating set}) < 2 \left(\frac{j+s}{n} \right)^{j(d-\frac{1}{2}j+\frac{1}{2})} n^{j+s}.$$

Proof Firstly, we can assume $j < 2d$ and $s < d^2$ since otherwise there is nothing to prove. If G is a graph, the lack of loops and multiple edges permits replacement of the inequality $r \ge (jd+t)/2$ in the proof of Lemma 2.14 by $r \ge jd - \binom{j}{2}$. Repeating the first part of that argument, and bounding (12) by $(j+t)^{\frac{1}{2}d}/n^{\frac{1}{2}d}$, shows that the worst case is $t = s$ and $r = (jd+t)/2$. The sum over t and r can be bounded by say twice this largest term, which gives the lemma. ∎

Proof of Theorem 2.10 For $n \ge d+1$, a graph with n vertices which is not d-connected contains a $(d-1, j)$-separating set for some $1 \le j \le \frac{1}{2}n$. For d-regular graphs, $j = 1$ can be excluded immediately. For $2 \le j \le d$ the bound in Lemma 2.15 is $O(n^{2-d}(d+1)^{2d-1})$, with the worst case always occurring at $j = 2$, as well as $j = 3$ in the case $d = 3$, the other cases contributing negligibly. For $j > d$ we use the bound in Lemma 2.14, which is at most

$$\frac{27(j+d)^{j(\frac{1}{2}d-1)+\frac{3}{2}d}}{n^{j(\frac{1}{2}d-1)-\frac{1}{2}(d-1)}}. \quad (14)$$

This is $O(n^{2-d})$ for $j = d+1$ and $j = d+2$ (worst case $d = 3$), $O(n^{1-d})$ for $j = d+3$, and much smaller for $j = \frac{1}{2}n$. For the values in between, the second derivative with respect to j of the logarithm of (14) has the same sign as $j(d-2) + d(2d-7)$, which is positive for $j \ge 3$. Hence it is a convex function of j, and the sum of (14) over $d+3 \le j \le \frac{1}{2}n$ is $O(n^{2-d})$. So for $P \in \mathcal{P}_{n,d}$, the probability that $G(P)$ is a graph which is not d-connected is $O(n^{2-d}(d+1)^{2d-1})$.

The theorem now follows by Lemma 2.1 and the estimate for $\mathbf{P}$(Simple) in (5). ∎

Proof of Theorem 2.12 A cyclically k-edge-connected d-regular graph with sufficiently many vertices must have girth at least $k/(d-2)$. On the other hand, let G be d-regular, of girth at least $k/(d-2)$, and with a cycle-separating vertex set S such that $|S| = s < k$. Then $G - S$ has two components with cycles and hence one of them has j vertices for $k/(d-2) \le j \le \frac{1}{2}n$. Since s and d are bounded, the upper bound in Lemma 2.14 is

$$\frac{O(1)(j+s)^{j(\frac{1}{2}d-1)+\frac{3}{2}s}}{n^{\frac{1}{2}j(d-2)-\frac{1}{2}s}}.$$

Since $j(d-2) \ge s+1$, this is $O(n^{-\frac{1}{2}})$ for j near its lower bound, and analysis as for (14) shows the sum over all relevant j is $O(n^{-\frac{1}{2}})$. Hence for $P \in \mathcal{P}_{n,d}$, $G(P)$ a.a.s. either has girth less than $k/(d-2)$ or is cyclically k-vertex-connected and hence (for n large enough) also cyclically k-edge-connected. It follows by Lemma 2.1, (3) and Theorem 2.5 that the probability that $G \in \mathcal{G}_{n,d}$ is cyclically k-vertex-connected, cyclically k-edge-connected, or of girth at least $k/(d-2)$, all tend to the same non-zero constant as $n \to \infty$. ∎

Note The proof of Theorem 2.12 is easily adapted to permit d and k to grow slowly with n, using the result mentioned in Section 2.5 to estimate the probability of having large girth. However, even for d fixed, it is only easy to reach $k = o(\log\log n)$. To go past this point would require more careful estimation of the probability of having a cycle-separating set in the conditional space of large girth, which would appear to be even more difficult than extending the range of d in Theorem 2.10. Again, the switching method would presumably be required here.

We close this subsection with mention of a similar result for $\mathcal{G}_{n,\mathbf{d}}$.

Theorem 2.16 ([117]) *Fix $3 \le d \le D$, and for each n let $\mathbf{d} = \mathbf{d}(n) = (d_1, \dots, d_n) \in [d, \dots, D]^n$ with $\sum d_i$ even. The probability that $G \in \mathcal{G}_{n,\mathbf{d}}$ is not d-connected is $O(n^{2-d})$.*

Łuczak's theorem in [63] shows that this is also valid with the upper bound $d \le n^{0.02}$, but with the bound $O(n^{2-d})$ weakened to $o(1)$. Theorem 2.16 is the degree-restricted counterpart of the result of Erdős and Rényi [36] that a random graph with n vertices and $\frac{1}{2}n\log n + O(n)$ edges a.a.s. has connectivity equal to its minimum degree. Without degree restrictions, a much higher edge density is required before higher connectivity occurs, but for both models of random graphs the basic effect is the same: vertices of degree less than d are a.a.s. the only obstruction to d-connectedness.

In [63], Łuczak also considers random graphs with given degree sequence where the minimum entry is 2, obtaining results on the number and nature of the components. Molloy and Reed [85] gave a nice extension of these results, to determine asymptotically necessary and sufficient conditions on the degree sequence $\mathbf{d}$ (with minimum degree 2) for a random graph in $\mathcal{G}_{n,\mathbf{d}}$ to have a.a.s. a giant component.

2.7 Automorphisms and unlabelled graphs

What is the number $U_{n,d}$ of unlabelled d-regular graphs on n vertices? Let $\mathcal{S}$ be any set of unlabelled graphs on n vertices, and $\mathcal{S}'$ the set of all labelled versions of graphs in $\mathcal{S}$. A simple application of the Frobenius–Burnside lemma gives

$$|\mathcal{S}| = \frac{1}{n!}\sum_{\sigma\in S_n} |\mathrm{fix}(\sigma)| \tag{15}$$

where S_n denotes the group of permutations of the label set $\{1, 2, \ldots, n\}$ and $\mathrm{fix}(\sigma)$ denotes the set of graphs in S' fixed by σ. (See Harary and Palmer [54] for a variety of applications of this principle to the enumeration of unlabelled graphs.) With id denoting the identity permutation, $|\mathrm{fix}(\mathrm{id})| = |\mathcal{G}_{n,d}|$ and so

$$\sum_{\sigma\in S_n\setminus\{\mathrm{id}\}} |\mathrm{fix}(\sigma)| = o(|\mathcal{G}_{n,d}|) \tag{16}$$

is equivalent to

$$U_{n,d} \sim |\mathcal{G}_{n,d}|/n!. \tag{17}$$

Rewrite $\mathrm{fix}(\sigma) = \sum_{G\in\mathcal{S}'} x(G,\sigma)$, where $x(G,\sigma)$ is 1 if $G \in \mathrm{fix}(\sigma)$ and 0 otherwise. Reversing the order of summation in (15) then gives

$$|\mathcal{S}| = \frac{1}{n!}\sum_{G\in\mathcal{S}'} \phi(G),$$

where $\phi(G)$ denotes the number of $\sigma \in S_n$ which fix G.

Hence (16), via (17), is equivalent to

$$\mathbf{E}(\phi(G)) \to 1 \quad \text{for} \quad G \in \mathcal{G}_{n,d}. \tag{18}$$

Bollobás [14] showed that this condition holds for fixed $d \geq 3$. In a more general investigation of McKay and Wormald, this was extended to the following result.

Theorem 2.17 ([78]) *Let $\epsilon > 0$ and $3 \leq d = d(n) = O(n^{\frac{1}{2}-\epsilon})$. Then $\mathbf{E}(\phi(G)) \to 1$ for $G \in \mathcal{G}_{n,d}$.*

This and Corollary 2.4 imply the following.

Corollary 2.18 *For $\epsilon > 0$ and $3 \leq d = d(n) = O(n^{\frac{1}{2}-\epsilon})$,*

$$U_{n,d} = \frac{(dn)!}{(dn/2)!2^{dn/2}(d!)^n n!}\exp\left(\frac{1-d^2}{4} - \frac{d^3}{12n} + o(1)\right).$$

Versions of Theorem 2.17 are invariably proved by going back to (16) and noting that if $\sigma \in S_n \setminus \{\mathrm{id}\}$, then for any graph in $\mathrm{fix}(\sigma)$, the arrangement of edges incident with vertices moved by σ is severely restricted. The various

possible subgraphs induced by these edges have a low probability of occurring in $G \in \mathcal{G}_{n,d}$, and the sum of this probability over all such subgraphs and σ is $o(1)$. Bounding $|\mathrm{fix}(\sigma)|$ is achieved by classifying $G \in \mathrm{fix}(\sigma)$ according to the numbers of edges joining vertices in cycles of σ of given small lengths. This leads to optimisation over several variables. A proof is given for bounded d in [120] which is significantly simpler than the others because of the way in which it handles these variables.

Computations to bound the probability of subgraphs occurring can be done directly in the pairing model as in [14] and [120] for fixed d, but for d growing significantly with n as in [78] the switching method (Section 2.4) is required. No doubt Theorem 2.17 is still valid with the upper bound on d increased to $d \le n - 4$, but a proof of this seems out of reach at present.

The results of McKay and Wormald [78] actually treat the model $\mathcal{G}_{n,\mathbf{d}}$ with varying degrees, and results are obtained like Theorem 2.17, but the upper bound on degrees depends on how many vertices of low degree are permitted.

2.8 Packing, covering and colouring

The independence property of the pairing model has been used heavily in packing and covering problems (including colouring, which is equivalent to covering with independent sets). Some algorithms (say, for finding independent sets) can be analysed by generating the pairs of the pairing in the order in which they are examined by the algorithm, as long as the algorithm is "greedy" in the sense that it does not have to know too much about the part of the pairing not yet generated. An upper or lower bound can then be obtained a.a.s. on the graph parameter of interest, by determining the behaviour of the algorithm a.a.s. Several results along these lines have been obtained by showing that the parameter of interest a.a.s. closely tracks a function given as the solution of differential equations throughout the generation process. Convergence is proved using supermartingale inequalities. This is referred to below as the *differential equation method*. It is described in [122], but see [124] for a much more comprehensive and up to date exposition.

Quite apart from this, there is a general concentration result using martingales which applies to many random regular graph parameters which are large. A general statement of the result seems to be lacking so far (though a particular case was referred to by Frieze and Suen [45]), and so we give the following. We write $P \sim P'$ to denote that two pairings P and P' differ by a simple switching.

Theorem 2.19 *If X_n is a random variable defined on $\mathcal{P}_{n,d}$ such that $|X_n(P) - X_n(P')| \le c$ whenever $P \sim P'$, then*

$$\mathbf{P}(|X_n - \mathbf{E}X_n| \ge t) \le 2\exp\left(\frac{-t^2}{dnc^2}\right)$$

for all $t > 0$.

Proof Hold n fixed and write $X = X_n$. Generate an element P of $\mathcal{P}_{n,d}$ by choosing the pairs consecutively. After each pair is completed, canonically choose the next point to be paired by some convention (such as the lowest-numbered unpaired point) and choose its mate randomly. Let P_0 be any pairing with the order of pairs specified according to the order of generation, let $P_0(m)$ denote the set of the first m pairs in P_0, and define

$$Y_m(P_0) = \mathbf{E}(X(P) \mid P_0(m) \subseteq P)$$

for $P \in \mathcal{P}_{n,d}$. Then $Y_0(P), Y_1(P), \ldots, Y_{dn/2}(P)$ is a martingale (in fact a martingale constructed this way is called a Doob martingale; see McDiarmid [65] for example), and we have $Y_0(P)$=$\mathbf{E}X$ and $Y_{dn/2}(P) = X(P)$. The theorem now follows from Hoeffding's inequality (also called Azuma's; see [65]), once we show that $|Y_{m+1}(P) - Y_m(P)| < c$ for all m; that is, the martingale has differences bounded by c.

For given P_0, let i denote the canonical next point after the pairs in $P_0(m)$ are chosen, and let S_j denote those pairings in $\mathcal{P}_{n,d}$ containing all of $P_0(m)$ as well as the pair $\{i, j\}$. Then for any j and k which occur in no pairs in $P_0(m)$, each $P \in S_j$ corresponds to a unique $P' \in S_k$ by the simple switching $\{\{i, j\}, \{k, l\}\} \leftrightarrow \{\{i, k\}, \{j, l\}\}$, where l is determined as the mate of k in P. This gives a bijection between S_j and S_k. By assumption $|X(P') - X(P)| \le c$, and the required bound on the differences now follows because $\mathcal{P}_{n,d}$ is a uniform space. ∎

One way of deriving results on $\mathcal{G}_{n,d}$ using Theorem 2.19 is to consider Y_n defined on $\mathcal{G}_{n,d}$ such that $Y(G(P)) = X(P)$ for all $P \in$ Simple where X_n satisfies the hypotheses of the theorem. Then

$$\mathbf{P}(|Y_n - \mathbf{E}Y_n| \ge t) \le \frac{2}{p} \exp\left(\frac{-t^2}{dnc^2}\right)$$

in $\mathcal{G}_{n,d}$, where p is the probability (in $\mathcal{P}_{n,d}$) of Simple. This gives concentration of X_n and normally implies that $\mathbf{E}X_n$ is well approximated by $\mathbf{E}Y_n$ (but $\mathbf{P}$(Simple) and the range of values of Y_n can affect quantification of this).

2.8.1 Independent sets, dominating sets and star forests An *independent set* or stable set in a graph G is a set of vertices no two of which are adjacent, and the *independence number* $\alpha(G)$ of G is the cardinality of a maximum independent set in G. A *dominating set* is a set S of vertices such that every vertex not in S has at least one neighbour in S, and the *dominating number* $\alpha'(G)$ of G is the cardinality of a minimum dominating set in G.

Since a simple switching in a pairing P can change $\alpha(G(P))$ by at most 1, Theorem 2.19 and Corollary 2.3 imply that $\alpha(G)/|V(G)|$ (called the *independence ratio*) is concentrated when d is fixed. For instance, it is a.a.s. within $\log n/\sqrt{n}$ of its expected value for $G \in \mathcal{G}_{n,d}$. A similar statement holds for the

dominating number. In spite of this, no argument has been forthcoming to show that the expected value of $\alpha(G)/n$ tends to some constant for $G \in \mathcal{G}_{n,d}$ as $n \to \infty$. This (and a similar result for $\alpha'(G)/n$) would imply the following.

