Bounds on the max and min bisection of random cubic and random 4-regular graphs ⋆

J. Díaz [a,1], N. Do [b], M.J. Serna [a,1], N.C. Wormald [b,1,2,3]

[a] *Dept. Llenguatges i Sistemes, Universitat Politecnica de Catalunya, Jordi Girona Salgado 1–3, 08034 Barcelona, Spain*

[b] *Department of Mathematics and Statistics, University of Melbourne, VIC 3010, Australia*

Abstract

In this paper we present a randomized algorithm to compute the bisection width of cubic and 4-regular graphs. The analysis of the proposed algorithms on random graphs provides asymptotic upper bounds for the bisection width of random cubic and random 4-regular graphs with n vertices, giving upper bounds of $0.174039n$ for random cubic, and of $0.333333n$ for random 4-regular. We also obtain asymptotic lower bounds for the size of the maximum bisection, for random cubic and random 4-regular graphs with n vertices, of $1.32697n$ and $1.66667n$, respectively. The randomized algorithms are derived from initial greedy algorithm and their analysis is based on the differential equation method.

Key words: bisection width, random cubic graphs, random 4-regular graphs.

1 Introduction

Given a graph $G = (V, E)$ with $|V|$ even, a *bisection* of G is a partition of V into two parts with the same cardinality, and its *size* is the number of edges crossing between the parts. A *minimum bisection* is a bisection of V with minimal size (to avoid parity problems, throughout this paper we assume that

⋆ Part of the results in this paper were presented at RANDOM-2002 [6]

[1] Supported by the IST programme of the EU under contract IST-1999-14186 (ALCOM-FT)

[2] Supported by the Australian Research Council and the CRM (Bellaterra)

[3] Present address: Dept. of Comb. and Optimization U. of Waterloo, Canada.

$n = |V|$ is even). The size of a minimum bisection is called the *bisection width* and the *min bisection problem* consists of finding a minimum bisection in a given G. In the same manner, we can also consider a *maximum bisection*, i.e. a bisection that maximizes the number of crossing edges. A related problem is that of finding the largest bipartite subgraph of a graph, i.e. a bipartite subgraph with as many edges as possible. This problem is known as the *max cut problem* (see for example [12]). Given a graph, the size of a maximum bisection is clearly a lower bound on the size of a max cut in the graph.

The min bisection problem has received a lot of attention, as the bisection width plays an important role in finding lower bounds to the routing performance of a network. The decisional version of the problem is known to be NP-complete [12], even for cubic graphs [5]. In [1] it is shown that min bisection has a PTAS for everywhere dense graphs (graphs with minimum degree $\Omega(n)$). Moreover, there exists an $O(\log^2 n)$ approximation for the min bisection on general graphs and an $O(\log n)$ approximation for planar graphs [8]. The min bisection problem can be solved in polynomial time for bounded treewidth graphs [13]. Several other exact and heuristic positive results are known for particular cases of the problem (see for example [7]).

With respect to lower bounds, well known is the spectral lower bound of $\lambda_2 n/4$ for the bisection width of any graph, where λ_2 is the second eigenvalue of the Laplacian of the graph [11]. Bollobas provided a lower bound of $(d - 2\sqrt{d \ln 2})n/4$, for almost all d-regular graphs [3]. In the same paper, he gave a lower bound of $0.22n$ for the particular case of 4-regular graphs. In [15] it is shown that almost all cubic graphs have bisection width greater than $0.101n$. Recently, using spectral techniques, Bezrukov et al. have given lower bounds for of $0.082n$ for the bisection width of cubic Ramanujan graphs, and of $0.176n$ for the case of 4-regular Ramanujan graphs [2].

Recently, Monien and Preis [17] gave upper bounds on the bisection width of $(1/6 + \epsilon)n$ for 3-regular graphs and of $(0.4 + \epsilon)n$ for 4-regular graphs, for any $\epsilon > 0$ (n sufficiently large, depending on ϵ).

The problem of finding the maximum bisection has also received a lot of attention. This is again NP-hard even for planar graphs [13]. It is known to be solvable in polynomial time for bounded treewidth graphs [13]. There exists several approximations algorithms for the problem; the max bisection problem has a PTAS for planar graphs [14]. In the case of regular graphs, there is a 0.795 approximation algorithm for the max bisection [10]. In the same paper, the authors gave an 0.834 approximation algorithm for the special case of max bisection on cubic graphs. The best approximation algorithm for max bisection on general graphs has an approximation ratio of 0.7027 [9]. We are not aware of any non-trivial lower bounds on the size of the maximum bisection.

We will use standard notation and we refer the reader to [16], for the definitions of u.a.r. (uniformly at random) and a.a.s. (asymptotically almost surely).

In this paper we present asymptotic results for the typical bisection width and the expected maximum bisection of random cubic and random 4-regular graphs, where our graphs have no loops or multiple edges. In particular we prove the following theorems.

Theorem 1 *For all $\epsilon > 0$, the bisection width of a random 4-regular graph on n vertices is a.a.s. smaller than $n/3 + \epsilon n$.*

This upper bound improves the previously best known upper bound of $0.4n$ which follows from the upper bound on the bisection width of *all* 4-regular graphs by Monien and Preis. Recall that Bollobas' lower bound for random 4-regular graphs is $0.22n$.

Theorem 1 is proved by setting ϵ sufficiently small in the simple randomized algorithm in Figure 3. Doing the same ($\epsilon < 10^{-6}$) for the algorithm in Figure 4 gives the following.

Theorem 2 *The algorithm in Figure 4 a.a.s. finds a bisection of width at most $0.17404n$ in a random cubic graph on n vertices.*

This asymptotic upper bound is close to but weaker than the asymptotic upper bound of $1/6 + \epsilon$ (any $\epsilon > 0$) for all cubic graphs, by Monien and Preis [17]. We give this result here for two reasons: our algorithm is much simpler, and our method also gives Theorem 3 below.

Slightly modifying the proposed algorithms and using the same analysis as for the previous results, we can get the following lower bounds for the maximum bisection of random cubic and random 4-regular graphs.