Conjecture 2.20 *For fixed d, there are constants $\beta(d)$ and $\beta'(d)$ such that for all $\epsilon > 0$, $|\alpha(G) - n\beta(d)| < \epsilon n$ and $|\alpha'(G) - n\beta'(d)| < \epsilon n$ a.a.s. for $G \in \mathcal{G}_{n,d}$.*

What bounds are known on $\alpha(G)/n$? Bollobás [12] used the expected number of independent sets of a given size to give an upper bound on $\alpha(G)/n$ a.a.s. for $G \in \mathcal{G}_{n,d}$, the upper bound being given in terms of the solution of an equation and being strictly less than $\frac{2\log d}{d}$ for all d. It can be checked that this bound is asymptotic to

$$\gamma_1(d) = \frac{2n}{d}(\log d - \log\log d + 1 - \log 2) \tag{19}$$

as $d \to \infty$. By sharpening the method, McKay [73] improved these slightly to functions $\gamma_2(d)$, a.a.s. for G of sufficiently large (but fixed) girth. In view of the concentration mentioned above and Theorem 2.5, $\gamma_2(d) + \epsilon$ applies, for any $\epsilon > 0$, with no girth restriction. (Beware of McKay's sign error in the terms $1 - \log 2$ in his expression for $\gamma_1(d)$ defined above.) The definition of γ_2 is complicated but numerical computations give $\alpha(G)/n < \gamma_2(3) < 0.4554$ a.a.s. for $G \in \mathcal{G}_{n,3}$, and similarly $\gamma_2(4) < 0.4164$ and $\gamma_2(5) < 0.3845$. McKay expressed disappointment in the small improvement he obtained over $\gamma_1(d)$ as $d \to \infty$, but as we mention below this was inevitable because $\gamma_1(d)$ is quite sharp to order $o(n/d)$ in this asymptotic sense.

Lower bounds were obtained by several authors; the best (at least for small d) comes from analysing what is called in [124] the degree-greedy algorithm: choose vertices to put in the growing independent set randomly; the first, v_1, is chosen uniformly at random, and in general after adding a vertex v_i to the set, delete v_i and its neighbours from the graph and randomly choose the next vertex v_{i+1} (uniformly) from those vertices now of minimal degree. The analysis uses the differential equation method. This was done first by Frieze and Suen [45] for $d = 3$, and independently for arbitrary $d \geq 3$ in [122]. It gives asymptotically almost sure lower bounds $\beta_1(d)$ on $\alpha(G)/n$. For example, $\beta_1(3) = 6\log\frac{3}{2} - 2 = 0.4328$, $\beta_1(4) = 0.3901$ and $\beta_1(5) = 0.3566$. These constants are computed numerically for $d \geq 4$ by solving systems of first order nonlinear differential equations, and the asymptotic behaviour for large d is unknown. Further values are in [122] and [124]. The best known simple formula for a lower bound when $d \geq 4$ is $\frac{1}{2} - \frac{1}{2}\left(\frac{1}{d-1}\right)^{2/(d-2)}$ from [122]. However, this is far from $\beta_1(d)$ for d small, and approaches only $\frac{1}{2}\gamma_1(d)$ for large d. It is obtained from the simpler algorithm which randomly greedily chooses vertices for the independent set regardless of their neighbours, analysed again by the differential equation method. Weaker lower bounds were obtained earlier by Bollobás [12] through lower bounds on the independent set size of all d-regular triangle-free graphs, and arguing that deleting all vertices in short cycles a.a.s.

decreases the independent set size only marginally. Interestingly, the non-deterministic bounds used here are obtained by analysing the greedy algorithm on triangle-free graphs of given average degree. That line of reasoning was also improved by Shearer [105] but with only an iterative formula for arbitrary d which seems always to be less than $\beta_1(d)$.

Frieze and Łuczak [44] proved that functions $\gamma_1(d)+o(n/d)$ for $d \to \infty$, $d < n^{1/3} - \epsilon$, provide upper and lower bounds a.a.s. on $\alpha(G)$, $G \in \mathcal{G}_{n,d}$. However, this result implies nothing for small d. They used an argument partitioning the pairing into large pieces which are then analysed almost separately, and part of the proof relies on the results of McKay and Wormald [79] which were obtained using switchings.

The dominating number seems to have been well studied only in $\mathcal{G}_{n,3}$. Molloy and Reed [86] showed that the obvious lower bound $\frac{1}{4}n$ on $\alpha'(G)$ can be improved slightly to $\alpha'(G) > 0.2636n$ a.a.s. by simply computing the expected number of dominating sets of that size. On the other hand they gave the almost sure upper bound $0.3126n$ by analysing a greedy algorithm which walks along a Hamilton cycle. The analysis uses the differential equation method and relies on Theorem 4.5 which implies that when proving events a.a.s., one can assume $G \in \mathcal{G}_{n,3}$ is a random Hamilton cycle plus a random matching. This is easier to analyse than the algorithm corresponding to the degree-greedy algorithm for independent sets; Duckworth and Wormald [34] applied the differential equation method to such an algorithm and improved the upper bound to $0.2746n$. To summarise, $0.2636 < \alpha'(G)/n < 0.2746$ a.a.s. for $G \in \mathcal{G}_{n,3}$.

As a slight variation of the last result, by analysing almost the same algorithm, it is shown in [34] that $G \in \mathcal{G}_{n,3}$ a.a.s. has a forest of stars containing at least $0.7178n$ edges. The obvious upper bound is $0.75n$, which can be reduced to $0.7390n$ using the usual argument of computing the expected number of star packings with a given number of edges.

2.8.2 Chromatic number Frieze and Łuczak [44] also gave functions asymptotic to $\frac{d}{2\log d}$ for $d \to \infty$ which a.a.s. are upper and lower bounds on the chromatic number $\chi(G)$ for $G \in \mathcal{G}_{n,d}$. Like their bounds on independent set size, these give no restriction for small d. A related result by Bollobás and Clark [18] considers a generalised version of chromatic number.

The chromatic number of G is at least $n/\alpha(G)$, and so almost sure lower bounds for small d are given by the almost sure upper bounds on $\alpha(G)$ mentioned in Section 2.8.1. Better lower bounds come from considering the expected number of k-colourings. For instance, counting bicoloured graphs shows that for fixed $d \geq 3$, $G \in \mathcal{G}_{n,d}$ a.a.s. cannot be 2-coloured. For $r \geq 3$, Molloy and Reed [84] performed calculations of the expected number of r-colourings of $G \in \mathcal{G}_{n,d}$ which show that its nth root is at most $r(\frac{r-1}{r})^{d/2} + o(1)$ (for fixed r and d). Hence, for instance, a.a.s. $\chi(G) \geq 4$ for $G \in \mathcal{G}_{n,d}$ when $d \geq 6$.

Brooks' Theorem (see Diestel [33] for example) gives the upper bound $\chi(G) \leq d$ a.a.s. Non-trivial asymptotically almost sure upper bounds for

small d seem to be harder to obtain. Borodin and Kostochka [23], Catlin [25] and Lawrence [62] all showed that for K_4-free graphs G, $\chi(G) \leq \frac{3}{4}(\Delta(G)+2)$. Hence by Lemma 2.7, $\chi(G) \leq \frac{3}{4}(d+2)$ a.a.s. for $G \in \mathcal{G}_{n,d}$. For triangle-free graphs G there is the bound $\chi(G) \leq \frac{2}{3}\Delta(G)+2$, which has been attributed to Kostochka, and by Theorem 2.5 implies $\chi(G) \leq \frac{2}{3}d+2$ with probability at least $e^{-(d-1)^3/6} - o(1)$ for $G \in \mathcal{G}_{n,d}$.

These results still leave open the question of whether a random 4-regular or 5-regular graph is a.a.s. 3-colourable. An answer to this would be a significant breakthrough.

Knowledge of the edge-chromatic number $\chi'(G)$ is in a much more satisfactory state. It is shown below (Corollary 4.16) that $G \in \mathcal{G}_{n,d}$ a.a.s. has a partition of the edge set into perfect matchings if n is even. Hence $\chi'(G) = d$ a.a.s. for n even, and clearly for n odd it must attain the upper bound $d+1$ of Vizing's theorem.

2.8.3 Linear arboricity The *linear arboricity* of a graph is the minimum number of edge-disjoint linear forests (i.e. forests in which each component is a path) which cover the edges of the graph. The linear arboricity conjecture [1] is that every d-regular graph has linear arboricity precisely equal to the obvious lower bound of $\lceil \frac{1}{2}(d+1) \rceil$. Alon [2] proved using the Lovász Local Lemma that all regular graphs with sufficiently large girth satisfy this. Based on his approach, McDiarmid and Reed proved that a random d-regular graph a.a.s. satisfies the conjecture.

Theorem 2.21 (McDiarmid and Reed [66]) *For fixed d, the linear arboricity of a random d-regular graph is a.a.s. $\lceil \frac{1}{2}(d+1) \rceil$.*

This theorem is proved easily from the more recent results on the structure of random regular graphs, as shown in Section 4.3. It is also shown there that the paths in all but one of the factors can be forced to be Hamilton paths.

2.9 Perfect matchings and Hamilton cycles

The aim of this section is to whet the appetite with some of the results on the existence of certain types of spanning subgraphs of random regular graphs. The method of proof of the strongest results in this direction is presented in Section 4, along with many extensions of these results which assert the existence of factorisations of regular graphs into subgraphs of prescribed types.

Every d-regular $(d-1)$-edge-connected graph of even order has a perfect matching (see, for example, Berge [8]). Hence Theorem 2.10 has an interesting consequence.

Corollary 2.22 (Bollobás [16]) *For fixed $d \geq 3$, $G \in \mathcal{G}_{n,d}$ a.a.s. has a perfect matching if dn is even.*

This fails for $d = 2$ since in that case $G \in \mathcal{G}_{n,d}$ a.a.s. has an odd cycle, and hence no perfect matching. One way to see this is to observe from Theorem 2.5 that, for fixed $k \geq 2$, the probability that $G \in \mathcal{G}_{n,2}$ has no odd cycle of length less than $2k$ is $\exp(-\prod_{i=2}^{k} \frac{1}{4i-2}) + o(1)$, which is $\exp(-\frac{1}{4}\log k) + o(1) \to 0$ as $k \to \infty$. (Here k goes "slowly" to ∞ while n goes "quickly".) Alternatively, this probability can be computed using generating function methods and then evaluated asymptotically (see Wilf [112, Section 5.3] for simlar examples). In $\mathcal{P}_{n,2}$ there is the following exact expression, which is asymptotic to $Cn^{-1/4}$ for a constant C. It then follows from Lemma 2.1 that the probability is also $o(1)$ for $G \in \mathcal{G}_{n,d}$.

Lemma 2.23 *For $P \in \mathcal{P}_{n,2}$, the probability that $G(P)$ has no odd cycles is precisely*

$$\prod_{i=1}^{\lfloor \frac{1}{2}n \rfloor} \frac{4i-2}{4i-1} \ .$$

Proof By the independence property of the pairing model, the random pairing $P \in \mathcal{P}_{n,2}$ can be generated in such a way that the next point paired is always a point in a cell which already contains a paired point (if such a cell exists). Then $G(P)$ is generated by tracing along paths, not starting a new path until the old one becomes a cycle. Let F_i denote the event that the i'th pair chosen completes a cycle in $G(P)$ with the earlier pairs. In order to create no odd cycle, the requirement is that F_{2i-1} must be false for each $1 \leq i \leq \frac{1}{2}n$. (Conditioning on F_{2i-1} being false for all $i \leq j$, if F_{2j} holds then the cycle created will have even length and if F_{2j+1} holds the cycle will have odd length.) Given all earlier choices of pairs, the conditional probability of F_i is $(2n-4i+3)^{-1}$ regardless of the earlier choices. The product of $(1-(2n-4i+3)^{-1})$ over $1 \leq i \leq \frac{1}{2}n$ gives the stated probability. ∎

The expected number and variance of the number of perfect matchings were computed asymptotically by Bollobás and McKay.

Theorem 2.24 (Bollobás and McKay[20]) *For $3 \leq d = d(n) \leq (\log n)^{1/3}$, let M_n denote the number of perfect matchings in $G \in \mathcal{G}_{n,d}$. Then with n restricted to even integers,*

$$\mathbf{E}M_n \sim \sqrt{2}e^{1/4}\left(\frac{(d-1)^{d-1}}{d^{d-2}}\right)^{\frac{1}{2}n}, \qquad \frac{\mathbf{E}M_n^2}{(\mathbf{E}M_n)^2} \to e^{-\frac{1}{4}(2d-1)/(d-1)^2}\sqrt{\frac{d-1}{d-2}}.$$

This result on the first moment is readily derived in the pairing model by counting how many ways one can first choose a perfect matching of the n vertices, then lay down a set of pairs in P which create the chosen matching in $G(P)$, and then complete these pairs to a full pairing. After this, divide by the total number of pairings. To obtain the result in $\mathcal{G}_{n,d}$ one can condition these counts on having no loops or multiple edges, using for instance the method of

moments as in Section 2.3. (Lemma 4.4 below performs this task in general.) The second moment $\mathbf{E}M_n^2$ is calculated in a similar fashion by laying down the pairs corresponding to an ordered pair of perfect matchings, which is much more complicated than the first moment calculation due to an extra variable denoting the number of edges in common between the two perfect matchings.

Since the variance is $\mathbf{Var}M_n = \mathbf{E}M_n^2 - (\mathbf{E}M_n)^2$, it tends to a non-zero constant multiple of $(\mathbf{E}M_n)^2$. This gives some information on the distribution of M_n via *Chebyshev's inequality*

$$\mathbf{P}(|X - \mathbf{E}X| \geq t) \leq \frac{\sigma^2}{t^2},$$

valid for any real-valued random variable X with variance σ^2 and for any $t > 0$. Putting $t < \mathbf{E}X$ gives an upper bound on $\mathbf{P}(X = 0)$, which is loosely called the *second moment method*. This is not sharp enough in this case to imply Corollary 2.22. However, the elaborate refinement of the second moment method presented in Section 4 is sufficient to obtain not only Corollary 2.22, but also the full asymptotic shape of the distribution of the number of perfect matchings.

Ever since the earliest results on random regular graphs, the natural conjecture on Hamiltonicity was (by Bollobás [11], for example) that $G \in \mathcal{G}_{n,d}$ is a.a.s. Hamiltonian for all fixed $d \geq 3$. For $d = 2$ this is false, by (11) for example. The second moment method implies that asymptotically at least a certain percentage of large random 3-regular graphs are Hamiltonian, which was first shown by Robinson and Wormald in [98]. In fact, it was shown there that by restricting to graphs with no triangles, a higher proportion $(2 - 3e^{-13/12})$ is obtained, which can be shown by bijective type methods to be valid back in the unrestricted space $\mathcal{G}_{n,3}$. This idea evolved into the small subgraph conditioning method for estimating the distribution of the number of large subgraphs. The results in the following theorem were derived for $d = 3$ in [98], stated in general in [100] without being used, and then derived and used by Frieze et al. [42].

Theorem 2.25 *For fixed $d \geq 3$, let H_n denote the number of Hamilton cycles in $G \in \mathcal{G}_{n,d}$. Then with dn restricted to even integers,*

$$\mathbf{E}H_n \sim e\sqrt{\frac{\pi}{2n}}\left(\frac{(d-2)^{d-2}(d-1)^2}{d^{d-2}}\right)^{\frac{1}{2}n}, \qquad \frac{\mathbf{E}H_n^2}{(\mathbf{E}H_n)^2} \to \frac{d}{(d-2)e^{2/(d-1)}}.$$

The calculations are similar to those for Theorem 2.24 but derivation of the second moment in particular is considerably more difficult. Again, two copies of the subgraph in question (this time a Hamilton cycle) are laid down and the rest of the cubic graph is completed. (All calculations can be made in the pairing model or directly in the graph model.) The number of ways of doing the laying down is computed as a function of the structure of the graph which

is the intersection of the two Hamilton cycles. Two variables are required: the number of paths in this intersection and the number of isolated vertices. The resulting expression is summed over these two variables and evaluated asymptotically. (In the case $d = 3$ the number of isolated vertices must be 0, and so that case is much simpler than for general d.)

Further sharpening of the small subgraph conditioning method, as described in Section 4, produced the following.

Theorem 2.26 (Robinson and Wormald [99, 100]) *For fixed $d \geq 3$, $G \in \mathcal{G}_{n,d}$ is a.a.s. Hamiltonian.*

This was proved first for large d only, using an entirely different method, by Fenner and Frieze [38] ($d \geq 796$) and independently by Bollobás [15] ($d \geq 10^7$). The method used was developed by Fenner and Frieze [37] for a model of non-regular graphs. It combines the technique which Pósa [90] and Komlós and Szemerédi [61] successfully applied to determine the threshold for Hamiltonicity for $G \in \mathcal{G}(n,p)$ and a colouring argument by which the random graph is split into semi-random pieces of different colours. (This resembles the partitioning argument mentioned in Section 2.8.1 regarding independence number.) Neither of these works was attempting to achieve a sharpest possible result from this method. Frieze [41] later made improvements to cover all $d \geq 85$, giving a polynomial time constructive proof, before the appearance of [99] and [100] settled the matter for all $d \geq 3$.

Theorem 2.26 has now been improved in various ways to give much stronger results. It is now known that for $d \geq 3$, $G \in \mathcal{G}_{n,d}$ is a.a.s. decomposable into edge-disjoint Hamilton cycles (plus a matching if d is odd). In addition, Hamilton cycles are in a certain sense dense in random regular graphs: for $d \geq 3$ there is a.a.s. a Hamilton cycle containing a randomly selected set of $o(\sqrt{n})$ edges of $G \in \mathcal{G}_{n,d}$. These results and others are discussed more fully in Section 4.3.

By putting vertices of degree 2 in each selected edge, the last-mentioned result translates to a result about random graphs with all vertices of degree 3 except for a small number of degree 2. This is possibly the only non-trivial result known on the existence of perfect matchings or Hamilton cycles in $G \in \mathcal{G}_{n,\mathbf{d}}$, and yet the following is certain to be true.

Conjecture 2.27 *Restricting $\mathbf{d} = \mathbf{d}(n)$ to those sequences with elements d_i in the range $3 \leq d_i \leq D$ for some fixed d and D, $G \in \mathcal{G}_{n,\mathbf{d}}$ is a.a.s. Hamiltonian.*

It would be very interesting to prove this even for $d = 3$ and $D = 4$; probably all the difficulties for this are encountered in the special case where half the vertices are of degree 3 and half of degree 4 (mentioned in Section 1.1).

3 Other uniform models

3.1 Unlabelled regular graphs

How does one select an unlabelled graph uniformly at random? As in Section 2.7, let $\mathcal{S}$ be a set of unlabelled graphs on n vertices and let $\mathcal{S}'$ be the set of all labelled versions of them. By applying (15) to each singleton subset of $\mathcal{S}$, it follows that an unlabelled graph can be selected uniformly at random from $\mathcal{S}$ by selecting a labelled $G \in \mathcal{S}'$ with weight proportional to $|\phi(G)|$, and then ignoring the labels on the vertices. From Theorem 2.17 it follows that for $3 \le d = o(n^{\frac{1}{2}-\epsilon})$, a structural property is a.a.s. true of $G \in \mathcal{G}_{n,d}$ if and only if it is a.a.s. true of a random unlabelled d-regular graph n-vertex graph.

3.2 Random bipartite graphs

For a model of random bipartite regular graphs we assume that n is even and that the vertices of one colour are labelled $1, 2, \ldots, \frac{1}{2}n$, as are the vertices of the other colour. The natural pairing model for d-regular bipartite graphs is obvious: the cells containing points are the same as for ordinary graphs, but the random perfect matching is equivalent to a bijection between the points in cells of one colour and those in cells of the other colour. As a result, calculations are usually even easier than in the graph case. O'Neil [89] found that the number of bicoloured d-regular graphs on n vertices (n even) is asymptotic to

$$\frac{(\frac{1}{2}dn)!e^{-\frac{1}{2}(d-1)^2}}{(d!)^n} \tag{20}$$

for $3 \le d < (\log n)^{\frac{1}{4}-\epsilon}$ with $\epsilon > 0$, and the same result (at least for d fixed) was found in [4] using a model equivalent to the pairing model. O'Neil also obtained a few results on random bipartite graphs.

More recently McKay [71] extended (20) to higher d using switchings, and then McKay and Wang [76] extended to $d = o(\sqrt{n})$ using the new type of switchings exemplified by the proof of Lemma 2.9.

For simplicity when we refer to *the* model of random bipartite regular graphs we mean the model described above, which is actually uniform on *bicoloured* graphs. It is *not* uniform on bipartite d-regular labelled graphs because the number of 2-colourings of a bipartite graph G is 2^k where k is the number of components of G. To deal with graphs which are not bicoloured but are rather bipartite, one notes (see below) that, at least for $d \ge 3$, the bicoloured graphs are a.a.s. connected. This implies that the bipartite graphs (taken as $\mathcal{G}_{n,d}$ restricted to bipartite graphs) are similarly a.a.s. connected. For $d = 2$ the situation is a bit more complicated but results can be obtained by exact enumeration and we do not pursue this here. So, in this article, random regular bipartite graphs have the uniform bicoloured distribution.

The distribution of short cycles can be derived just as for graphs. The expected number of cycles of length i is asymptotically $\frac{(d-1)^i}{i}$ for all even

$i \geq 4$, and the numbers of cycles of given bounded lengths are asymptotically independent Poisson. This was all derived in [115].

Ellingham [35] showed the analogues of Theorems 2.10 (for d bounded) and 2.12 on connectivity of random regular bipartite graphs. The main result on the expected number of automorphisms of a random regular graph, Theorem 2.17, also applies in the bipartite case as shown by the very general results of McKay and Wormald [78, Note before Corollary 3.5].

The expectation and variance of the number of perfect matchings in random d-regular bipartite graphs were given asymptotically in [89] and [20] (for quite small d). For Hamilton cycles these quantities were given in [98], where the case $d = 3$ of the following result was proved simply from Chebyshev's inequality.

Theorem 3.1 (Robinson and Wormald [100]) *For fixed $d \geq 3$, a random d-regular bipartite graph is a.a.s. Hamiltonian.*

Godsil and McKay [51] obtained asymptotics for the number of $k \times n$ Latin rectangles for $k = o(n^{6/7})$. As part of the argument they find an asymptotic formula for the number of perfect matchings (respecting the bipartition) in a random edge-coloured bicoloured k-regular graph whose edges have been coloured such that each colour class is a perfect matching. This is equivalent to the bicoloured analogue of the random graph model $k\mathcal{G}_{n,1}$ defined in Section 4.3. Their derivation uses an integral formula to convert the problem into one about short cycles, and from here switchings are used (necessarily, for such a large k).

3.3 Directed graphs

Let $\mathcal{DG}_{n,d}$ denote the uniform model of random d-regular digraphs (in which each vertex has in- and out-degree d).

Cooper [28] showed for $D \in \mathcal{DG}_{n,2}$ that D is a.a.s. strongly 2-connected and [29] that a.a.s. every pair of vertices lies on a common directed cycle. For graphs, 2-connectedness implies such a property, but not so for digraphs.

For $d = 2$ the expected number of directed Hamilton cycles goes to 0, so digraphs in $\mathcal{DG}_{n,d}$ are a.a.s. not Hamiltonian. For larger d the situation is different.

Theorem 3.2 (Cooper et al. [31]) *For $d \geq 3$ a random digraph in $\mathcal{DG}_{n,d}$ a.a.s. has a directed Hamilton cycle.*

The proof of this employs the ideas from Section 4.3, and is outlined near the end of that section. A different but related proof was given by Janson using the small subgraph conditioning method directly (see Theorem 4.12).

The d-regular bipartite graph corresponding to a d-regular digraph has a perfect matching by Hall's theorem (see [22] for example), and so *all* d-regular digraphs have an edge decomposition into 1-regular factors.

3.4 Multigraphs

Nothing much has been done with uniformly distributed multigraphs. The asymptotic number with a given degree sequence (with an upper bound on the degrees) is found in [118], with loops permitted as well as multiple edges. Bender and Canfield [5] have a similar result, but they effectively count a loop as contributing degree 1 to a vertex. These results should be possible to extend to higher degrees.

Of course many results on the uniform model can be deduced from the pairing model, which is not uniform but is close. This can be quantified as follows (here multigraphs permit loops as well).

Theorem 3.3 (Janson [56]) *For d fixed, an event is true a.a.s. for $G(P)$, $P \in \mathcal{P}_{n,d}$, if and only if it is true a.a.s. for d-regular multigraphs chosen u.a.r.*

This does not follow from Theorem 2.6, because multigraphs with many multiple edges have a reduced probability of arising as $G(P)$. Hence Theorem 2.6 leaves open the possibility that a random d-regular multigraph a.a.s. has at least say $\log n$ multiple edges. Some further analysis of the appropriate moments was required to prove this theorem.

3.5 Hypergraphs

There are some results on uniformly distributed d-regular, r-uniform hypergraphs with d and r fixed, as $n \to \infty$ (subject to dn being a multiple of r). For many of the results on random regular graphs, similar techniques will suffice. For instance, Bollobás [10] stated an asymptotic formula for the number of r-uniform d-regular hypergraphs, and Cooper et al. [32] applied the small subgraph conditioning method to show that for fixed positive integers r and s,

$$\lim \mathbf{P}(G_{n,r,s} \text{ has a perfect matching }) = \begin{cases} 0 & s > \sigma_r \\ 1 & s < \sigma_r \end{cases}$$

where

$$\sigma_r = \frac{\log r}{(r-1)\log\left(\frac{r}{r-1}\right)} + 1.$$

3.6 Tournaments

In 1974 Spencer [106] used an argument akin to switchings to prove properties of subgraphs of random regular tournaments. This gave the logarithm of the number of n-vertex regular tournaments asymptotically. This result was superseded by McKay [74], who found an asymptotic formula for the number itself using saddle-point techniques. This method, also used for asymptotic enumeration of tournaments with given score sequence by McKay and Wang [77], is useful for studying random regular tournaments. In this way Gao et al. [46] obtained some results on subgraphs of random tournaments

with a given score sequence, and have shown that the expected number of automorphisms tends to 1.

4 The small subgraph conditioning method, contiguity, and superposition models

When it was finally shown in [99] and [100] that almost all d-regular graphs are Hamiltonian for $d \geq 3$ (Theorem 2.26), the central work involved two applications of essentially the same general method. The method has since proved useful for gaining information on a variety of spanning subgraphs, as well as other large subgraphs, of random regular graphs. The strongest consequences are given in Section 4.3.

The way in which the method works can be described as follows. (A precise development of how to use it comes in the following subsections.) The random variable $Y = Y_n$ of interest (for instance, the number of Hamilton cycles in $G \in \mathcal{G}_{n,d}$) has a variance which is of a size comparable with the square of its expectation (as in, for example, Theorems 2.24 and 2.25). Thus, Chebyshev's inequality is not strong enough to show that a.a.s. $Y > 0$. However, the explanation for the large variance lies in the fact that the distribution of Y is affected by the presence of certain small but not too common subgraphs in the random graph—usually the short cycles of given lengths. (By not too common, we mean the expected number is bounded.) It turns out (in the cases where the method works) that conditioning on the small subgraph counts (up to some preselected size of small subgraphs) affects $\mathbf{E}Y$, altering it by some constant factor. However, luckily and yet mysteriously, such conditioning reduces the variance of Y, to the point that conditioning on the numbers of enough small subgraphs reduces the variance to any desired small fraction of $(\mathbf{E}Y)^2$.

Perhaps this is not a total mystery. For instance, one can imagine that the number of ways in which a Hamilton cycle or perfect matching can "go through" a short cycle depends on the length of the short cycle. Probably the main mystery is not that conditioning on the small cycle counts affects $\mathbf{E}Y$, but that the variance can be made *so* small by conditioning on the small subgraph counts. The computations which are performed in carrying out the small subgraph conditioning method demonstrate this as a fact but leave the user with no underlying explanation of the effect. These computations can be described as follows. The set of all regular graphs can be divided into groups according to the small subgraph counts. The variance $\mathbf{Var}Y$ of Y can be written as the expected value of the variance within a group, plus the variance of the group mean. The calculations show that the latter consumes almost all of the variance of Y, so the variance within any group is small. One deduction which can be made from this is that $Y > 0$ a.a.s. This explains the basic argument used in [99], where Y counts Hamilton cycles, to show that $G \in \mathcal{G}_{n,3}$ is a.a.s. Hamiltonian.

Moreover, the asymptotic distribution of Y is determined from the distri-

bution of the group means and the sizes of the groups. The fact that the group means of the most common groups are close to each other implies that the value of Y does not vary too much in the random graphs (except for a rare event). For instance, taking Y as the number of perfect matchings in $G \in \mathcal{G}_{n,d}$, this shows that the number of perfect matchings in such a random graph usually varies by only a constant factor. This was used in [100] to deduce that adding a random perfect matching to an edge-disjoint random $(d-1)$-regular graph produces a random d-regular graph with something close to the uniform distribution. This relationship is called contiguity of the two models—perfect matching plus $(d-1)$-regular on the one hand, and d-regular on the other hand—as defined precisely below. Hence, if random $(d-1)$-regular graphs are a.a.s. Hamiltonian, so are random d-regular graphs, and so Theorem 2.26 follows by induction on d from the case $d = 3$. This is the second major use of the small subgraph conditioning method, and it leads naturally to the establishment of relationships between various non-uniform models of random regular graphs which are called here superposition models.

4.1 Contiguity of models

One of the most important conclusions from the small subgraph conditioning method is well described in terms of altered probabilistic models. Suppose that Y is a non-negative integer random variable defined on a space $\mathcal{G}$ with $\mathbf{E}Y \neq 0$. We define[1] a new model $\mathcal{G}^{(Y)}$ with the same underlying set as $\mathcal{G}$ by weighting the probability of each element G by $Y(G)$. That is, the probability of G in $\mathcal{G}^{(Y)}$ equals the probability in $\mathcal{G}$ multiplied by $Y(G)/\mathbf{E}Y$. Thus the probability of an event H in $\mathcal{G}^{(Y)}$ is $\mathbf{E}_{\mathcal{G}}(Y \wedge 1_H)/\mathbf{E}_{\mathcal{G}}Y$, where 1_H is the indicator function of H and $\wedge$ denotes the conjuction of events. We have already encountered this idea in the proof of Lemma 2.8 with $Y = [X_{k,n}]_{r_k}$.

It is a very interesting property of two different random graph models if all events true a.a.s. in one model are true a.a.s. in the other. To make this precise, suppose that $(\mathcal{G}_n)_{n\geq 1}$ and $(\widehat{\mathcal{G}}_n)_{n\geq 1}$ are two sequences of probability spaces such that $\mathcal{G}_n$ and $\widehat{\mathcal{G}}_n$ differ only in the probabilities. We say that these sequences are *contiguous*[2] if a sequence of events A_n is a.a.s. true in $\mathcal{G}_n$ if and only if it is a.a.s. true in $\widehat{\mathcal{G}}_n$, in which case we write

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

Contiguity is clearly an equivalence relation on these sequences of spaces.

[1]This definition is for discrete probabilistic spaces, a context which can be assumed for this article. The method of this section holds in continuous spaces as well, by defining $dQ/dP = Y/\mathbf{E}_{\mathcal{G}}Y$ where P and Q are the probability measures in $\mathcal{G}$ and $\mathcal{G}^{(Y)}$ respectively, and the σ-algebra is the same for both spaces.