Theorem 3 *The maximum bisection of a random cubic graph with n vertices is a.a.s. greater than $1.32595n$.*

Theorem 4 *For all $\epsilon > 0$, the maximum bisection of a random 4-regular graph with n vertices is a.a.s. greater than $5n/3 + \epsilon n$.*

As mentioned above, the size of the maximum bisection is a trivial lower bound on the size of the maximum cut, but removing the balance constraint does not permit our method to obtain any better result. We conjecture that the largest balanced bipartite subgraph of a random d-regular graph is a.a.s. almost the same size as the largest bipartite subgraph. For the particular example of random cubic graphs, we can state this even more strongly and precisely, as follows.

Conjecture 1 *For every $\epsilon > 0$, a.a.s. the largest bipartite subgraph of a random cubic graph has a 2-colouring with the difference in the numbers of vertices of the two colours less than ϵn.*

The techniques in the present paper should extend in some fashion to regular graphs of higher degree. However, it then becomes unclear even what algorithms should be used, so we do not pursue this question here.

2 Greedy algorithms for the minimum bisection of random cubic and 4-regular graphs

In this section we give the basic randomized greedy procedures to find a bisection for random cubic and random 4-regular graphs. We also introduce some notation to be used in the analysis.

Given a graph, and given a partial assignment of colours red (R) and blue (B) to its vertices, we classify the non-coloured vertices according with the number of their coloured neighbours:

A vertex is of **Type** (r, b) if it has r neighbours coloured R and b neighbours coloured B.

We say that a pair of uncoloured vertices is (r, b)-*symmetric* if their types are (r, b) and (b, r).

The greedy procedures work by colouring vertices chosen randomly in symmetric pairs, to maintain balance, and repeatedly use one of the following two operations: the *majority operation* (Maj), that colours each vertex with the majority colour among its neighbours, and the *random operation* (Rand) that randomly colours one vertex R and the other B.

The greedy procedure 4-min greedy for 4-regular graphs is given in Figure 1, while the greedy procedure 3-min greedy for cubic graphs is given in Figure 2.

A major difference between the two algorithms is that while the algorithm for 4-regular graphs consists of only one phase (followed by a "cleaning up" operation), for cubic graphs the algorithm consists of three phases. This fact makes the analysis of the 4-regular case simpler, so it is presented here first.

The algorithm 4-min greedy considers only (0,1)-, (0,2)- and (1,2)-symmetric pairs of uncoloured vertices and gives higher priority to the (0,2)- and (1,2)-symmetric pairs than to the (0,1)-symmetric pairs. Note that the size of the bisection is the number of *bicoloured* edges, with one vertex of each colour, so

Initial step: select two non-adjacent vertices u.a.r., colour one with R and the other with B

Phase 1: **repeat**

if there are vertices of both types (2,0) and (0,2)
or vertices of types (2,1) and (1,2)
then
select u.a.r. a (0,2)- or (1,2)-symmetric pair
and perform Maj;
else if there are vertices of both types (1,0) and (0,1)
then
select u.a.r. a (0,1)-symmetric pair and perform Maj;

until no new vertex is coloured

Cleanup: colour any remaining uncoloured vertices, half of them R and half B, in any manner, and output the bisection R, B.

Fig. 1. Algorithm 4-min greedy for obtaining a bisection of 4-regular graphs

only each Maj operation on a (1,2)-symmetric pair contributes, with 2, to the bisection.

On the other hand, each phase of the algorithm 3-min greedy considers two types of symmetric pairs and gives priority to one of them. Observe that in the first phase there is no contribution to the bisection, while in the second and third phases, every time a (1,1)- or (1,2)-symmetric pair is coloured, the bisection is increased by 2.

One method of analyzing the performance of a randomized algorithm is to use a system of differential equations to express the expected changes in the variables describing the state of the algorithm during its execution. An exposition of this method can be found in [18], which includes various examples of graph-theoretic optimization problems. For purposes of exposition, we continue for the present to discuss the proposed algorithms, without giving full justification. After this, in order to reduce the complexity of the justification, it is in fact a different but related algorithm which we will analyse to yield our claimed bounds. We call this variation of algorithm a *deprioritized* algorithm as in [20], where this technique was first used.

We use the pairing model to generate n-vertex d-regular graphs u.a.r. Briefly, to generate such a random graph, it is enough to begin with dn *points* in n *cells*, and choose a random perfect matching of the points, which we call a *pairing*. The corresponding pseudograph (possibly with loops or multiple edges) has the cells as vertices and the pairs as edges. Since d is fixed, any property a.a.s. true of the random pseudograph is also a.a.s. true of the restriction to random graphs, with no loops or multiple edges, and this restricted probability space is uniform (see for example [4,19] for a full description). Without loss of generality, when stating such asymptotic results, we restrict n to being even to avoid parity problems.

Initial step: select two non-adjacent vertices u.a.r., colour one with R and the other with B

Phase 1: **repeat**
  if there are vertices of both types (2,0) and (0,2)
  then
    select u.a.r. a (0,2)-symmetric pair and perform Maj;
  else if there are vertices of both types (1,0) and (0,1)
    then
      select u.a.r. a (0,1)-symmetric pair and perform Maj;
until no new vertex is coloured

Phase 2: **repeat**
  if there are vertices of both types (1,0) and (0,1)
  then
    select u.a.r. a (0,1)-symmetric pair and perform Maj;
  else if there are at least two vertices of type (1,1)
    then
      select u.a.r. a (1,1)-symmetric pair and perform Rand;
until no new vertex is coloured

Phase 3: **repeat**
  if there are vertices of both types (3,0) and (0,3)
  then
    select u.a.r. a (0,3)-symmetric pair and perform Maj;
  else if there are vertices of both types (2,1) and (1,2)
    then
      select u.a.r. a (1,2)-symmetric pair and perform Maj;
until no new vertex is coloured

Cleanup: colour any remaining uncoloured vertices, half of them R and half B, in any manner, and output the bisection R, B.

Fig. 2. Algorithm 3-min greedy for obtaining a bisection of cubic graphs

We consider the greedy algorithms applied directly to the random pairing. As discussed in [18], the random pairing can be generated pair by pair, and at each step a point p can be chosen by any rule whatsoever, as long as the other point in the pair is chosen u.a.r. from the remaining unused points. We call this step *exposing* the pair containing p.