[2]Formally, two sequences P_n and Q_n of probability measures are contiguous if P_n and Q_n are defined on the same measurable space $(\mathcal{G}_n, \mathcal{F}_n)$ for each n, and for every sequence of measurable sets A_n, $\lim_{n\to\infty} P_n(A_n) = 0$ iff $\lim_{n\to\infty} Q_n(A_n) = 0$. Here I have defined the (sequences of) spaces to be contiguous if the probabilities are.

The original approach in [99] and [100] is sufficiently strong for all the applications of the small subgraph conditioning method which have occurred to date and was presented in general by Molloy et al. [87, Theorem 1]. However Janson [56] was able to streamline the procedure by adjusting the method of proof so as to eliminate several of the conditions required in [87]. This simplification has made little difference in applications of the method: it turns out that the extra conditions have been trivially satisfied in every application so far. However, the simplification is definitely advantageous.

Theorem 4.1 ([56], see also [87]) *Let $\lambda_i > 0$ and $\delta_i \geq -1$, $i = 1, 2, \ldots$, be real numbers and suppose that for each n there are random variables $X_i = X_i(n)$, $i = 1, 2, \ldots$, and $Y = Y(n)$ defined on the same probability space $\mathcal{G} = \mathcal{G}(n)$ such that X_i is non-negative integer valued, Y is non-negative and $\mathbf{E}Y > 0$ (for n sufficiently large). Suppose furthermore that*

(a) For each $k \geq 1$ X_i, $i = 1, 2, \ldots, k$ are asymptotically independent Poisson random variables with $\mathbf{E}X_i \to \lambda_i$;

(b)

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

for every finite sequence $j_1, \ldots, j_k$ of non-negative integers;

(c) $\sum_i \lambda_i \delta_i^2 < \infty$;

(d) $\dfrac{\mathbf{E}Y_n^2}{(\mathbf{E}Y_n)^2} \leq \exp\left(\sum_i \lambda_i \delta_i^2\right) + o(1) \qquad \text{as } n \to \infty.$

Then

$$\mathbf{P}(Y_n > 0) = \exp\left(-\sum_{\delta_i = -1} \lambda_i\right) + o(1),$$

and, provided $\sum_{\delta_i=-1} \lambda_i < \infty$,

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

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

The proof of this theorem is based on the ideas in the discussion at the start of Section 4. The role of condition (b) is to deduce by the method of moments the joint distribution of the X_i in $\mathcal{G}^{(Y)}$, which then gives the expected value of Y_n conditioned on $X_1, \ldots, X_k$. For the proof the factor $(1 + \delta_i)$ need not be interpreted. It is an adjustment to the expected value of X_i in the space $\mathcal{G}^{(Y)}$, so that $\delta_i = -1$ only if conditioning on the event $X_i > 0$ in $\mathcal{G}$ suppresses the

value of Y asymptotically. Condition (d) has been at other times stated with "=" instead of "≤", but "≤" implies "=" anyway (see Janson [56, Theorem 1, Note 4]), basically because variance is always non-negative.

The hypotheses of the theorem are normally verified in much the same way on each occasion. Most of the times this method has been used so far, the X_i are short cycle counts and so (a) is simply Theorem 2.5 or Theorem 2.6. In all other cases, the method of moments has sufficed for verifying (a). Part (d) is verified like Theorems 2.24 or 2.25, and (b) is usually like a combination of the two, harder than (a) but easier than (d).

It follows immediately from contiguity that the value of Y in $\overline{\mathcal{G}}$ is usually "close" to its expected value.

Corollary 4.2 *If the conditions of the theorem are satisfied then*

$$\lim_{\epsilon\to 0^+}\left\{\lim_{n\to\infty}\mathbf{P}_n\left(\epsilon<\frac{Y_n}{\mathbf{E}Y_n}<\frac{1}{\epsilon}\right)\right\}=1$$

where $\mathbf{P}_n$ denotes probability in $\overline{\mathcal{G}}$.

From the arguments in [99] and [100] it was clear that when the hypotheses of the theorem are satisfied, the distribution of Y is determined asymptotically. Janson made this explicit.

Theorem 4.3 (Janson [56]) *With the same hypotheses as Theorem 4.1,*

$$\frac{Y}{\mathbf{E}Y}\xrightarrow{d}W=\prod_{i=1}^{\infty}(1+\delta_i)^{Z_i}e^{-\lambda_i\delta_i}\qquad \text{as } n\to\infty,$$

where the variables Z_i are independent Poisson variables with $\mathbf{E}Z_i=\lambda_i$ for $i\geq 1$. Moreover, this convergence and the convergence of the X_i to the Z_i expressed in Theorem 4.1 (a) all hold jointly.

Thus in each case where contiguity is proved using Theorem 4.1, we can if so desired deduce the asymptotic distribution of the appropriate variable immediately by Theorem 4.3. Moreover, the joint convergence mentioned here means that the Z_i are linked to the X_i. Thus the distribution of Y conditioned on any values of a finite number of X_i is also determined asymptotically, for example the distribution of Y conditioned on $X_i=k$ is asymptotically that of W conditioned on $Z_i=k$. This makes the contiguity assertion of Theorem 4.1 easy to verify in view of the proof of [87, Corollary 1].

One more result is often useful when the X_i are short cycle counts in $\mathcal{P}_{n,d}$. It is readily proved by applying Lemma 2.8 to the space $\mathcal{G}^{(Y)}$.

Lemma 4.4 (Janson [56]) *If the conditions of Theorem 4.1 hold and $Y'(n)$ has the distribution of $Y(n)$ conditioned on $X_1(n)=X_2(n)=0$ then*

$$\frac{\mathbf{E}Y'(n)}{\mathbf{E}Y(n)}\to\exp(-\lambda_1\delta_1-\lambda_2\delta_2),\quad \frac{\mathbf{E}(Y'(n)^2)}{(\mathbf{E}Y'(n))^2}\to\exp(-\lambda_1\delta_1-\lambda_2\delta_2)\frac{\mathbf{E}(Y(n)^2)}{(\mathbf{E}Y(n))^2}.$$

We mention one application of contiguity immediately. Janson showed in [56] that $G(P)$, for $P \in \mathcal{P}_{n,d}$, is contiguous to a uniformly chosen random regular multigraph (with loops permitted). Theorem 3.3 follows directly from this.

4.2 Applications to regular spanning subgraphs and long cycles

In order to discuss in some detail how to verify the hypotheses of Theorem 4.1, we consider the result of the arguments in [99], expressed in the language of contiguity as follows.

Theorem 4.5 *Let H_n denote the number of Hamilton cycles in $G \in \mathcal{G}_{n,3}$. Then $\mathcal{G}_{n,3}^{(H_n)} \approx \mathcal{G}_{n,3}$.*

Proof The first and second moments of H_n are given by Theorem 2.25 with $d = 3$. However, although the final result will be in $\mathcal{G}_{n,3}$, it is a little simpler to work with Theorem 4.1 where $\mathcal{G}(n) = \mathcal{P}_{n,3}$. In this case $Y = Y(n)$ is defined as the number of Hamilton cycles in the multigraph corresponding to $P \in \mathcal{P}_{n,3}$. We have

$$\frac{\mathbf{E}Y^2}{(\mathbf{E}Y)^2} \to 3. \tag{21}$$

(In $\mathcal{P}_{n,d}$ for $d \geq 3$ the limit is $\frac{d}{d-2}$, as shown in [42].) This second moment calculation is the most difficult part of this whole proof, and is virtually the same as the calculation required for $\mathcal{G}_{n,d}$.

Let $X_i = X_i(n)$ denote the number of cycles of length i in $G(P)$ for $P \in \mathcal{P}_{n,3}$. The condition of Theorem 4.1 (a) follows from Theorem 2.6 with

$$\lambda_i = \frac{2^{i-1}}{i} \tag{22}$$

($i \geq 1$). Working towards Theorem 4.1 (b), we next show that for any fixed $i \geq 1$

$$\frac{\mathbf{E}(YX_i)}{\mathbf{E}Y} \to \lambda_i(1 + \delta_i) \tag{23}$$

where

$$\delta_i = \frac{(-1)^i - 1}{2^i}.$$

Let D be some fixed set of pairs of points corresponding to a Hamilton cycle in pairings in $\mathcal{P}_{n,3}$. By symmetry all copies of D are equivalent and so in $\mathcal{P}_{n,3}$

$$\frac{\mathbf{E}(YX_i)}{\mathbf{E}Y} = \mathbf{E}(X_i \mid D \subseteq P).$$

If C is the set of pairs corresponding to an i-cycle (in which case we also call C itself an i-cycle), we classify C according to the configuration of paths in $G(D \cap C)$. We can assume that there must be at least one such path as all

pairs in C cannot be in D provided $n > i$. Give these paths a consistent orientation along C (which multiplies counts by 2) and distinguish one path as first (which multiplies counts by the number of paths and induces a linear ordering of paths around C). Thus

$$\frac{\mathbf{E}(YX_i)}{\mathbf{E}Y} = \sum_Q \frac{1}{2|Q|}\mathbf{E}(X_i(Q) \mid D \subseteq P) \tag{24}$$

where Q denotes the sequence of lengths of paths, $|Q|$ is the number of paths in Q and $X_i(Q)$ is the number of i-cycles in P consistent with such a configuration Q. Fix on such a Q with k paths. There are asymptotically $(2n)^k$ ways to choose the starting points of the paths on D together with their directions along D. Almost all such points are well spaced for n large, and once they are chosen the pairs in C, if it is to correspond to an i-cycle yielding Q, are determined. The probability that these pairs all occur in $P \in \mathcal{P}_{n,3}$ conditional upon $D \subseteq P$ is asymptotically n^{-k}. Hence $\mathbf{E}(X_i(Q) \mid D \subseteq P) \to 2^{|Q|}$ and so (24) becomes

$$\frac{\mathbf{E}(YX_i)}{\mathbf{E}Y} \to \sum_{k\geq 1} \frac{2^k}{2k} \mid \{Q : |Q| = k\} \mid. \tag{25}$$

The ordinary generating function for the number of configurations Q with x marking the total number of vertices involved and y marking the number of paths is $\frac{g(x,y)}{1-g(x,y)}$ where $g(x,y)$ is the generating function for one path; that is, $\frac{yx^2}{1-x}$. Thus, with square brackets denoting extraction of coefficients,

$$\begin{aligned}
\frac{\mathbf{E}(YX_i)}{\mathbf{E}Y} &\to \sum_{k\geq 1} \frac{2^k}{2k}[x^i y^k]\frac{yx^2}{1-x-yx^2} \\
&= [x^i]\sum_{k\geq 1}\frac{1}{2k}[y^{k-1}]\frac{2x^2}{1-x-2yx^2} \\
&= [x^i]\frac{1}{2}\int_0^1 \frac{2x^2}{1-x-2yx^2}\,\mathrm{d}y \\
&= -[x^i]\tfrac{1}{2}\log(1-x-2x^2) + \tfrac{1}{2}\log(1-x),
\end{aligned}$$

and (23) follows. To obtain the condition of Theorem 4.1 (b), one observes that this argument works in general for higher moments and gives the required result.

Noting $\sum_{i\geq 1}\lambda_i\delta_i^2 = \log 3$ and recalling (21), we see that the hypotheses of Theorem 4.1 are satisfied. As $\delta_i = -1$ only for $i = 1$, $\overline{\mathcal{P}}_{n,3}$ is the restriction of $\mathcal{P}_{n,3}$ to pairings P with $X_1 = 0$, i.e. for which $G(P)$ has no loops. The conclusion is that $\overline{\mathcal{P}}_{n,3}^{(Y)} \approx \overline{\mathcal{P}}_{n,3}$. The theorem now follows using the definition of contiguity by restricting to $X_2 > 0$, since in $\mathcal{P}_{n,3}$ the probability of this event tends to a non-zero constant. ∎

Notes (i) An alternative way of proceeding from (25) to (23), which in some applications leads to significantly simpler computations (see Janson [56]), is the following. Each intersection $D \cap C$ corresponds to a sequence in $\{0,1\}^i$ determined by walking along C in a given direction from an arbitrary starting edge (which multiplies counts by $2i$) and writing 0 for an edge not in $D \cap C$ and 1 for an edge in $D \cap C$. Each 0 must be followed by 1 and contributes a factor 2 (by an argument analogous to (25)). Regarding 0 and 1 as two states in a Markov chain (or something similar), and imposing the condition that the final state is equal to the initial state (since the first edge must be returned to) we see that the sum over sequences is $\mathrm{Tr}(A^i)$ where

$$A = \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}.$$

As the sequence of all 1's is impossible, it follows that $\mathbf{E}(YX_i)/\mathbf{E}Y \to \frac{1}{2i}(\mathrm{Tr}(A^i) - 1)$. Then by noting that A has eigenvalues 2 and -1, we obtain $\mathrm{Tr}(A^i) = 2^i + (-1)^i$ and then (23) as required.

(ii) Working the proof in $\mathcal{G}_{n,3}$ instead of $\mathcal{P}_{n,3}$ merely requires eliminating loops and multiple edges in a couple of places, which can always be done using the method of moments. Alternatively, as in [99, 100], one can appeal directly to the enumeration results of Bender and Canfield [5] to count the ways of adding edges to an existing graph without creating multiple edges, since their main theorem allows arbitrary forbidden edges as long as a bounded number are incident with a vertex.

From the proof of Theorem 4.5 and Theorem 4.3 we obtain the following.

Corollary 4.6 (Janson[56]) *Let H_n denote the number of Hamilton cycles in $G \in \mathcal{G}_{n,3}$. Then*

$$\frac{H_n}{\mathbf{E}H_n} \xrightarrow{d} W = \prod_{i=3}^{\infty} (1+\delta_i)^{Z_i} e^{-\lambda_i \delta_i} \qquad \text{as } n \to \infty,$$

where $\lambda_i = \frac{2^{i-1}}{i}$, $\delta_i = \frac{(-1)^i - 1}{2^i}$, and Z_i are independent Poisson variables with $\mathbf{E}Z_i = \lambda_i$ for $i \geq 3$.

Proof These are the λ_i and δ_i obtained in the proof of Theorem 4.5, so Theorem 4.3 gives the claimed result for $G(P)$, $P \in \mathcal{P}_{n,3}$ but with $i \geq 1$. By the joint convergence of the variables, restricting to $\mathcal{G}_{n,3}$ (i.e. $X_1 = X_2 = 0$) is asymptotically equivalent to restricting to $Z_1 = Z_2 = 0$. ∎

Similar arguments, using [42] for the variance calculation, give the following.

Theorem 4.7 (Frieze et al. [42], Janson [56]) *Let H_n denote the number of Hamilton cycles in $G \in \mathcal{G}_{n,d}$. Then $\mathcal{G}_{n,d}^{(H_n)} \approx \mathcal{G}_{n,d}$ for $d \geq 3$.*

To obtain a result generalising Corollary 4.6, one only needs to note the values $\lambda_i = \frac{(d-1)^i}{2i}$, $\delta_i = \frac{(-1)^i - 1}{(d-1)^i}$ and that X_i is again the number of i-cycles (so the product in the definition of W is again over $i \geq 3$ for distribution in $\mathcal{G}_{n,d}$). This was done by Janson [56].

Proof of Theorem 2.26 This now follows immediately by the definition of contiguity, since the event $\{H_n > 0\}$ has probability 1 in $\mathcal{G}_{n,d}^{(H_n)}$, so it is true a.a.s. in $\mathcal{G}_{n,d}$. ∎

The proof of Theorem 2.26 in [100] was different. It involved proving the analogous property of perfect matchings given in Theorem 4.8, and then argued using contiguity inductively as developed in Section 4.3. In the case of n odd and d even this result could not be used directly, but instead another result was proved which related to the distribution of the number of matchings which miss just one vertex. (However, Janson [56] was the first to state the contiguity result for perfect matchings explicitly.) Other results on the distribution of numbers of spanning regular subgraphs in $\mathcal{G}_{n,d}$ have been obtained using the same method, with the conclusions summarised in the next theorem.

A *k-factor* of a graph is a k-regular spanning subgraph. (A perfect matching is the edge set of a 1-factor.) A *1-factorisation* of a d-regular graph is a set of d edge-disjoint 1-factors, equivalent to a partition of the edge set into perfect matchings. For enumeration purposes, we assume these matchings are ordered. (For $d = 3$ the number of 1-factorisations is then equal to the number of ordered pairs of disjoint 1-factors, called double 1-factors in [87].) For n even, let M_n (defined in $\mathcal{G}_{n,d}$) denote the number of perfect matchings and let T_n (defined in $\mathcal{G}_{n,3}$) denote the number of 1-factorisations of G. For all n let D_n denote the number of 2-factors in $G \in \mathcal{G}_{n,d}$ (D comes from dwa, which is Polish for 2), and let B_n denote the number of ordered pairs of edge-disjoint Hamilton cycles in $G \in \mathcal{G}_{n,4}$ (a good name for these is Hamilton bicycles; hence the letter B). The following result on M_n essentially appears in [100] and was given explicitly by Janson [56], that on T_n was obtained independently by Janson [56] and by Molloy et al. [87], that on D_n is by Robalewska [95], and the one on B_n is by Kim and Wormald [60].

Theorem 4.8 *Restricting to n even, $\mathcal{G}_{n,d}^{(M_n)} \approx \mathcal{G}_{n,d}$ for all $d \geq 3$, and $\mathcal{G}_{n,3}^{(T_n)} \approx \mathcal{G}_{n,3}$. With no restrictions on n, $\mathcal{G}_{n,d}^{(D_n)} \approx \mathcal{G}_{n,d}$ for $n \geq 4$ and $\mathcal{G}_{n,4}^{(B_n)} \approx \mathcal{G}_{n,4}$.*

Proving these results is like proving Theorem 4.7, but the calculations are much easier in the case of M_n, somewhat easier for T_n and D_n. The result for 2-factors (D_n) when $d = 4$ was conjectured in [56]. (More precisely, the stronger multigraph version was conjectured for $G(P)$ where $P \in \mathcal{P}_{n,d}$ restricted to

loopless multigraphs, which no doubt follows in the same way. As noted in Section 2, this space of multigraphs is not uniform.) For edge-disjoint Hamilton cycles (B_n) the same method was again used, but a really new idea was required to compute the variance, including a proof of Theorem 4.11 below. In fact, it was a conjecture in [100] that a random 4-regular graph a.a.s. decomposes into two edge-disjoint Hamilton cycles, and the related contiguity result was later conjectured by Janson [56]. (Actually, the multigraph version was conjectured, which also follows from the results in [60].)

In each of these applications, the X_i are the short cycle counts as before, and so $\lambda_i = \frac{(d-1)^i}{2i}$ as in (22). Other relevant values (all in the graph space $\mathcal{G}_{n,d}$) are given in the following table. Hamilton cycles (H_n) are included for comparison.

Y_n	$\mathbf{E}Y_n \sim$	$\mathbf{E}Y_n^2/(\mathbf{E}Y_n)^2 \to$	δ_i
H_n	$e\sqrt{\frac{\pi}{2n}}\left(\frac{(d-2)^{d-2}(d-1)^2}{d^{d-2}}\right)^{\frac{1}{2}n}$	$\frac{d}{(d-2)e^{2/(d-1)}}$	$\frac{(-1)^i-1}{(d-1)^i}$
M_n	$\sqrt{2}e^{1/4}\left(\frac{(d-1)^{d-1}}{d^{d-2}}\right)^{\frac{1}{2}n}$	$e^{-\frac{1}{4}(2d-1)/(d-1)^2}\sqrt{\frac{d-1}{d-2}}$	$\frac{(-1)^i}{(d-1)^i}$
T_n	$2\sqrt{e}(\frac{4}{3})^{n/2}$	$4e^{-5/4}$	$\frac{(-1)^i}{2^{i-1}}$
D_n	$e^{1/4}\sqrt{2}\left(\frac{(d-1)(d-2)^{(d-2)/2}}{d^{(d-2)/2}}\right)^n$	$e^{-\frac{1}{4}(2d-1)/(d-1)^2}\sqrt{\frac{d-1}{d-2}}$	$\frac{(-1)^i}{(d-1)^i}$
B_n	$e^{7/4}\frac{\pi}{\sqrt{8}n}(\frac{3}{2})^n$	$e^{-55/36}\sqrt{24}$	$\frac{-2+(-1)^i}{3^i}$

Theorem 2.24 is used for the second moment of M_n. The values of the first two moments of M_n and T_n for $G(P)$, $P \in \mathcal{P}_{n,d}$ (and hence the corresponding contiguity results for this multigraph space) are given in [56]; in all applications examined so far these are obtained by omitting the powers of e, since these and only these factors arise when sieving to remove loops and multiple edges (see Lemma 4.4). Similarly, the contiguity relating to B_n for $G(P)$, $P \in \mathcal{P}_{n,d}$ restricted to loopless multigraphs, follows from the results in [60]. For general d, Janson also derives the value $\delta_i = \frac{(-1)^i}{(d-1)^{i-1}}$ relevant to 1-factorisations of d-regular graphs, but it seems to be too difficult to obtain the variance for general d.

In all cases, the limiting distribution of Y_n is given immediately from Theorem 4.3 (and this is done in [56] for M_n and T_n, and in [95] for D_n).

The obvious corollary of the result for T_n is that $G \in \mathcal{G}_{n,3}$ a.a.s. has a 1-factorisation for even n, but this already follows from Theorem 2.26 with $d = 3$. The result for M_n gives another proof of Corollary 2.22. The result for D_n shows that $G \in \mathcal{G}_{n,4}$ a.a.s. has a 2-factor, but this event has probability 1 anyway by Petersen's theorem (see [33] for example). So the only new result we obtain of this type is with B_n.

Corollary 4.9 *A random 4-regular graph a.a.s. decomposes into two edge-disjoint Hamilton cycles.*

The surprising equality of the values in the last two columns of the table for M_n and D_n has the following consequence.

Corollary 4.10 (Robalewska [95]) *For even n,*

$$\frac{M_n}{\mathbf{E}M_n} \qquad \text{and} \qquad \frac{D_n}{\mathbf{E}D_n}$$

have the same limit distribution in $\mathcal{G}_{n,d}$ ($d \geq 3$).

Moreover, from the joint convergence given by Theorem 4.1 and the fact that the X_i are the same variables in each of these two cases, it follows that M_n and D_n are far from independent: they are asymptotically linked, and we can write

$$\frac{M_n}{\mathbf{E}M_n} \sim \frac{D_n}{\mathbf{E}D_n} \qquad \text{a.a.s.}$$

in $\mathcal{G}_{n,d}$.

Computing the variance of B_n in $\mathcal{G}_{n,4}$ required the following result, which interestingly relates to the models of random regular graphs in Section 4.3.

Theorem 4.11 (Kim and Wormald [60]) *Let n be even. Given four independent random matchings M_i, $i = 0, \ldots, 3$, of n vertices, the probability that $M_i \cup M_{i+1}$ induces a Hamilton cycle for each i (with subscripts mod 4) is asymptotic to p_H^4 as $n \to \infty$, where $p_H \sim \sqrt{\frac{\pi}{2n}}$ is the probability that two independent random matchings of n vertices induce a Hamilton cycle.*

There are recent extensions of Theorem 4.5 in other directions. Robinson and Wormald [101] established the analogue of this theorem in which $o(\sqrt{n})$ edges of $G \in \mathcal{G}_{n,3}$ are chosen u.a.r., and the variable H_n counts only Hamilton cycles which pass through these edges. A corollary is that if that number of edges of $G \in \mathcal{G}_{n,3}$ are chosen at random, then a.a.s. there is a Hamilton cycle passing through them all. The limiting distribution of the probability was also found, if $c\sqrt{n}$ edges are chosen. This is an interesting example in that the variables X_i in the use of Theorem 4.1 are not just the short cycles counts, but also the counts of short paths joining distinguished edges. A similar result is by Janson and Wormald [57] that if the edges of $G \in \mathcal{G}_{n,d}$ are randomly coloured in n colours, $\frac{1}{2}d$ of each colour, then a.a.s. there is a Hamilton cycle using precisely one of each colour (provided $d \geq 8$).

Moving away from regular spanning subgraphs, in [98] the expectation and variance of the number of decompositions of $G \in \mathcal{G}_{n,3}$ into a tree and a cycle were computed asymptotically. (The length of cycle is necessarily $\frac{1}{2}n + 1$.) Janson [55] computed the joint distribution with short cycles and verified that the small subgraph conditioning method performs as expected; that is, almost all cubic graphs have such a decomposition. The difficult part is as usual the

variance, covered in [98], but the short cycle part of the calculation is perhaps more complicated than all the other examples done so far.

Finally, as an example of use of this method with non-spanning subgraphs, there are the results of Garmo mentioned in Section 2.5.

Several results have appeared on the random regular bipartite model. The calculations for Hamilton cycles in bipartite graphs for the case $d = 3$ are given in [98], in which case the simple application of Chebyshev's inequality to the expectation and variance is enough to deduce that the probability of Hamiltonicity tends to 1. The same is true of the general case $d \geq 4$ (but this does not seem to have been published). The bipartite analogue of Theorem 2.26 was proved in [100] *not* by proving the bipartite analogue of Theorem 4.7, but it did involve verifying the bipartite analogue of the statement in Theorem 4.8 about M_n. The variance of H_n in the bipartite case requires a function to be maximised virtually the same as that encountered for $\mathcal{G}_{n,d}$ in [42]. This calculation is almost the same as in the general graph case, and no doubt calculations for D_n also follow a similar pattern in bipartite graphs. The bipartite analogue for T_n (1-factorisations) of the statement in Theorem 4.8 is verified in [87].

Let H_n denote the number of (directed) Hamilton cycles in a digraph in $\mathcal{DG}_{n,d}$ (defined in Section 3.3). Janson [56] used the small subgraph conditioning method to obtain the following.

Theorem 4.12 (Janson [56]) $\mathcal{DG}_{n,d}^{(H_n)} \approx \mathcal{DG}_{n,d}$ *for all* $d \geq 3$.

Theorem 3.2 is a corollary of this.

Although the results discussed in this section are strong in some ways, the method is extremely difficult to extend to examination of $\mathcal{G}_{n,\mathbf{d}}$ in general, and no good results in this direction seem to be known. It is not known for instance if Conjecture 2.27 has any chance of being proved by contiguity.

4.3 The superposition arithmetic of contiguity classes

Originating from the argument in [100], the notion of the union of two random regular graphs on the same vertex set is very useful for proving asymptotic properties of $\mathcal{G}_{n,d}$. If $\mathcal{G}$ and $\widehat{\mathcal{G}}$ are two probability spaces of random graphs or multigraphs on the same vertex set, we define their *sum* $\mathcal{G} + \widehat{\mathcal{G}}$ to denote the space whose elements are defined by the random multigraph $G \cup \widehat{G}$ (called the *superposition* of G and $\widehat{G}$) where $G \in \mathcal{G}$ and $\widehat{G} \in \widehat{\mathcal{G}}$ are generated independently. Similarly, define the *graph-restricted sum* of $\mathcal{G}$ and $\widehat{\mathcal{G}}$, denoted by $\mathcal{G} \oplus \widehat{\mathcal{G}}$, to be the space which is the restriction of $\mathcal{G} + \widehat{\mathcal{G}}$ to simple graphs (i.e. with no multiple edges—we have no cause to use this operation when loops are present). This is defined only if $\mathcal{G} + \widehat{\mathcal{G}}$ contains at least one simple graph. In order to ensure that all our probability spaces can be sensibly related, we assume when using these operations that the underlying sets are extended to cover all graphs (or multigraphs, as the case may be) on that vertex set. Thus

any n-vertex graph which cannot be formed by the operation in question is included in the space, but with probability 0. We loosely call models defined by sums *superposition models*.

The two sum operations $+$ and $\oplus$ are clearly commutative and associative. We define

$$k\mathcal{G} = \mathcal{G} \oplus \cdots \oplus \mathcal{G},$$

with k terms on the right. (The analogous product of k by $\mathcal{G}$ using $+$ instead of $\oplus$ is not so interesting for our purposes here.)

It is also straightforward to derive the following result.

Lemma 4.13 (Janson[56]) *Suppose that $\mathcal{G}_n \approx \mathcal{G}'_n$ and $\widehat{\mathcal{G}}_n \approx \widehat{\mathcal{G}}'_n$ where all four spaces $\mathcal{G}_n$, $\mathcal{G}'_n$, $\widehat{\mathcal{G}}_n$ and $\widehat{\mathcal{G}}'_n$ are of graphs on n vertices. Then*

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

Janson [56] discussed contiguity of multigraph models defined by both the sum $+$ and the graph-restricted sum $\oplus$, but in that paper the emphasis is on the sum. Results for the graph-restricted sum can be obtained using the multigraph results, and it is desirable to put this on a systematic base. A crucial property required of the graph models in order for this idea to work is that the probability that a simple graph is created in the sum space must be bounded away from 0 (which incidentally ensures that the graph-restricted sum is defined). To ensure this, we use the property of every natural model, that it is *label-independent*, by which we mean that the probability of any multigraph is the same as that of any particular relabelled version of that multigraph.

Lemma 4.14 *Suppose that $\mathcal{G}_n \approx \mathcal{G}'_n$ and $\widehat{\mathcal{G}}_n \approx \widehat{\mathcal{G}}'_n$ where all four spaces $\mathcal{G}_n$, $\mathcal{G}'_n$, $\widehat{\mathcal{G}}_n$ and $\widehat{\mathcal{G}}'_n$ are label-independent and all graphs in the spaces have bounded degree (and n vertices). Then*

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

Proof Let A_n be any sequence of events which is a.a.s. true in $\mathcal{G}_n \oplus \widehat{\mathcal{G}}_n$. We can regard A_n as an event in $\mathcal{G}_n + \widehat{\mathcal{G}}_n$ also (after all, it is just a set of graphs). Let B_n be the event that $G \in \mathcal{G}_n + \widehat{\mathcal{G}}_n$ has a multiple edge. Then $A_n \vee B_n$ is a.a.s. true in $\mathcal{G}_n + \widehat{\mathcal{G}}_n$ (where $\vee$ denotes the union of events) and hence also in $\mathcal{G}'_n + \widehat{\mathcal{G}}'_n$ by Lemma 4.13. It then follows that A_n is a.a.s. true in $\mathcal{G}'_n \oplus \widehat{\mathcal{G}}'_n$ once we show that the probability of the complement of B_n in $\mathcal{G}'_n + \widehat{\mathcal{G}}'_n$ is bounded below by a positive constant.

For this it suffices to treat each pair of graphs $G_1 \in \mathcal{G}'_n$ and $G_2 \in \mathcal{G}'_n$ separately, and consider a random relabelling of each. We can use the method of moments. It is easily verified that the expected number λ of edges in common between two such random relabellings is exactly half the product of the average degrees of vertices in G_1 and G_2. The other moments are calculated and

satisfy the relation (8) in Section 2.3 (with $k = 1$). Hence (by Brun's sieve or Lemma 2.8) the probability that the superposition of the relabelled copies of G_1 and G_2 creates no multiple edges is asymptotically $e^{-\lambda}$, which is bounded below in view of the upper bound on vertex degrees.