At each point in the algorithm, let Z_{rb} represent the number of uncoloured vertices of type (r, b), and let W denote the number of points not yet involved in exposed pairs. It follows that, for d-regular graphs,

$$W = \sum_{r+b \leq d} (d - r - b) Z_{rb}.$$

To analyse the algorithm, when a vertex is coloured we immediately expose all pairs involved in that vertex. In this way, the numbers Z_{rb} are always determined. Furthermore, W points are available for the pairs that will be exposed during the next step.

3 Analysis of an algorithm for random 4-regular graphs

When considering algorithm 4-min greedy run on a random pairing, at any time step, the number of points not yet involved in exposed pairs is

$$W = 4Z_{00} + 3Z_{10} + 3Z_{01} + 2Z_{02} + 2Z_{20} + 2Z_{11} + Z_{12} + Z_{21} + Z_{03} + Z_{30}.$$

Consider what happens when a vertex u is newly coloured R and one of the pairs containing a point p in that cell is exposed. The other point will lie in some vertex v. The probability that v has type (i, j) will be $(4-i-j)Z_{ij}/(W-1)$ (except for a correction due to the change in status of u). Let d_{rb} denote the expected contribution to the increment $\Delta(Z_{rb})$ in Z_{rb} due to the change in the status of v. Up to terms $O(1/W)$, this contribution is gains $(4-i-j)Z_{ij}/W$ from the case $(i, j) = (r-1, b)$, and $(4-(i+j))Z_{ij}/(W-1)$ from the case $(i, j) = (r, b)$. The error term $O(1/W)$ is due to the replacement of $W-1$ by W and an adjustment occurring when v happens to be the same as u. This gives (ignoring $O(1/W)$ terms)

$$\begin{array}{llll}
d_{00} = -\dfrac{4Z_{00}}{W} & d_{01} = -\dfrac{3Z_{01}}{W} & d_{02} = -\dfrac{2Z_{02}}{W} & d_{03} = -\dfrac{Z_{03}}{W}\ d_{04} = 0 \\
d_{10} = \dfrac{4Z_{00} - 3Z_{10}}{W} & d_{11} = \dfrac{3Z_{01} - 2Z_{11}}{W} & d_{12} = \dfrac{2Z_{02} - Z_{12}}{W} & d_{13} = \dfrac{Z_{03}}{W} \\
d_{20} = \dfrac{3Z_{10} - 2Z_{20}}{W} & d_{21} = \dfrac{2Z_{11} - Z_{21}}{W} & d_{22} = \dfrac{Z_{12}}{W} & \\
d_{30} = \dfrac{2Z_{20} - Z_{30}}{W} & d_{31} = \dfrac{Z_{21}}{W} & & \\
d_{40} = \dfrac{Z_{30}}{W}. & & &
\end{array}$$

The corresponding equations when a vertex is coloured B form a symmetric set with these: they are the same but with the index pair on all variables swapped. Therefore, denoting by $\bar{d}_{r,b}$ the expected increments due to a dual step, consisting of colouring two vertices R and B, we get (again ignoring $O(1/W)$ terms)

$$
\begin{aligned}
\bar{d}_{00} &= -\frac{8Z_{00}}{W} \\
\bar{d}_{01} &= \frac{-6Z_{01} + 4Z_{00}}{W} \; \bar{d}_{11} = \frac{3Z_{01} + 3Z_{10} - 4Z_{11}}{W} \\
\bar{d}_{02} &= \frac{-4Z_{02} + 3Z_{01}}{W} \; \bar{d}_{12} = \frac{2Z_{02} - 2Z_{12} + 2Z_{11}}{W} \; \bar{d}_{22} = \frac{Z_{12} + Z_{21}}{W} \\
\bar{d}_{03} &= \frac{-2Z_{03} + 2Z_{02}}{W} \; \bar{d}_{13} = \frac{Z_{03} + Z_{12}}{W} \\
\bar{d}_{04} &= \frac{Z_{03}}{W}
\end{aligned} \tag{1}
$$

and symmetric equations for $\bar{d}_{rb}$ when $r > b$. We now make the assumption of having *rb-symmetry*: for all i and j, $Z_{ij} = Z_{ji}$. Later we will see how to remove it, but when it holds, the values of $\bar{d}_{11}$ and $\bar{d}_{22}$ can be simplified, and the equations are:

$$
\begin{aligned}
\bar{d}_{00} &= -\frac{8Z_{00}}{W} \\
\bar{d}_{01} &= \frac{-6Z_{01} + 4Z_{00}}{W} \; \bar{d}_{11} = \frac{6Z_{01} - 4Z_{11}}{W} \\
\bar{d}_{02} &= \frac{-4Z_{02} + 3Z_{01}}{W} \; \bar{d}_{12} = \frac{2Z_{02} - 2Z_{12} + 2Z_{11}}{W} \; \bar{d}_{22} = \frac{2Z_{12}}{W} \\
\bar{d}_{03} &= \frac{-2Z_{03} + 2Z_{02}}{W} \; \bar{d}_{13} = \frac{Z_{03} + Z_{12}}{W} \\
\bar{d}_{04} &= \frac{Z_{03}}{W}.
\end{aligned} \tag{2}
$$

The rest of our discussion, until considering the deprioritized algorithm, is nonrigorous, mainly for motivation, but also including the derivations of some formulae used later. The difficulty of analysis is caused by the prioritization. To proceed, define ϕ_1 to be the probability of processing a $(0,1)$-symmetric pair, let ϕ_2 be the probability of processing a $(0,2)$-symmetric pair, and let ϕ_3 be the probability of processing a $(1,2)$-symmetric pair, at a given step in the algorithm. Then immediately

$$
\phi_1 + \phi_2 + \phi_3 = 1. \tag{3}
$$

Moreover, every dual colouring of a $(0,1)$-symmetric pair produces in expectation $3\bar{d}_{02}$ $(0,2)$-symmetric pairs and $3\bar{d}_{12}$ $(1,2)$-symmetric pairs. Every dual colouring of a $(0,2)$-symmetric pair produces $2\bar{d}_{02}$ $(0,2)$-symmetric pairs and $2\bar{d}_{12}$ $(1,2)$-symmetric pairs; and every dual colouring of $(1,2)$-symmetric pair produces $\bar{d}_{02}$ $(0,2)$-symmetric pairs and $\bar{d}_{12}$ $(1,2)$-symmetric pairs. Therefore,