The reverse argument, from $\mathcal{G}'_n \oplus \widehat{\mathcal{G}}'_n$ to $\mathcal{G}_n \oplus \widehat{\mathcal{G}}_n$, is identical. ∎

For bicoloured graphs the analogue of Lemma 4.14 holds if the definition of label-independence is specified to include only those relabellings which rearrange the labels within each of the two label sets.

If $Y = Y(n)$ is the number of spanning k-regular subgraphs of $G \in \mathcal{G}_{n,d}$ which lie in some specified set S_n, then $G \in \mathcal{G}_{n,d}^{(Y)}$ has the distribution of $F(n) \cup F'(n)$ where the ordered pairs $(F(n), F'(n))$ are sampled uniformly such that $F(n) \in S_n$, $F'(n)$ is a $(d-k)$-regular graph on the same vertex set, and $F(n)$ and $F'(n)$ are edge-disjoint. Thus $\mathcal{G}_{n,d}^{(Y)} = \mathcal{U}_n \oplus \mathcal{G}_{n,d-k}$ where $\mathcal{U}_n$ is the uniform space on S_n. Similarly, if Y is the number of decompositions of the edge set of $G \in \mathcal{G}_{n,d}$ into j spanning k-regular subgraphs of $G \in \mathcal{G}_{n,d}$ each of which lie in a uniform space $\mathcal{U}_n$ (so $jk = d$) then $\mathcal{G}_{n,d}^{(Y)} = j\mathcal{U}_n$. So, for example, $\mathcal{G}_{n,3}^{(T_n)} = 3\mathcal{G}_{n,1}$ for n even. We can now restate Theorems 4.7 and 4.8 in the following form. Let $\mathcal{H}_n$ denote a uniformly random Hamilton cycle (n even), on the same n vertices as $G \in \mathcal{G}_{n,d}$ for all d. The only item in the following not explicitly covered already is (iv) for $d = 3$, which is equivalent to (ii) for $d = 3$.

Theorem 4.15

(i) $\mathcal{G}_{n,d-2} \oplus \mathcal{H}_n \approx \mathcal{G}_{n,d}$ *for* $d \geq 3$.

(ii) $\mathcal{G}_{n,d-1} \oplus \mathcal{G}_{n,1} \approx \mathcal{G}_{n,d}$ *for* $d \geq 3$ *and* n *even.*

(iii) $3\mathcal{G}_{n,1} \approx \mathcal{G}_{n,3}$ *(*n *even).*

(iv) $\mathcal{G}_{n,d-2} \oplus \mathcal{G}_{n,2} \approx \mathcal{G}_{n,d}$ *for* $d \geq 3$.

(v) $2\mathcal{H}_n \approx \mathcal{G}_{n,4}$.

We next examine implications of this theorem, taking an interesting special case first. From (ii) (used repeatedly), and (iii), and Lemma 4.14,

$$\mathcal{G}_{n,d} \approx d\mathcal{G}_{n,1} \tag{26}$$

for n even and $d \geq 3$. Since every element of $d\mathcal{G}_{n,1}$ has a 1-factorisation, this implies the following.

Corollary 4.16 (Robinson and Wormald [100]) *For* $d \geq 3$, $G \in \mathcal{G}_{n,d}$ *a.a.s. has a 1-factorisation when* n *is even.*

Another way to say this is that the edge-chromatic number of these graphs is a.a.s. equal to d when n is even, which is best possible since it must be $d+1$ by Vizing's theorem when n is odd.

The main problem with $d = 2$ is that $\mathcal{G}_{n,1} \oplus \mathcal{G}_{n,1} \not\approx \mathcal{G}_{n,2}$, since, as seen in Section 2.9, $G \in \mathcal{G}_{n,2}$ a.a.s. has an odd cycle, but $G \in 2\mathcal{G}_{n,1}$ clearly does not. Similarly, $\mathcal{G}_{n,1} \oplus \mathcal{G}_{n,1} \not\approx \mathcal{H}_n \not\approx \mathcal{G}_{n,2}$, but these three form the only exceptions in the rather pleasant arithmetic of contiguity classes of regular graph models. Earlier, weaker versions of this result [56, 87] appeared at the times when various parts of Theorem 4.15 were proved.

Corollary 4.17 *Let $d \geq 3$, and suppose $d = 2j + \sum_{i=1}^{d-1} ik_i$ with all terms non-negative. Then*

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

with n restricted to even integers if $k_i \neq 0$ for any odd i.

Proof Lemma 4.14 is used abundantly throughout. From (i), (iv) and (v),

$$\mathcal{G}_{n,d} \approx j\mathcal{H}_n \oplus k\mathcal{G}_{n,2} \oplus \mathcal{G}_{n,d-2j-2k} \tag{27}$$

for any $k \leq \frac{1}{2}d - 2j$. If $k_i = 0$ for all odd i, take $k = \frac{1}{2}d - 2j$ and combine the copies of $\mathcal{G}_{n,2}$ into the desired spaces $\mathcal{G}_{n,i}$ using the same result in reverse (with $j = 0$) for each space. If not, we can assume n is even, use (27) with $k = k_2$ and (26) to get

$$\mathcal{G}_{n,d} \approx j\mathcal{H}_n \oplus k_2\mathcal{G}_{n,2} \oplus (d - 2j - 2k_2)\mathcal{G}_{n,1}$$

(unless $d - 2j - 2k_2 = 2$). Then recombine the copies of $\mathcal{G}_{n,1}$ into all the other terms required using (26) in reverse for each term $\mathcal{G}_{n,i}$, $i \geq 3$. The required k_1 copies of $\mathcal{G}_{n,1}$ will be surplus. The only case left is $d - 2j - 2k_2 = 2$ and $k_1 = 2$, whence either $j > 0$ or $k_2 > 0$. From above we have $\mathcal{G}_{n,d} \approx j\mathcal{H}_n \oplus (k_2 + 1)\mathcal{G}_{n,2}$, and any two of these spaces can be recombined to give $\mathcal{G}_{n,4}$. This can then be split as desired since $\mathcal{G}_{n,4} \approx \mathcal{G}_{n,3} \oplus \mathcal{G}_{n,1} \approx \mathcal{G}_{n,2} \oplus 2\mathcal{G}_{n,1}$ by (ii) (twice) and $\mathcal{G}_{n,3} \approx \mathcal{H}_n \oplus \mathcal{G}_{n,1}$ by (i). ∎

A *complete Hamiltonian decomposition* of a d-regular graph is a partition of its edge set into the edges of $\frac{d}{2}$ Hamilton cycles (for d even), or $\frac{d-1}{2}$ Hamilton cycles and a perfect matching (for d odd).

Corollary 4.18 (Kim and Wormald [60]) *For fixed $d \geq 3$, $G \in \mathcal{G}_{n,d}$ a.a.s. has a complete Hamiltonian decomposition.*

One restatement of Corollary 4.9 is that a random 4-regular graph a.a.s. has four 1-factors such that a certain pairwise union gives two Hamilton cycles. The following strengthening of this has not even been proved yet for $d = 3$. A *perfect 1-factorisation* of a d-regular graph is an edge-decomposition into d pairwise disjoint perfect matchings such that the union of any two of them gives a Hamilton cycle.

Conjecture 4.19 *For $d \geq 3$, $G \in \mathcal{G}_{n,d}$ a.a.s. has a perfect 1-factorisation when n is even.*

I also conjecture the contiguity version of this; i.e. that $\mathcal{G}_{n,d}^{Y_n} \approx \mathcal{G}_{n,d}^{Y_n}$ where Y_n is the number of perfect 1-factorisations of G.

There is strong evidence for this conjecture from the results in [60], where it is shown that imposing Hamiltonicity conditions on pairs of perfect matchings does not asymptotically alter the probability of such conditions holding between other pairs (at least, not for a small number of pairs). If this is true in general then the expected number of perfect 1-factorisations will be large. I believe that Theorem 4.11 extends as follows.

Conjecture 4.20 *Let n be even. The probability that $d \geq 4$ random matchings of n vertices create a perfect 1-factorisation of a d-regular graph is asymptotic to $(\frac{\pi}{2n})^{d(d-1)/4}$ as $n \to \infty$.*

The case $d = 3$ is not conjectured here, because it is proved in [60], which means that the expected number of perfect 1-factorisations in $G \in \mathcal{P}_{n,3}$ or $\mathcal{G}_{n,3}$ is known. The variance, as usual, is much harder to compute. The expected number of 1-factorisations of $G \in \mathcal{G}_{n,d}$ grows exponentially with n, so by the contiguity expressed in (26) it is a.a.s. exponentially large. It seems unlikely that none of these 1-factorisations is perfect, even if Conjecture 4.20 is in error by a large factor. This supports Conjecture 4.19.

Corollary 4.18 is quite a strong statement about the structure of $G \in \mathcal{G}_{n,d}$ and enables quick proof of some results already known. For even d it immediately shows that $G \in \mathcal{G}_{n,d}$ is a.a.s. d-edge-connected, giving the edge connectivity version of Theorem 2.10 for d fixed. Here are some other examples.

Proof of Theorem 2.21 Take the decomposition in Corollary 4.18, and remove one edge from each Hamilton cycle. Since there are only $\lfloor \frac{1}{2}d \rfloor$ such cycles, for n sufficiently large we can easily arrange that the removed edges form a matching (d even) or form a linear forest with the perfect matching in the decomposition (d odd). The Hamilton paths form the other linear forests, to obtain linear arboricity $\lfloor \frac{1}{2}d \rfloor + 1 = \lceil \frac{1}{2}(d+1) \rceil$. ∎

This proof also shows that the conclusion of Theorem 2.21 can be considerably strengthened: all but one of the forests in this proof are Hamilton paths, and the other has maximum path length 1 (even d) or 3 (odd d).

The model $\mathcal{H}_n \oplus \mathcal{G}_{n,1}$ has already been studied for its diameter, but without the convenience of contiguity. A weaker version of the following result was obtained by Bollobás and Chung [19], with lower bound $\log_2 n - 10$ and upper bound $\log_2 n + \log_2 \log n + 10$.

Corollary 4.21 *A Hamilton cycle plus a random matching a.a.s. gives a graph with diameter between $\log_2 n + \log_2 \log n - 4$ and $\log_2 n + \log_2 \log n + 4$.*

Proof This comes immediately from Theorems 4.5 and 2.13 with $d = 3$. ∎

Bollobás and Chung [19] also investigated the diameter of a fixed regular graph plus a random matching in order to show that it is possible to find large regular graphs with small diameter using randomisation for only a small part of the graph.

Janson [56] conjectured that the model of random 2-regular graphs obtained by taking a random permutation digraph and ignoring the directions of edges fits in with the arithmetic of models just like the models $\mathcal{H}_n$ and $\mathcal{G}_{n,d}$ (again requiring $d \geq 3$). However, no results in this direction have been forthcoming, probably because of the lack of a suitable framework for combining this model with $\mathcal{G}_{n,d}$. We conjecture that the model obtained from $\mathcal{G}_{n,2}$ by restricting to graphs with even cycles fits in the same way. (This is *not* the same as $2\mathcal{G}_{n,1}$—the probabilities are different.)

The theory of the superposition arithmetic of contiguity classes has progressed further for random regular graph models than for related models. For bipartite graphs, Theorem 3.1 is proved in [100] in the same way as Theorem 2.26 by treating 1-factors in random $(d-1)$-regular bipartite graphs, and examination of the argument will verify the analogue of Theorem 4.15 (ii) in the bipartite case. The proof of Theorem 3.2 in [31] also uses contiguity in just the same way: the bipartite regular graph related to a regular digraph is contiguous to the graph-restricted sum of a perfect matching and a bipartite $(d-1)$-regular graph. The result is proved first for $d = 3$ working with the two random parts of this decomposition, and the results for larger d again follow by contiguity.

Some of the other necessary base results in this area are missing (see the comments on this near the end of Section 4.2). But this is due mainly to the fact that they are almost the same as the corresponding graph case. For directed graph models, even more results are missing, but this is due to real difficulties. Permutation digraphs, or 1-regular digraphs (with loops permitted) are equivalent to matchings in the bipartite graphs corresponding to d-regular digraphs, so contiguity is known for these. On the other hand, in view of Corollary 4.18 and Theorem 4.12 we have the following.

Conjecture 4.22 *A random digraph in $\mathcal{DG}_{n,d}$ a.a.s. has d edge-disjoint directed Hamilton cycles.*

We also of course conjecture the contiguity version of this. From Theorem 4.12 this would imply contiguity of sums of directed Hamilton cycles.

5 The generation problem

The difficulty of sampling from $\mathcal{G}_{n,d}$, i.e. of generating regular graphs u.a.r., and the related problem for $\mathcal{G}_{n,\mathbf{d}}$, are part of the motivation for the "algorithmic" models given in Section 6. Those models do not have uniform, or even well understood, distributions. In this section we focus on uniform and near-uniform distribution.

5.1 Uniform generation for d large

Generating an element of $\mathcal{G}_{n,d}$ is easy in the following sense: just generate an element P of $\mathcal{P}_{n,d}$ and use $G(P)$ if it has no loops or multiple edges. If it is not suitable, repeat. This is implicit in the presentation of the pairing model by Bollobás [10], and in the model used by Bender and Canfield in [5]. It was discussed explicitly in [119], along with a non-probabilistic algorithm for generating 3-regular graphs (which is much more complicated).

The difficulty of this simple procedure is the number of repetitions required before success. The expected number is exactly $1/\mathbf{P}(\text{Simple})$, which from (3) and (5) is prohibitively large for even quite small d, say $d = 8$. What can we do for larger d?

McKay and Wormald [79] used the modern version of switchings to give an algorithm for generating a random d-regular graph for $d = O(n^{1/3})$ in which the expected time per graph is polynomial ($O(n^2d^4)$ in a version that would be implementable, or $O(nd^3)$ in a version that qualifies as the programmer's nightmare). The idea is to generate a random pairing in the pairing model, where those pairings corresponding to multigraphs with given numbers of loops and multiple edges are all equally likely, and then use a random switching to get to a random pairing with fewer loops and multiple edges. An accept/reject procedure is used to make sure the resulting random pairing is uniformly distributed with those parameters. Iterating this procedure eventually reaches pairings P such that $G(P)$ is a simple d-regular graph uniformly distributed. The "old" switchings do not help with this problem even for $d = n^\epsilon$, and the idea definitely cannot be extended in any easy way to get an algorithm for degree $n^{1/3+\epsilon}$ with polynomial expected time per graph generated.

Earlier Tinhofer [109] considered choosing edges consecutively, each choice a uniform choice from some restricted set, so that a regular graph hopefully results. This is combined with an accept/reject procedure to produce the uniform distribution, which is based on *a posteriori* computation of the probability $p(G)$ of the generated graph. The probability of acceptance is the ratio of $p(G)$ to an upper bound on $p(G')$ for all G'. Unfortunately basically nothing is known about such upper bounds, with the result that such algorithms do not seem to be of practical use for uniform generation, except perhaps for very small graphs.

The description of how to generate a random unlabelled graph given in Section 3.1 was applied directly in [121] to generating random unlabelled d-regular graphs. The expected time complexity per graph is linear in n for d-regular graphs on n vertices if d is fixed. Achieving this requires judicious use of accept/reject procedures, as well as the simplification given in [120] of the multidimensional optimisation problem involved in Theorem 2.17. However, in an uncommon twist, the algorithm is not at all practicable for $d = 3$ and 4 unless n is quite *large*, say $n \geq 200$. This is because symmetries are more common for smaller n than for larger n.