the expected number of $(0,2)$-symmetric pairs produced in a given step is

$$(3\phi_1 + 2\phi_2 + \phi_3)\bar{d}_{02}, \tag{4}$$

and the expected number of $(1,2)$-symmetric pairs is

$$(3\phi_1 + 2\phi_2 + \phi_3)\bar{d}_{12}. \tag{5}$$

In a large number of consecutive steps, the prioritization ensures that virtually all of the $(0,2)$- and $(1,2)$-symmetric pairs are used up (unless ϕ_1 has reached 0), implying that ϕ_2 and ϕ_3 should be equated to the expressions in (4) and (5) respectively. Solving these together with (3), we get

$$\phi_1 = \frac{1 - 2\bar{d}_{02} - \bar{d}_{12}}{1 + \bar{d}_{02} + 2\bar{d}_{12}}; \quad \phi_2 = \frac{3\bar{d}_{02}}{1 + \bar{d}_{02} + 2\bar{d}_{12}}; \quad \phi_3 = \frac{3\bar{d}_{12}}{1 + \bar{d}_{02} + 2\bar{d}_{12}}. \tag{6}$$

Phase 1 will finish when $Z_{00} = Z_{02} = Z_{12} = 0$. Continuing our non-rigorous computation, we can find the expected increments of the random variables Z_{ij} in each iteration in phase 1 (assuming rb-symmetry). Using linearity of expectation,

$$\mathbf{E}\left[\Delta(Z_{ij})\right] = \bar{d}_{ij}(\phi_3 + 2\phi_2 + 3\phi_1) - \delta_{01}\phi_1 - \delta_{02}\phi_2 - \delta_{12}\phi_3 \tag{7}$$

for any i, j with $i \leq j$ and $i + j \leq 3$, where $\delta_{pq} = 1$ if $(p,q) = (i,j)$, and 0 otherwise.

We may express the above expected increments as a set of differential equations, where each $\mathbf{E}\left[\Delta(Z_{ij})\right]$ is expressed as the differential Z'_{ij} (as a function of the number t of iterations). We scale both time, t, and the variables by dividing by n, and denote Z_{ij}/n by z_{ij}, t/n by x and $W/n = W(t)/n$ by $w = w(x)$. This yields

$$\begin{aligned} z'_{00} &= -8z_{00}\frac{D}{w}, & z'_{01} &= (4z_{00} - 6z_{01})\frac{D}{w} - \theta_1, \\ z'_{11} &= (6z_{01} - 4z_{11})\frac{D}{w}, \end{aligned} \tag{8}$$

where

$$\begin{aligned} D &= D(x) = \theta_3 + 2\theta_2 + 3\theta_1, \\ w &= w(x) = 4z_{00} + 6z_{01} + 4z_{02} + 2z_{11} + 2z_{12} + 2z_{03} \end{aligned} \tag{9}$$

and $\theta_i = \theta_i(x)$, representing $\phi(t/n)$, is defined as ϕ_i in (6) but with Z_{ij} replaced by $z_{ij}(x)$ in the definition of the variables $\bar{d}$. For instance,

$$\theta_2 = \frac{3(-4z_{02} + 3z_{01})}{w + (-4z_{02} + 3z_{01}) + 2(2z_{02} - 2z_{12} + 2z_{11})}.$$

As long as θ_1 remains positive (which, as we see later, is the case in phase 1), it will follow that the number of vertices of types $(0,2)$, $(2,0)$, $(1,2)$ and $(2,1)$ remain small and that there is regularly no symmetric pair of either of these types (since it is the only time that $(0,1)$-symmetric pairs are processed). This implies that a negligible number of vertices of types $(1,3)$ or $(2,2)$ are ever created. It follows that the z_{ij} whose derivatives are not included in (8) should remain constant at 0, and so for the present discussion we write

$$z_{ij} \equiv 0, \quad j \geq 2. \tag{10}$$

It also follows that the size of the bisection in phase 1 is approximately equal to the twice the total number of $(1,2)$-symmetric pairs of vertices which are processed (since colouring a $(0,2)$ or $(0,1)$ vertex does not add to the bisection). Letting $Y(t)$ be a random variable keeping track of the number of times a $(1,2)$-symmetric pair is processed, we have the expected change in Y in one step equal to $\bar{d}_{12}$. Let $y(x)$ represent $Y(t)/n$. Then the suggested equation for y is

$$y'(x) = z_{11}\frac{D(x)}{w(x)}. \tag{11}$$

Solving this together with (8), (9) and (10) with the initial conditions

$$z_{00}(0) = 1;\ z_{ij}(0) = 0 \text{ for } 0 < i \leq j,\ i+j \leq 4;\ y(0) = 0, \tag{12}$$

is equivalent to solving

$$z_{00}' = \frac{-24z_{00}}{R}, \quad z_{01}' = \frac{8z_{00} - 18z_{01}}{R} \quad z_{11}' = \frac{18z_{01} - 12z_{11}}{R}, \quad y' = \frac{12y}{R} \tag{13}$$

where $R = 4z_{00} + 9z_{01} + 6z_{11}$, with the same initial conditions. We do not know the explicit solution to these equations, but we can deduce all that we need as follows.

Let x_1 be the infimum of those $x > 0$ for which either $\theta_1 = 0$ or $z_{ij}(x) = 0$ for $ij = 00$, 01 or 11. From (13) we have

$$(z_{00} + 2z_{01} + z_{11})' = -2$$

and hence

$$z_{00} + 2z_{01} + z_{11} = 1 - 2x. \tag{14}$$

Note that for small $x > 0$, all the variables z_{00}, z_{01} and z_{11} must be positive. Can any be positive at x_1? All are nonnegative, so if one of them is strictly positive, this implies $R(x_1) > 0$. But then from (13) it follows that $z_{00}(x_1) > 0$,

and from this $z_{01}(x_1) > 0$ and $z_{11}(x_1) > 0$ in turn. We also find that

$$\theta_1 = \frac{4z_{00}}{R}, \qquad \theta_2 = \frac{9z_{01}}{R}, \qquad \theta_3 = \frac{6z_{11}}{R}. \tag{15}$$

Thus θ_1 is then positive, and this contradicts the definition of x_1. Hence

$$z_{00}(x_1) = z_{01}(x_1) = z_{11}(x_1) = 0 \tag{16}$$

and hence R and θ_1 are 0 as well at x_1. So from (13),

$$x_1 = \frac{1}{2}. \tag{17}$$

Furthermore, from (13),

$$R' = \frac{-24z_{00} - 54z_{01} - 72z_{11}}{R} = -6 - \frac{32z_{11}}{R} = -6 - 6y'$$

and solving $R' + 6y' = -6$ with the given initial conditions, we have $R(0) = 4$ and $R + 6y = 4 - 6x$. Since $R(1/2) = R(x_1) = 0$, it follows that

$$y(1/2) = 1/6. \tag{18}$$

So the solution of the differential equation system at the end of phase 1 represents the situation that all nodes are coloured, and indicates a bisection of size asymptotically $2z_{11}n = n/3$.

Now we are in position to carry out the formal analysis. We wish to avoid the complications created by the prioritization, which makes it difficult to rigorously establish the meaining of the ϕ_i. For a given sufficiently small $\epsilon > 0$ (let us say $\epsilon < 10^{-6}$ in order to derive Theorem 1), consider the deprioritized algorithm given in Figure 4. Pre-phase 1, where a large number of vertices with no coloured neighbours are coloured in pairs, is just to ensure a good supply of symmetric pairs of vertices of types $(0,1), (0,2)$ and $(1,2)$ before entering phase 1. Note that the way the deprioritization is carried out here differs slightly from that in [20], since in that paper the ϕ_i were precomputed from the solution of the differential equations analogous to (8), whereas here, they are computed dynamically in the algorithm. The net effect is the same, and either version will work here, the only difference being some aspects of the justification which determine which version may be preferred.

In pre-phase 1, we have a unique operation which colours two vertices of type $(0,0)$. Working as in the lead-up to (1), we find that the expected increment of Z_{ij} due to the vertex v is $\mathbf{E}[\Delta(Z_{ij})] = 4\bar{d}_{ij} - \delta_{00}$. Each operation involves four such vertices v (except for cases that one vertex is involved more than once in the same operation, which happens with probability $O(1/W)$ and

Pre-phase 1: **do** the following $\lfloor \epsilon n \rfloor$ times:
select u.a.r. a non-adjacent (0,0)-symmetric pair
and perform Rand;
Phase 1: **while** all of Z_{01}, Z_{10}, Z_{02}, Z_{20}, Z_{21} and Z_{12} are non-zero
let $\phi_1 = \dfrac{1-2\bar{d}_{02}-\bar{d}_{12}}{1+\bar{d}_{02}+2\bar{d}_{12}}$; $\phi_2 = \dfrac{3\bar{d}_{02}}{1+\bar{d}_{02}+2\bar{d}_{12}}$;
$\phi_3 = \dfrac{3\bar{d}_{12}}{1+\bar{d}_{02}+2\bar{d}_{12}}$;
with probability ϕ_1
select u.a.r. a $(0,1)$-symmetric pair and perform Maj;
with probability ϕ_2
select u.a.r. a $(0,2)$-symmetric pair and perform Maj;
with probability ϕ_3
select u.a.r. a $(1,2)$-symmetric pair and perform Maj;
Cleanup: colour any remaining uncoloured vertices, half of them R
and half B, in any manner, and output the bisection R, B.

Fig. 3. Algorithm deprioritized 4-min greedy for bisection of random 4-regular graphs

is therefore ignored.) At this stage, we entirely avoid using the rb-symmetry assumption (as a reminder of which we reinstate r and b as general subscripts). Referring back to (1) and the ensuing derivation of (8), the *suggested* system of differential equations is

$$\begin{aligned} z'_{rb} = {} & \tfrac{-8(4-r-b)z_{rb}+4(5-r-b)z_{(r-1)b}\delta_{r>0}+4(5-r-b)z_{r(b-1)}\delta_{b>0}}{w(x)} \\ & -\delta_{rb=00} \end{aligned} \tag{19}$$

where

$$\begin{aligned} w(x) = {} & 4z_{00}(x)+3z_{01}(x)+3z_{10}(x)+2z_{02}(x)+2z_{20}(x) \\ & +2z_{11}(x)+z_{30}(x)+z_{03}(x) \end{aligned} \tag{20}$$

and for any statement S, δ_S denotes 1 if S is true and 0 otherwise. These apply for $0 \le r+b \le 4$, together with the additional equation $y'(x)=0$, with the initial conditions (12). It follows from the symmetric nature of the equations that the solution (which exists uniquely by standard theory of differential equations) must be the symmetric one, satisfying $z_{ij}(x) = z_{ji}(x)$ for all i, j and x. We write $\tilde{z}_{ij}(x)$ for the (unique) solutions of this initial value problem, $0 \le x \le \epsilon$.

Let $Z_{rb}(t)$ denote the value of Z_{rb} after t steps of the algorithm. Regarding $z_{rb}(t/n)$ as $Z_{rb}(t)/n$, the right hand side of give the expected one-step change in the variables Z_{rb} with error $O(1/n)$. (This error is due to the changing value of the variables between when one vertex of type $(0,0)$ is chosen and the next.) We may now apply the differential equation method (using, for example, [18, Theorem 5.1] or the simplified version [20, Theorem 3]) to deduce that during

pre-phase 1, we have a.a.s.

$$Z_{rb}(t) = n\tilde{z}_{rb}(t/n) + o(n) \tag{21}$$

for each r and b. This applies until either $t = \lfloor \epsilon n \rfloor$ or one of the derivatives approaches a singularity, which we can prevent by restricting to a domain in which $w > \epsilon$, or the differential equations no longer apply for some other reason, which in this case only occurs if Z_{00} reaches 0. Note that the derivatives are all $O(1)$, so $\tilde{z}_{00}(x)$ stays close to 1 for $x < \epsilon$ (recall $\epsilon > 0$ is arbitrarily small). We conclude that a.a.s.

$$Z_{rb}(t_0) = n\tilde{z}_{rb}(t_0/n) + o(n), \qquad t_0 := \lfloor \epsilon n \rfloor. \tag{22}$$

We also note that z'_{01} must be strictly positive, and so $\tilde{z}_{01}$, $\tilde{z}_{02}$ and $\tilde{z}_{12}$ are strictly positive on $(0, \epsilon)$. Thus, in particular, for sufficiently small $\epsilon_1 = \epsilon_1(\epsilon) > 0$,

$$\tilde{z}_{01}(\epsilon) \geq \epsilon_1, \quad \tilde{z}_{02}(\epsilon) \geq \epsilon_1, \quad \tilde{z}_{12}(\epsilon) \geq \epsilon_1. \tag{23}$$