For random graphs in $\mathcal{G}_{n,\mathbf{d}}$, the result in [79] applies to degree sequences for which the maximum degree is $O(m^{1/4})$, where m is the sum of the degrees. The result is a polynomial time algorithm for uniform generation. Those in [109] apply to any degree sequence (but tell us little).

5.2 Near-uniform generation

When uniform generation is difficult, it is worth looking at near-uniform generation. If the approximation to uniform can be made to a given accuracy, then probabilities can be obtained experimentally to the same accuracy.

Jerrum and Sinclair [58] gave a *fully polynomial almost uniform generator* for d-regular graphs on n vertices; that is, a generation algorithm which, for $\epsilon > 0$, runs in time polynomial in n and $\log(1/\epsilon)$, such that all graphs are generated with probabilities varying by a factor of at most $1 + \epsilon$. This result is remarkable in that it applies for *all* d. The analysis required to prove it uses sophisticated eigenvalue techniques for estimating the rate of convergence of a Markov chain to the stationary distribution. Unfortunately the polynomial does not have very low order so this algorithm seems to be of little practical significance.

Łuczak and Wormald [64] studied a generalised form of the random processes described below in Section 6 which suggested the algorithm studied by Steger and Wormald [107]. This seems to be a very practical algorithm (time $O(nd^2)$ expected per graph for small d, and good practical results for all $d \leq \frac{1}{2}n$, which suffices by complementation). Although it does not give the uniform distribution, for $d < n^c$ where c is a small positive constant, it produces all graphs with asymptotically the same probability. That is, for such d there are upper and lower bounds on the probabilities of a graph on n vertices being generated which are asymptotically equal as $n \to \infty$. Unfortunately, there are indications that this algorithm does not give such a close approximation to uniform distribution when d gets past $n^{1/3}$.

This algorithm has useful properties even though it is equivalent to the quite naive idea of generating (non-uniformly) at random an element of the pairing model one pair at a time, at each step selecting only from those pairs which will not create loops or multiple edges with those already selected. This process occasionally gets stuck, with unpaired points remaining and not legally able to be paired, but experimental evidence strongly supports the following conjecture, with the value of the constant approximately $\frac{1}{3}$.

Conjecture 5.1 *Let $f(n, d)$ denote the probability that this procedure terminates with a regular graph. Then $f(n, d)$ is bounded below by a constant for all $d \leq \frac{1}{2}n$.*

Jerrum et al. [59] showed that the algorithm in [58] applies to a very wide class of degree sequences. The proof relies on showing that if the degrees do not vary by too much, the number of graphs does not changes radically with

small perturbations of the degrees (maintaining the same number of vertices). This incidentally gives an explanation of the difficulty of enumerating graphs or generating random graphs with given degrees when the degrees vary wildly.

6 Algorithmically defined models

As explained in Section 5, it can be difficult to efficiently generate members of $\mathcal{G}_{n,d}$ uniformly at random. There are no good practical schemes for large d, and the state of affairs for $\mathcal{G}_{n,\mathbf{d}}$ is even worse. It is substantially these difficulties which motivate the "algorithmic" models given here, but aside from this they are often also of intrinsic interest due to their simplicity.

Tinhofer's scheme (see Section 5.1) is based on selection of edges by repeating a uniform selection from a restricted set. Wilson [113] uses a similar approach (ignoring questions of distribution). A simple form of such an algorithm is the following. Given d and n, start with n isolated vertices and repeatedly add edges joining vertices of degree strictly less than d. Each time, the edge added is chosen uniformly at random from all unfilled positions. This is the *degree restricted graph process* with parameter d, or d-process for short. The d-process stops when no more edges can be added, i.e. when the graph induced by the vertices of degree less than d is a clique.

This process was considered by Erdős, who asked for the asymptotic distribution of the number of vertices of degree less than d in the final graph (d fixed, $n \to \infty$). This question was settled in [102] using the differential equation method together with some other arguments. It was shown that a.a.s. the final graph is regular if dn is even, and almost regular, with one vertex of degree $d-1$ and the rest of degree d, otherwise.

The final graph of the d-process, conditioned on it being d-regular, gives a model of d-regular graphs which we denote here by $\mathcal{G}^{\deg}_{n,d}$. Using the fact that the numbers of vertices of given degree follow close to the solutions of differential equations, Ruciński and Wormald [103] also determined the limiting distribution of the numbers of short cycles in $G \in \mathcal{G}^{\deg}_{n,2}$. These are asymptotically independent Poisson, but the expected number of cycles of length i involves an i-dimensional integral. For $i = 3$ this simplifies to $\frac{1}{2}\int_0^\infty \frac{(\log(1+x))^2\,\mathrm{d}x}{xe^x} = 0.1887\ldots$, which is a little different from the expected number $\frac{1}{6}$ of 3-cycles in the uniform model $\mathcal{G}_{n,2}$. More recently, the same authors [104] showed that $G \in \mathcal{G}^{\deg}_{n,d}$ is a.a.s. connected for $d \geq 3$, and the following is conjectured there.

Conjecture 6.1 *For $G \in \mathcal{G}^{\deg}_{n,3}$, $\mathbf{P}(G$ is disconnected$) \sim cn^{-2}$ where $c \approx 0.25$.*

The corresponding statement in $\mathcal{G}_{n,3}$ holds with $c = \frac{2}{27}$.

If we do not condition on the final graph of the d-process being regular, the probability of disconnectedness seems, from simulation results, to be much higher than this. Nevertheless, the following seems quite plausible.

Conjecture 6.2 *Let $d \geq 3$ be fixed. For $G \in \mathcal{G}^{\deg}_{n,d}$, G is a.a.s. d-connected.*

Telcs and Wormald [108], using a delicate application of the differential equation method, found the probability that $G \in \mathcal{G}^{\deg}_{n,2}$ is a single Hamilton cycle is asymptotic to $\sqrt{\frac{\pi e^\tau}{2n}} \approx 1.819 n^{-1/2}$ where $\tau = \int_0^\infty \frac{(\log(1+x))^2\,\mathrm{d}x}{xe^x}$, and also determined the expected number of cycles in $G \in \mathcal{G}^{\deg}_{n,2}$ quite accurately. These values are also close to those in the uniform model $\mathcal{G}_{n,2}$ (see (11) for example). Although the evidence is not very strong, the following is suspiciously hard to disprove.

Conjecture 6.3 *For fixed $d \geq 1$ and dn even, $\mathcal{G}_{n,d} \approx \mathcal{G}^{\deg}_{n,d}$.*

Motivated by [109] and [113], in which the aim was to conveniently generate random graphs with given degrees, Robalewska [96] studied a random process by which stars are added to fill the required vertex degrees. We consider the regular case here, and take d fixed. Beginning with n isolated vertices, at each step choose u.a.r. a vertex v of minimum degree, and choose u.a.r. $d - d(v)$ other vertices of degree strictly less than d. Edges are added from v to these vertices, then the step is repeated. As with the d-process, stop when the required edges do not exist or the graph is d-regular. This is called the *star d-process*. Multiple edges are impossible by the degree-filling nature of the process.

In the case $d = 2$ the short cycle distribution of the final graph was obtained in [96], as well as the asymptotic probability of Hamiltonicity and the distribution of the number of cycles. It was also shown that the final graph is a.a.s. 2-regular. The latter result was extended to d-regular for fixed d, provided dn is even, in [97]. The methods are similar to those used for d-processes. Again, we can consider restricting to those processes in which the final graph to d-regular, to obtain a probability space which we call here $\mathcal{G}^{\mathrm{stardeg}}_{n,d}$.

Conjecture 6.4 *For fixed $d \geq 1$ and dn even, $\mathcal{G}_{n,d} \approx \mathcal{G}^{\mathrm{stardeg}}_{n,d}$.*

7 A wider perspective

There are several interesting random graph models which involve either special regular graphs or in some sense near-regular graphs. After mentioning some cubic examples, we look at results on planar graphs and near-regular graphs in a bit more detail.

McKay et al. [75] found the asymptotic number of claw-free cubic graphs. Together with the result in [101], their result implies that almost all of these graphs are Hamiltonian.

Garmo [48, 49] defined random railways to be random cubic multigraphs with one of the three half-edges at each vertex distinguished, and studies special connectivity properties of these (properties defined with respect to the distinguished half-edges).

7.1 Random regular planar graphs

For enumeration purposes, planar graphs are normally studied embedded in the sphere and with a rooting which consists of distinguishing an edge and a vertex incident with that edge. For 3-connected graphs the enumeration is equivalent to labelled planar graphs (just multiply the number of rooted graphs by $\frac{n!}{4n}$ to get the corresponding number of labelled planar graphs).

Many enumeration results and several on random planar graphs were obtained in this way, for instance Tutte [110] showed that a random 3-connected planar graph a.a.s. has no automorphisms. Then a more general theory of random planar graphs began with the paper of Richmond et al. [92] showing that a random 3-connected cubic planar graph has exponentially small probability of being Hamiltonian. Richmond and Wormald [93] extended this result to more classes of planar maps. These papers also give results on the frequency of subgraphs in such graphs.

Up to this point, results were quite dependent on generating function analysis, but Bender et al. [6] freed the topic from generating functions to a large extent (although their argument to show this uses a modification of the generating function singularity arguments in [92, 93]). Building on this, Richmond and Wormald [94] showed that many classes of planar maps are almost all asymmetric, thus including a simpler proof of the result of Tutte mentioned above (amongst other things).

Bender et al. [7], returning to generating function methods, showed that in almost all 3-connected cubic planar graphs with n edges, the largest cyclically 4-edge-connected cubic component has about $n/2$ edges. See also Gao and Wormald [47] for more general results.

7.2 Non-regular models

Some random graph models cannot easily be specialised to models of regular graphs, and yet have features in common, in particular vertices of low degree but with a guaranteed lower bound on the degrees. Here is only a brief mention of some of the results of this type.

Several authors have considered the random graph coming from a random digraph in which k arcs are chosen out of each vertex at random, and then the orientations are suppressed and multiple edges are coalesced. Fenner and Frieze [37] showed for $k \geq 23$ that this is a.a.s. Hamiltonian. The big open problem in this area is whether this is true for all $k \geq 3$. For $k = 2$ the opposite is true, as sketched in an exercise in Bollobás [16, Section VIII.5], because of the asymptotically almost sure occurrence of small subgraphs called spiders which kill Hamilton cycles. The result for $k \geq 23$ has been far superseded by the recent result of Cooper and Frieze [30] that a random 3-in, 3-out digraph (in which from each vertex three arcs are randomly chosen in and three out) is a.a.s. Hamiltonian.

An early result of Walkup [111], subsequently used by others, is that a

random d-out directed bicoloured graph with n vertices in each part (in which from each vertex d arcs are randomly chosen out) a.a.s. contains a perfect matching provided $d \geq 2$ (and a.a.s. does not if $d = 1$). Frieze [40] later showed that the superposition of two random 1-out-regular digraphs a.a.s. has a perfect matching (n even). Also in the spirit of the models in Section 4, Frieze et al. [43] showed that the superposition of five random trees is a.a.s. Hamiltonian. The big open problem in this direction is whether three is enough.

Recently a model of random regular graphs with edge faults has been studied. Here the edges of $G \in \mathcal{G}_{n,d}$ are randomly deleted independently with probability p each. For example, Goerdt [52] looked at the giant component, and Nikoletseas and Spirakis [88] obtained concentration results on the second eigenvalue of the adjacency matrix of the giant component.

7.3 Further unsolved problems

Besides the various conjectures sprinkled through this paper, there is one issue which benefits from emphasising. Large values of d are in general a problem. For many of the results mentioned in this article, the obvious conjecture is that the result holds for much higher values of d (such as Conjecture 2.11). Can the generating function method in [80] help here? It certainly does permit computation of the probability of sets of edges being present, to some extent. For another example, can one in this way show that for $d \sim \frac{1}{2}n$, $G \in \mathcal{G}_{n,d}$ a.a.s. has no non-trivial automorphisms? At least it would be good to extend Theorem 2.17 past $d = o(n^{1/2})$. Also can a practical near-uniform generation algorithm be found for very large d (say $d \approx \sqrt{n}$ or $d = cn$)? In addition, further results on random tournaments (Section 3.6) would be interesting, but this suffers for one thing from a problem similar to the case of $\mathcal{G}_{n,d}$ with large d: high edge density restricts switchings.

For some of these questions, asymptotic enumeration of $|\mathcal{G}_{n,d}|$ in the range $\sqrt{n} < d < cn/\log n$ may help. This is still wide open, and has strong implications for the model of the degree sequence of a random graph in $\mathcal{G}(n,p)$ which is given in [82].

Acknowledgements

This research was supported by the Australian Research Council. I am grateful to Brendan McKay for reading a version of the manuscript carefully, and to the anonymous referee for many suggestions and corrections to the presentation.

References

[1] J. Akiyama, G. Exoo & F. Harary, Covering and packing in graphs III, cyclic and acyclic invariants, *Mathematica Slovaca*, **30** (1980), 405–417.

[2] N. Alon, The linear arboricity of graphs, *Israel Journal of Mathematics*, **62** (1988), 311–325.

[3] N. Alon & J. H. Spencer, *The Probabilistic Method*, Wiley, New York (1992).

[4] A. Békéssy, P. Békéssy & J. Komlós, Asymptotic enumeration of regular matrices, *Studia Scientiarum Mathematicarum Hungarica*, **7** (1972), 343–353.

[5] E. A. Bender & E. R. Canfield, The asymptotic number of non-negative integer matrices with given row and column sums, *Journal of Combinatorial Theory, Series A*, **24** (1978), 296–307.

[6] E. A. Bender, Z. C. Gao & L. B. Richmond, Submaps of Maps I. General 0-1 Laws, *Journal of Combinatorial Theory, Series B*, **55** (1992), 104–117.

[7] E. A. Bender, L. B. Richmond & N. C. Wormald, Largest 4-connected components of 3-connected planar triangulations, *Random Structures & Algorithms*, **7** (1995), 273–285.

[8] C. Berge, *Graphes et Hypergraphes*, Dunod, Paris (1970).

[9] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, Preprint Series, Matematisk Institut, Aarhus Universitet, 1979.

[10] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European Journal of Combinatorics*, **1** (1980), 311–316.

[11] B. Bollobás, Random graphs, in *Combinatorics* (ed. H. N. V. Temperley), *London Mathematical Society Lecture Note Series*, 52, Cambridge University Press, Cambridge (1981), pp. 80–102.

[12] B. Bollobás, The independence ratio of regular graphs, *Proceedings of the American Mathematical Society*, **83** (1981), 433–436.

[13] B. Bollobás, Distinguishing vertices of random graphs, in *Graph Theory* (ed. B. Bollobás), *North-Holland Math. Stud.*, 62, North-Holland, Amsterdam (1982), pp. 33–49.

[14] B. Bollobás, The asymptotic number of unlabelled regular graphs, *Journal of the London Mathematical Society*, **26** (1982), 201–206.

[15] B. Bollobás, Almost all regular graphs are Hamiltonian, *European Journal of Combinatorics*, **4** (1983), 97–106.

[16] B. Bollobás, *Random Graphs*, Academic Press, London (1985).

[17] B. Bollobás, The isoperimetric number of a random graph, *European Journal of Combinatorics*, **9** (1988), 241–244.

[18] B. Bollobás & L. Clark, Generalized chromatic numbers of random regular graphs, in *Graph Theory, Combinatorics and Algorithms, Vol. 1, 2* (eds. Y. Alavi & A. Schwenk), Wiley, New York (1995), pp. 209–219.