Now consider phase 1. Note that the values of the ϕ_i are defined in an asymmetric way, but they in turn affect the expected changes symmetrically. Arguing as in the lead-up to (8) (but with the discussion around the ϕ_i simplified because they are prescribed), the expected changes in the Z_{rb} can easily be computed with error $O(1/W)$. For $(r, b) = (0, 0)$, $(0, 1)$ and $(1, 1)$, these expected changes are given by the right hand sides of the equations in (8) (reading z_{rb} as Z_{rb}), with $w = 4z_{00} + 3z_{01} + 3z_{10} + 2z_{02} + 2z_{20} + 2z_{11} + z_{12} + z_{21} + z_{30} + z_{03}$ and the replacement equation

$$z'_{11} = (3z_{01} + 3z_{10} - 4z_{11})\frac{D}{w} \tag{24}$$

to avoid the rb-symmetry assumption. At this point we do not try to argue (as in the informal discussion) that the other variables can be ignored. Some of the analogous equations for those variables are

$$\begin{aligned} z'_{02} &= (3z_{01} - 4z_{02})\frac{D}{w} - \theta_2, & z'_{03} &= (2z_{02} - 2z_{03})\frac{D}{w}, & z'_{04} &= z_{03}\frac{D}{w}, \\ z'_{12} &= (2z_{02} + 2z_{11} - 2z_{12})\frac{D}{w} - \theta_3, & z'_{13} &= (z_{03} + z_{12})\frac{D}{w}, & z'_{22} &= (z_{12} + z_{21})\frac{D}{w}, \end{aligned} \tag{25}$$

and the symmetrically reversed functions have symmetrically reversed equations (the θ_i are defined without symmetric reversal in these equations, of course, and w is defined in (3)). Continue the definition of the functions $\tilde{z}_{ij}(x)$ for $x > \epsilon$ by the solution of these equations (z'_{00} and z'_{01} given in (8), the other variables in (23) and (24), together with the symmetrically reversed versions)

with initial conditions given by the values of these functions at $x = \epsilon$ as determined above.

Again applying the differential equation method, we deduce that (20) holds a.a.s. as long as the solution set $\tilde{z}_{ij}$ stays within a predefined closed domain which does not contain singularities of the derivatives, and also the variables Z_{01}, Z_{02} and Z_{12} and their symmetrically reversed counterparts stay positive (so that the operations can be carried out when required) and the ϕ_i remain positive (so that the probability step in the algorithm is well defined). With ϵ_1 as before, we may select the domain L defined by $\tilde{z}_{rb} \geq \epsilon_1$ for $rb = 01$, 10, 02, 20, 12 and 21, $w \geq \epsilon_1$, and $\theta_i \geq \epsilon_1$ for $i = 1$, 2 and 3. By (22) and the symmetrically reversed versions, the first set of these inequalities hold at $x = \epsilon$. On the other hand the constraints on θ_i hold for ϵ_1 sufficiently small by a slightly deeper analysis which shows that $\tilde{z}_{rb} = \Theta(\epsilon^{r+b})$ (and then considering (1)). Finally, $w(0) = 1$, and the derivative of w is clearly bounded near 0. Thus, the solution at $x = \epsilon$ lies within the domain L. We need to study where the solution leaves L.

First, this is a convenient point to consider symmetry. It follows as before, from the symmetry in the differential equations and the initial conditions, that the unique solution is symmetric, with $\tilde{z}_{ij} = \tilde{z}_{ji}$. The symmetric solution must satisfy the equations (8) and (24), with w as in (9). In discussing the solution, we may therefore restrict our attention to the variables which appear in these equations, ignoring z_{rb} with $r > b$.

Arguing by continuity, z'_{01} as given in (8) is strictly positive for $x < \delta$, where $\delta > 0$ is an absolute constant independent of ϵ. In view of the equations (4) which was involved in the (circuitous) definition of θ_i to represent ϕ_i, we see that z'_{02} is identical to 0. (The reader may find it easiest to refer to (7) in verifying this.) Thus $\tilde{z}_{02} \equiv \tilde{z}_{02}(\epsilon)$, and we obtain $\tilde{z}_{12} \equiv \tilde{z}_{12}(\epsilon)$ in a similar fashion. Hence, for ϵ sufficiently small, the solution set $\tilde{z}_{ij}$ stays inside L for $x < \delta$, and can only leave L when, for some $x \geq \delta$,

$$\tilde{z}_{01} = \epsilon_1, \;\; w = \epsilon_1 \text{ or } \theta_i = \epsilon_1 \text{ for some } i \in \{1, 2, 3\}. \tag{26}$$

Note that for ϵ and ϵ_1 sufficiently small, the initial conditions for $\tilde{z}_{ij}$ are arbitrarily close to (12). Let us denote the solutions with initial conditions (12) by $\bar{z}_{ij}$. By standard theory of first order systems of differential equations, it follows that the functions $\tilde{z}_{ij}$ can be made arbitrarily close to $\bar{z}_{ij}$ in the domain L, by taking ϵ and ϵ_1 sufficiently small. By (15) and the definition of x_1, the conditions corresponding to (25) for $\bar{z}_{ij}$ are not reached until x approaches the x_1 given in (17), at which point all $\bar{z}_{ij}$ reach 0 by (16). It follows that, as ϵ and ϵ_1 tend towards 0, the value of x at the exit point of the $\tilde{z}_{ij}$ from D also tends towards $x_1 = 1/2$.

We also introduced the variable y, to keep track of the number Y of times a

(1,2)-symmetric pair is processed. In phase 1, this is the only contributor to the size of the bisection. The conclusion is that there is a deprioritized algorithm in which the values of the variables $Z_{ij}(t)$ are a.a.s. $\tilde{n}z_{ij}(t/n) + o(n)$, and $Y(t)$ is a.a.s. $n\tilde{y}(t/n) + o(n)$ where the functions $\tilde{z}_{ij}$ and $\tilde{y}$ solve (8), (9), (10) and (11), and moreover, that these functions can be made arbitrarily close to the solutions $\bar{z}_{ij}$ and $\bar{y}$ with initial conditions (12). It follows that the size of the bisection at the end of phase 1 is a.a.s.

$$2y(x_1)n + O(\delta' n)$$

where $\delta' \to 0$ as ϵ and ϵ_1 go to 0. By (16) the clean-up phase increases the size of the bisection by a negligible amount. In view of (18), this completes the proof of Theorem 1.

4 Analysis of the algorithm for random cubic graphs

Consider analyzing the algorithm 3-min greedy in the same way as we have done for 4-min greedy. In the discussion leading up to (8) we obtain for a vertex newly coloured R

$$\begin{aligned}
d_{00} &= -\frac{3Z_{00}}{W} \qquad d_{01} = -\frac{2Z_{01}}{W} \qquad d_{02} = -\frac{Z_{02}}{W}\ d_{03} = 0\\
d_{10} &= \frac{3Z_{00} - 2Z_{10}}{W}\ d_{11} = \frac{2Z_{01} - Z_{11}}{W}\ d_{12} = \frac{Z_{02}}{W}\\
d_{20} &= \frac{2Z_{10} - Z_{20}}{W} \quad d_{21} = \frac{Z_{11}}{W}\\
d_{30} &= \frac{Z_{02}}{W}
\end{aligned}$$

where $W = 3Z_{00} + 2Z_{01} + 2Z_{10} + Z_{02} + Z_{20} + Z_{11}$, and consequently, in place of (1)

$$\begin{aligned}
\bar{d}_{00} &= -\frac{6Z_{00}}{W}\\
\bar{d}_{01} &= \frac{3Z_{00} - 4Z_{01}}{W}\ \bar{d}_{11} = \frac{2Z_{01} + 2Z_{10} - 2Z_{11}}{W}\\
\bar{d}_{02} &= \frac{2Z_{01} - 2Z_{02}}{W}\ \bar{d}_{12} = \frac{Z_{02} + Z_{11}}{W}\\
\bar{d}_{03} &= \frac{Z_{02}}{W}
\end{aligned} \tag{27}$$

The difference in the analysis for algorithm 3-min greedy, given in Figure 2,

with respect to the analysis of the previous section is that we must analyse each phase separatly, feeding to it the solutions to the differential equation in the previous phase.

For the non-rigorous discussion of phase 1, assume that a given iteration in phase 1 colours a (0,1)-symmetric pair with probability ϕ_1, and (0,2) with probability ϕ_2, where $\phi_1 + \phi_2 = 1$.

Analogous to (13), we find eventually (see [6] for details of all but the final step)

$$z'_{00} = \frac{-12z_{00}}{R},\ z'_{01} = \frac{6z_{00} - 8z_{01}}{R},\ z'_{11} = \frac{8z_{01} - 4z_{11}}{R},\ z'_{12} = \frac{2z_{11}}{R},\ y' = 0 \tag{28}$$

where $R = 3z_{00} + 6z_{01} + z_{11}$, with initial conditions

$$z_{00}(0) = 1, \quad z_{01}(0) = z_{11}(0) = z_{12}(0) = y(0) = 0. \tag{29}$$

Note that y has zero derivative because the two operations in this phase do not add to the bisection size.

We are interested in the point that z_{01} first goes negative, which by numerical solution (using an order 2 Runge Kutta algorithm) occurs when

$$x = x_1 \approx 0.41178,\ z_{00} \approx 0.002405,\ z_{11} \approx 0.046633,\ z_{12} \approx 0.063700. \tag{30}$$

The whole algorithm "takes off" at the start because the derivative of z_{01} is strictly positive, so a.a.s. phase 1 does not quickly use up all vertices to be processed.

At the point given by (29), since z_{01} and z_{02} are both 0, phase 2 is entered. The situation is similar to phase 1, but with different operations. We pause to highlight one difference. When colouring a (1,1)-symmetric pair, there is one pair exposed from each of two vertices of type $(1,1)$, and the expected number of new vertices of type (0,1) arising from this is $2\bar{d}_{01} = (4Z_{01} - 4Z_{02})/W$, where *two* vertices of type $(1,1)$ are used in this operation. The rest of the argument is similar and the resulting differential equation is

$$\begin{aligned} z'_{00} &= \frac{-6z_{00}}{R'},\ z'_{01} = 0,\ z'_{11} = \frac{3z_{00} - 8z_{01} - 3z_{11} - 2z_{02}}{R'},\ z'_{12} = \frac{z_{02} + z_{11}}{R}, \\ z'_{02} &= \frac{2z_{01} - 2z_{02}}{R'},\ y' = \frac{-6z_{00} + 24z_{01} + 4z_{02} + 2z_{11}}{R'} \end{aligned} \tag{31}$$

where $R' = 8z_{01} + 2z_{02} + z_{11}$, with initial conditions given by (29) and $z_{01} = z_{02} = 0$. (Again, see [6] for details.) Note that the derivative of y comes from

the fact that the colouring of a (1,1)-symmetric pair increases de bisection by two, and this is the only cause of increase of the bisection.

The point of interest is

$$x_2 = \sup\{x \,:\, z_{11} > 0,\ w > 0\}. \tag{32}$$

This corresponds to the beginning of phase 3. During phase 3 the number of bicoloured edges created is 2 for every pair of vertices of types $(1,2)$ and $(2,1)$ (using rb-symmetry) and at most 6 for every other pair coloured except types $(0,3)$ and $(3,0)$, which give none. Since $z_{01} = z_{02} = z_{11} = 0$ at x_2, our upper bound for the size of the bisection is thus $(6z_{00} + 2z_{12} + 2y + \epsilon)n$ where the variables are evaluated at x_2. Solving numerically, we find

$$6z_{00}(x_2) + 2z_{12}(x_2) + 2y(x_2) < 0.174038. \tag{33}$$

As with the 4-regular case, we introduce a deprioritized algorithm, in Figure 3.