[19] B. Bollobás & F. R. K. Chung, The diameter of a cycle plus a random matching, *SIAM Journal on Discrete Mathematics*, **1** (1988), 328–333.

[20] B. Bollobás & B. D. McKay, The number of matchings in random regular graphs and bipartite graphs, *Journal of Combinatorial Theory, Series B*, **41** (1986), 80–91.

[21] B. Bollobás & W. Fernandez de la Vega, The diameter of random regular graphs, *Combinatorica*, **2** (1982), 125–134.

[22] J. A. Bondy & U. S. R. Murty, *Graph Theory with Applications*, Elsevier, New York (1976).

[23] O. Borodin & A. Kostochka, On an upper bound of a graph's chromatic number depending on the graph's degree and density, *Journal of Combinatorial Theory, Series B*, **23** (1977), 247–250.

[24] A. Broder & E. Shamir, On the second eigenvalue of random regular graphs, in *28th Annual Symposium on Foundations of Computer Science.* IEEE Comput. Soc. Press, Washington DC (1987), pp. 286–294.

[25] P. Catlin, A bound on the chromatic number of a graph, *Discrete Mathematics*, **22** (1978), 81–83.

[26] F. R. K. Chung, Constructing random-like graphs, in *Probabilistic Combinatorics and Its Applications* (ed. B. Bollobás), *Proc. Sympos. Appl. Math.*, 44, Amer. Math. Soc., Providence, Rhode Island (1991), pp. 21–55.

[27] K. L. Chung, *A Course in Probability Theory, 2nd ed.*, Academic Press, New York (1974).

[28] C. Cooper, A note on the connectivity of 2-regular digraphs, *Random Structures & Algorithms*, **4** (1993), 469–472.

[29] C. Cooper, On the 2-cyclic property in 2-regular digraphs, *Random Structures & Algorithms*, **6** (1995), 439–448.

[30] C. Cooper & A. M. Frieze, Hamilton cycles in a class of random directed graphs, *Journal of Combinatorial Theory, Series B*, **62** (1994), 151–163.

[31] C. Cooper, A. Frieze & M. Molloy, Hamilton cycles in random regular digraphs, *Combinatorics, Probability and Computing*, **3** (1994), 39–50.

[32] C. Cooper, A. Frieze, M. Molloy & B. Reed, Perfect matchings in random r-regular, s-uniform hypergraphs, *Combinatorics, Probability and Computing*, **5** (1996), 1–14.

[33] R. Diestel, *Graph Theory*, Springer, New York (1997).

[34] W. Duckworth & N. C. Wormald, Maximum star forests and minimum dominating sets in random cubic graphs, in preparation.

[35] M. N. Ellingham, The asymptotic connectivity of labelled coloured regular bipartite graphs, in *Combinatorial Mathematics X.* (ed. L. R. A. Casse), *Springer Lecture Notes in Mathematics*, 1036, Springer, Berlin (1983), pp. 177–188.

[36] P. Erdős & A. Rényi, On the strength of connectedness of a random graph, *Acta Mathematica Academiae Scientiarum Hungaricae*, **12** (1961), 261–267.

[37] T. I. Fenner & A. M. Frieze, On the existence of hamiltonian cycles in a class of random graphs, *Discrete Mathematics*, **45** (1983), 301–305.

[38] T. I. Fenner & A. M. Frieze, Hamiltonian cycles in random regular graphs, *Journal of Combinatorial Theory, Series B*, **37** (1984), 103–112.

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

[40] A. M. Frieze, Maximum matchings in a class of random graphs, *Journal of Combinatorial Theory, Series B*, **40** (1986), 196–212.

[41] A. M. Frieze, Finding Hamilton cycles in sparse random graphs, *Journal of Combinatorial Theory, Series B*, **44** (1988), 230–250.

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

[43] A. Frieze, M. Karoński & L. Thoma, On perfect matchings and Hamilton cycles in sums of random trees, *SIAM Journal on Discrete Mathematics*, in press.

[44] A. M. Frieze & T. Łuczak, On the independence numbers of random regular graphs, *Journal of Combinatorial Theory, Series B*, **54** (1992), 123–132.

[45] A. M. Frieze & S. Suen, On the independence number of random cubic graphs, *Random Structures & Algorithms*, **5** (1994), 649–664.

[46] Z. Gao, B. D. McKay & X. Wang, On the structure of a random tournament with a given score sequence, (submitted).

[47] Z. Gao & N. C. Wormald, The size of the largest components in random planar maps, *SIAM Journal on Discrete Mathematics*, in press.

[48] H. Garmo, Random railways modeled as random 3-regular graphs, *Random Structures & Algorithms*, **9** (1996), 113–136.

[49] H. Garmo, Asymptotic properties of the connectivity number of random railways, *Advances in Applied Probability*, **31** (1999), in press.

[50] H. Garmo, Random railways and cycles in random regular graphs, Doctoral thesis, Uppsala University, 1998.

[51] C. D. Godsil & B. D. McKay, Asymptotic enumeration of Latin rectangles, *Journal of Combinatorial Theory, Series B*, **48** (1990), 19–44.

[52] A. Goerdt, The giant component threshold for random regular graphs with edge faults, in *Mathematical foundations of computer science 1997* (eds. I. Prívara & P. Ružička), *Lecture Notes in Comput. Sci.*, 1295, Springer, Berlin (1997), pp. 279–288.

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

[54] F. Harary & E. M. Palmer, *Graphical Enumeration*, Academic Press, New York (1973).

[55] S. Janson, private communication, 1997.

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

[57] S. Janson & N. C. Wormald, Rainbow Hamilton cycles in random coloured regular graphs, in preparation.

[58] M. R. Jerrum & A. J. Sinclair, Fast uniform generation of regular graphs, *Theoretical Computer Science*, **73** (1990), 91–100.

[59] M. R. Jerrum, B. D. McKay & A. J. Sinclair, When is a graphical sequence stable?, in *Random Graphs Vol. 2* (eds. A. Frieze & T. Łuczak), Wiley, New York (1992), pp. 101–115.

[60] J. H. Kim & N. C. Wormald, Random matchings which induce Hamilton cycles, and hamiltonian decompositions of random regular graphs, in preparation.

[61] J. Komlós & E. Szemerédi, Limit distributions for the existence of Hamilton circuits in a random graph, *Discrete Mathematics*, **43** (1983), 55–63.

[62] J. Lawrence, Covering the vertex set of a graph with subgraphs of smaller degree, *Discrete Mathematics*, **21** (1978), 61–68.

[63] T. Łuczak, Sparse random graphs with a given degree sequence, in *Random Graphs Vol. 2* (eds. A. Frieze & T. Łuczak), Wiley, New York (1992), pp. 165–182.

[64] T. Łuczak & N. C. Wormald, Phase transition for random graph processes, in preparation.

[65] C. McDiarmid, On the method of bounded differences, in *Surveys in Combinatorics 1989* (ed. J. Siemons), *London Mathematical Society Lecture Note Series*, 141, Cambridge University Press, Cambridge (1989), pp. 148-188.

[66] C. McDiarmid & B. Reed, Linear arboricity of random regular graphs, *Random Structures & Algorithms*, **1** (1990), 443–445.

[67] B. D. McKay, The expected eigenvalue distribution of a large regular graph, *Linear Algebra and its Applications*, **40** (1981), 203–216.

[68] B. D. McKay, Spanning trees in random regular graphs, in *Proceedings of the Third Caribbean Conference on Combinatorics and Computing (Bridgetown, 1981)* (ed. C. C. Cadogan), Univ. West Indies, Cave Hill Campus, Barbados (1981), pp. 139–143.

[69] B. D. McKay, Subgraphs of random graphs with specified degrees, *Congressus Numerantium*, **33** (1981), 213–223.

[70] B. D. McKay, Spanning trees in regular graphs, *European Journal of Combinatorics*, **4** (1983), 149–160.

[71] B. D. McKay, Asymptotics for 0-1 matrices with prescribed line sums, in *Enumeration and Design* (eds. D. M. Jackson & S. A. Vanstone), Academic Press, Toronto (1984), 225–238.

[72] B. D. McKay, Asymptotics for symmetric 0-1 matrices with prescribed row sums, *Ars Combinatoria*, **19A** (1985), 15–25.

[73] B. D. McKay, Independent sets in regular graphs of high girth, in *Proc. Australia-Singapore Joint Conference on Information Processing and Combinatorial Mathematics (Singapore, 1986). Ars Combinatoria*, 23A, (1987), pp. 179–185.

[74] B. D. McKay, The asymptotic numbers of regular tournaments, eulerian digraphs and eulerian oriented graphs, *Combinatorica*, **10** (1990), 367–377.

[75] B. D. McKay, E. M. Palmer, R. C. Read & R. W. Robinson, The asymptotic number of claw-free cubic graphs, preprint, Australian National University, 1998.

[76] B. D. McKay & X. Wang, Asymptotic enumeration of bipartite graphs by degree sequence, in preparation.

[77] B. D. McKay & X. Wang, Asymptotic enumeration of tournaments with a given score sequence, *Journal of Combinatorial Theory, Series A*, **73** (1996), 77–90.

[78] B. D. McKay & N. C. Wormald, Automorphisms of random graphs with specified degrees, *Combinatorica*, **4** (1984), 325–338.

[79] B. D. McKay & N. C. Wormald, Uniform generation of random regular graphs of moderate degree, *Journal of Algorithms*, **11** (1990), 52–67.

[80] B. D. McKay & N. C. Wormald, Asymptotic enumeration by degree sequence of graphs of high degree, *European Journal of Combinatorics*, **11** (1990), 565–580.

[81] B. D. McKay & N. C. Wormald, Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$, *Combinatorica*, **11** (1991), 369–382.

[82] B. D. McKay & N. C. Wormald, The degree sequence of a random graph. I. The models, *Random Structures & Algorithms*, **11** (1997), 97–117.

[83] B. D. McKay, N. C. Wormald & B. Wysocka, Asymptotic enumeration of regular graphs of high girth, in preparation.

[84] M. Molloy, The chromatic number of sparse random graphs, M.Math. thesis, University of Waterloo, 1992.

[85] M. Molloy & B. Reed, A critical point for random graphs with a given degree sequence, *Random Structures & Algorithms*, **6** (1995), 161–179.

[86] M. Molloy & B. Reed, The dominating number of a random cubic graph, *Random Structures & Algorithms*, **7** (1995), 209–221.

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

[88] S. E. Nikoletseas & P. G. Spirakis, Expander properties in random regular graphs with edge faults, in *STACS 95* (eds. E. W. Mayr & C. Puech), *Lecture Notes in Comput. Sci.*, 900, Springer, Berlin (1995), pp. 421–432.

[89] P. E. O'Neil, Asymptotics and random matrices with row-sum and column sum-restrictions, *Bulletin of the American Mathematical Society*, **75** (1969), 1276–1282.

[90] L. Pósa, Hamiltonian circuits in random graphs, *Discrete Mathematics*, **14** (1976), 359–364.

[91] R. C. Read, The enumeration of locally restricted graphs (II), *Journal of the London Mathematical Society*, **35** (1960), 344–351.

[92] L. B. Richmond, R. W. Robinson & N. C. Wormald, On Hamilton cycles in 3-connected cubic maps, in *Cycles in Graphs* (eds. B. Alspach & C. D. Godsil), *Annals of Discrete Mathematics*, 27, North-Holland, Amsterdam (1985), pp. 141–150.

[93] L. B. Richmond & N. C. Wormald, Random triangulations of the plane, *European Journal of Combinatorics*, **9** (1988), 61–71.

[94] L. B. Richmond & N. C. Wormald, Almost all maps are asymmetric, *Journal of Combinatorial Theory, Series B*, **63** (1995), 1–7.

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

[96] H. Robalewska, Random graphs generated by the star 2-process, *Random Structures & Algorithms*, **11** (1997), 125–149.

[97] H. Robalewska & N. C. Wormald, Random star processes, preprint, University of Melbourne, 1997.

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

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

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

[101] R. W. Robinson & N. C. Wormald, Hamilton cycles containing randomly selected edges in random cubic graphs, in preparation.

[102] A. Ruciński & N. C. Wormald, Random graph processes with degree restrictions, *Combinatorics, Probability and Computing*, **1** (1992), 169–180.

[103] A. Ruciński & N. C. Wormald, Random graph processes with maximum degree 2, *Annals of Applied Probability*, **7** (1997), 183–199.

[104] A. Ruciński & N.C. Wormald, The probability that a random d-process results in a connected graph, preprint, University of Melbourne, 1997.

[105] J. B. Shearer, A note on the independence number of triangle-free graphs, *Journal of Combinatorial Theory, Series B*, **53** (1991), 300–307.

[106] J. H. Spencer, Random regular tournaments, *Periodica Mathematica Hungarica*, **5** (1974), 105–120.

[107] A. Steger & N. C. Wormald, A fast algorithm for generating random regular graphs, preprint, University of Melbourne, 1998.

[108] A. Telcs & N. C. Wormald, Hamiltonicity of random 2-processes, in preparation.

[109] G. Tinhofer, On the generation of random graphs with given properties and known distribution, *Applied Computer Science, Berichte zur Praktischen Informatik*, **13** (1979), 265–297.

[110] W. T. Tutte, On the enumeration of convex polyhedra, *Journal of Combinatorial Theory, Series B*, **28** (1980), 105–126.

[111] D. W. Walkup, Matchings in random regular bipartite digraphs, *Discrete Mathematics*, **31** (1980), 59–64.

[112] H. S. Wilf, *generatingfunctionology*, Academic Press, San Diego (1990).

[113] J. B. Wilson, Methods for detecting non-randomness in species co-occurrences: a contribution, *Oecologia*, **73** (1987), 579–582.

[114] N. C. Wormald, Triangles in labelled cubic graphs, in *Combinatorial Mathematics* (eds. D. A. Holton & J. Seberry), *Springer Lecture Notes in Mathematics*, 686, Springer, Berlin (1978), pp. 337–343.

[115] N. C. Wormald, Some Problems in the Enumeration of Labelled Graphs, PhD thesis, University of Newcastle, 1978.

[116] N. C. Wormald, Enumeration of labelled graphs II: Cubic graphs with a given connectivity, *Journal of the London Mathematical Society (2)*, **20** (1979), 1–7.

[117] N. C. Wormald, The asymptotic connectivity of labelled regular graphs, *Journal of Combinatorial Theory, Series B*, **31** 156–167. (1981),

[118] N. C. Wormald, The asymptotic distribution of short cycles in random regular graphs, *Journal of Combinatorial Theory, Series B*, **31** 168–182. (1981),

[119] N. C. Wormald, Generating random regular graphs, *Journal of Algorithms*, **5** (1984), 247–280.

[120] N. C. Wormald, A simpler proof of the asymptotic formula for the number of unlabelled r-regular graphs, *Indian Journal of Mathematics*, **28** (1986), 43–47.

[121] N. C. Wormald, Generating random unlabelled graphs, *SIAM Journal on Computing*, **16** (1987), 717–727.

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

[123] N. C. Wormald, The perturbation method and triangle-free random graphs, *Random Structures & Algorithms*, **9** (1996), 253–270.

[124] N. C. Wormald, The differential equation method for random graph processes and greedy algorithms, in *Lectures on Approximation and Randomized Algorithms* (eds. M. Karoński & H. J. Prömel), PWN, Warsaw, in press.

Department of Mathematics and Statistics
University of Melbourne
Parkville VIC 3052
Australia
nick@ms.unimelb.edu.au