Pre-phase 1: **do** the following $\lfloor \epsilon n \rfloor$ times:
select u.a.r. a non-adjacent (0,0)-symmetric pair,
and perform Rand;
Phase 1: **while** all of Z_{01}, Z_{10}, Z_{02} and Z_{20} are non-zero
let $\theta = \frac{2Z_{01}-2Z_{02}}{W}$ and $\phi = \frac{1-\theta}{1+\theta}$;
with probability ϕ
select a (2,0)-symmetric pair and perform Maj;
otherwise
select a (1,0)-symmetric pair and perform Maj;
Pre-phase 2: **do** $\lfloor \epsilon n \rfloor$ steps as in Pre-phase 1;
Phase 2: **while** $Z_{01} > 0$, $Z_{10} > 0$ and $Z_{11} > 1$
let $\theta_2 = \frac{3Z_{00}-4Z_{01}}{W}$ and $\phi_2 = \frac{1-2\theta_2}{1-\theta_2}$;
with probability ϕ_2
select a (1,0)-symmetric pair an perform Maj;
otherwise
select a (1,1)-symmetric pair an perform Rand;
Phase 3: as for Algorithm 3-min greedy.

Fig. 4. Algorithm deprioritized 3-min greedy for bisection

The formal analysis, completing the proof of Theorem 2, is essentially the same as the previous section within each phase, so is omitted from this paper (the interested reader may again consult [6]).

5 Maximum Bisection

Let us consider the variation of the alorithms 4-min greedy and 3-min greedy (given in Figure 1 and 2,respectively) obtained by changing the meaning of

Maj, now we will colour a vertex with the minority colour among its coloured neighbours.

Let us say that an edge is *fully coloured* when both its ends are finally coloured. A fully coloured edge is *monocoloured* if both ends have the same colour and *bicoloured* if both ends have different colour. So the monocoloured edges by min greedy get bicoloured by max greedy and vice versa, whenever the vertices of the graph are treated in the same order (which happens with the same probability, in both cases). That is, every edge that counts in the bisection for one algorithm does not count in the other and vice versa. Therefore, taking into account that the total number of edges in a 4-regular graph is $2n$, and in a cubic graph is $1.5n$, we have proved Therorems 3 and 4.

6 Further remarks

We have given an application of the differential equation method to analyse the bisection of random cubic and random 4-regular graphs, providing reasonable bounds both for max and min bisection.

One natural problem remains open, to find the exact solution of the system of differential equations for cubic graphs. By doing so, more accurate constants will be obtained, up to know we have been only able to solve them numerically. This may have also a bearing on extending the method to larger d.

References

[1] S. Arora, D. Karger, and M. Karpinski. Polynomial time approximation schemes for dense instances of NP-hard problems. In *36th ACM Symposium on Theory of Computing*, pages 284–293, 1995.

[2] S.L. Bezroukov, R. Elsasser, B. Monien, R. Preis, and J.P. Tillich. New spectral lower bounds on the bisection width of graphs. In U. Brandes and D. Wagner, editors, *26th. Graph-Theoretic Concepts in Computer Science*, volume 1928, pages 23–34. Springer, Lecture Notes in Computer Science, 2000.

[3] B. Bollobas. The isoperimetric number of random regular graphs. *European Journal of combinatorics*, 9:241–244, 1984.

[4] B. Bollobás. *Random Graphs.* Academic Press, London, 1985.

[5] T. Bui, S. Chaudhuri, T. Leighton, and M. Sipser. Graph bisection algorithms with good average case behavior. *Combinatorica*, 7:171–191, 1987.

[6] J. Díaz, N. Do, M. J. Serna, and N. Wormald. Bisection of random cubic graphs. In J. Rolim, editor, *RANDOM*, volume 2000 of *Lecture Notes in Computer Science*, pages 291–302. Springer-Verlag, 2002.

[7] Josep Díaz, Jordi Petit, and Maria Serna. A survey on graph layout problems. *ACM Computing Surveys*, 34:313–356, 2002.

[8] U. Feige and R. Krauthgamer. A polylogarithmic approximation of the minimum bisection. In *41st. IEEE Annual Symposium on Foundation of Computer Science*, pages 23–26, 2000.

[9] U. Feige and M. Langberg. The RPR^2 rounding technique for semidefinite programs. In P. Spirakis, F. Orejas, and J. van Leeuwen, editors, *28th. International Colloquium on Automata, Languages and Programming*, volume 2076, pages 213–224. Lecture Notes in Computer Science, Springer-Verlag, 2001.

[10] U. Feigue, M. Karpinski, and M. Langberg. A note on approximating MAX-BISECTION on regular graphs. *Information Processing Letters*, (79):181–188, 2000.

[11] M. Fiedler. A property of the eigenvectors of nonnegtive symetric matrices and its application to graph theory. *Czechoslovak Mathematical journal*, (25):619–633, 1974.

[12] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Freeman, San Francisco, 1979.

[13] K. Jansen, M. Karpinski, A. Lingas, and E. Seidel. Polynomial time approximation schemes for MAX-BISECTION on planar and geometric graphs. In A. Ferreira and H. Reichel, editors, *18th. Annual Symposium on Theoretical Aspects of Computer Science*, volume 2010 of *Lecture Notes in Computer Science*, pages 365–375. Springer, 2001.

[14] M. Karpinski, M. Kowaluk, and A. Lingas. Approximation algorithms for MAX-BISECTION on low degree regular graphs and planar graphs. Technical Report TR00-051, Electronic Colloquium on Computational Complexity, 2000. http://www.eccc.uni-trier.de/eccc/.

[15] A.V. Kostochka and L.S. Melnikov. On bounds of the bisection width of cubic graphs. In J. Nesetril and M. Fiedler, editors, *Fourth Czechoslovakian Symposium on Combinatorics, Graphs and Complexity*, pages 151–154. Elsevier Science Publishers, 1992.

[16] T. Luczak, S. Janson, and A. Ruciinski. *Random Graphs*. Wiley, 2000.

[17] B. Monien and R. Preis. Upper bounds on the bisection width of 3- and 4-regular graphs. In A. Pultr J. Sgall and P. Kolman, editors, *Mathematical Foundations of Computer Science*, volume 2136, pages 524–536. Springer, Lecture Notes in Computer Science, 2001.

[18] N.C. Wormald. The differential equation method for random graph processes and greedy algorithms. In M. Karoński and H. Prömel, editors, *Lectures on*

Approximation and Randomized Algorithms, pages 73–155. PWN, Warsaw, 1999.

[19] N.C. Wormald. Models of random regular graphs. In *Surveys in Combinatorics*, pages 239–298. Cambridge University Press, 1999.

[20] N.C. Wormald. Analysis of greedy algorithms on graphs with bounded degree. *Discrete Mathematics*, 2003.