Bisection of Random Cubic Graphs

J. Díaz*† N. Do‡ M.J. Serna*† N.C. Wormald‡§

Abstract

In this paper we present two randomized algorithms to compute the bisection width of random cubic graphs with n vertices, giving an asymptotic upper bound for the bisection width respectively of $0.174039n$ and $0.174501n$. We also obtain an asymptotic lower bound for the size of the max bisection of a random cubic graph with n vertices of $1.325961n$ and $1.325499n$. The analysis is based on the differential equation method.

1 Introduction

Given a graph $G = (V, E)$ with $|V| = n$ and n even, a *bisection* of V is a partition of V into two parts each of cardinality $n/2$, and its *size* is the number of edges crossing between the parts. A *minimum bisection* is a bisection of V with minimal size. 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 [GJ79]). 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 [GJ79], even for cubic graphs [BCLS87]. On the other hand, several exact and heuristic positive results are known (see for example [DPPS01]). In this paper, we deal with the problem of estimating the typical size of minimum and maximum bisections of random cubic graphs.

It is shown in [KM92] that all cubic graphs have bisection width at most $\frac{n}{4} + O(\sqrt{n}\log n)$, and there are cubic graphs with bisection width of at least $\frac{n}{9.9}$. Our first result is an asymptotic bound on the bisection width of random cubic graphs. We refer the reader to [JŁR00], for the definitions of u.a.r. (uniformly at random) and a.a.s. (asymptotically almost surely). For such statements, $n \to \infty$ and we restrict n to even integers.

Theorem 1 *The bisection width of a random cubic graph on n vertices is a.a.s. smaller than $0.174039n$.*

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

†Supported by the FET programme of the EU under contract IST-1999-14186 (ALCOM-FT)

‡Department of Mathematics and Statistics, University of Melbourne, VIC 3010, Australia

§Supported by the Australian Research Council

We actually give two quite different proofs of approximately equal upper bounds; the other is $0.17451n$ (see Theorem 3).

Regarding the size of the maximum bisection, we are not aware of any non-trivial lower bounds. Our second result provides an asymptotic bound on the maximum bisection of random cubic graphs.

Theorem 2 *The maximum bisection of a random cubic graph with n vertices is a.a.s. greater than $1.325961n$.*

We conjecture that the largest balanced bipartite subgraph of a random cubic graph is a.a.s. almost the same size as the largest bipartite subgraph. We can state this even more strongly, as follows.

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

2 Greedy algorithms for Minimum Bisection

In this section we prove Theorems 1 and 2. Given a random cubic graph, and given a partial assignment of colors red (R) and blue (B) to its vertices, we classify the non-colored vertices according with the number of their colored neighbors:

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

We will consider the greedy procedure Simple greedy given in Figure 1 to find w.h.p. a balanced partition (R, B) with small bisection. The algorithm colors vertices in pairs to maintain balancedness. It repeatedly uses three operations. Op1 consists of choosing one vertex of type $(1, 0)$ and one of type $(0, 1)$ u.a.r., and coloring each the same as its colored neighbor. Op2 consists of choosing one vertex of type $(2, 0)$ and one of type $(0, 2)$ u.a.r., and coloring each the same as its colored neighbors. Op3 consists of choosing two non-adjacent vertices of type $(1, 1)$ u.a.r., and coloring one with R and one with B.

Note that the size of the bisection is the number of *bicolored* edges, with one vertex of each color.

One method of analyzing the performance of a randomized algorithm is to use a system of differential equations to express the expected changes in variables describing the state of the algorithm during its execution. An exposition of this method is in [Wor99b], which includes various examples of graph-theoretic optimization problems. For purposes of exposition, we continue for the present to discuss Algorithm simple min greedy, 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 analyze to yield our claimed bounds. We call this variation of algorithm a *deprioritized* algorithm as in [Wor02], where this technique was first used.

We use the pairing model to analyses n-vertex cubic graphs, generated u.a.r. Briefly, to generate such a random graph, it is enough to begin with $3n$ *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. 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 [Bol85, Wor99a] for a full description).

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

Phase 1: **repeat**

  if there are vertices of both types (2,0) and (0,2) **then** perform Op2;
  else if there are vertices of both types (1,0) and (0,1) **then** perform Op1;

until no new vertex is colored

Phase 2: **repeat**

  if there are vertices of both types (1,0) and (0,1)**then** perform Op1;
  else if there are at least two vertices of type (1,1) **then** perform Op3;

until no new vertex is colored

Phase 3: **repeat**

  if there are vertices of both types (3,0) and (0,3)
    then choose one of each type at random, and color each the same as its colored neighbor;
  if there are vertices of both types (2,1) and (1,2)
    then choose one of each type at random, and color each with the majority of its colored neighbors;

until no new vertex is colored

color any remaining uncolored vertices, half of them R and half B, in any manner, and output the bisection R, B.

Figure 1: Algorithm simple min greedy for Min Bisection

Without loss of generality, when stating such asymptotic results, we restrict n to being even to avoid parity problems. We consider Algorithm simple min greedy applied directly to the random pairing. As discussed in [Wor99b], 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 uncolored vertices of type (r, b). To analyze the algorithm, when a vertex is colored we immediately expose all pairs involved in that vertex. In this way, the numbers Z_{rb} are always determined.

At any time, let W denote the number of points not yet involved in exposed pairs. These are the points available for the pairs that will be exposed during the next step. Then $W = 3Z_{00} + 2Z_{01} + 2Z_{10} + Z_{02} + Z_{20} + Z_{11}$.

Consider what happens when a vertex u is newly colored R and one of the pairs containing a point p in that cell is exposed. The other point will lie in some vertex v. 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. Then, up to terms $O(1/W)$,

$$d_{00} = -\frac{3Z_{00}}{W,}, \ d_{01} = -\frac{2Z_{01}}{W}, \ d_{02} = -\frac{Z_{02}}{W}, \ d_{03} = 0, \ d_{11} = \frac{2Z_{01} - Z_{11}}{W},$$
$$d_{12} = \frac{Z_{02}}{W}, \ d_{10} = \frac{3Z_{00} - 2Z_{10}}{W}, \ d_{20} = \frac{2Z_{10} - Z_{20}}{W}, \ d_{30} = \frac{Z_{02}}{W}, \ d_{21} = \frac{Z_{11}}{W}.$$

The error term $O(1/W)$ is due to adjustments occurring when v happens to be the same as u (and also saves us from specifying whether the variables refer to the graph before or after coloring u).

The corresponding equations when a vertex is colored B form a symmetric set with these: they are the same but with the index pair on all variables swapped. Therefore, the expected increments due to a dual step, consisting of one new pair from a vertex of each color, is $\bar{d}_{r,b}$:

$$\bar{d}_{00} = -\frac{6Z_{00}}{W},\ \bar{d}_{01} = \frac{3Z_{00} - 4Z_{01}}{W},\ \bar{d}_{02} = \frac{2Z_{01} - 2Z_{02}}{W},\ \bar{d}_{11} = \frac{2Z_{01} + 2Z_{10} - 2Z_{11}}{W},$$
$$\bar{d}_{12} = \frac{Z_{02} + Z_{11}}{W},\ \text{where } W = 3Z_{00} + 2Z_{01} + 2Z_{10} + Z_{02} + Z_{20} + Z_{11}. \tag{1}$$

Symmetrically corresponding variables have symmetrically corresponding equations. Note that $\bar{d}_{03}$ and its symmetric mate are not required, since vertices of type $(0,3)$ are just colored in phase 3 with the color of all their neighbors and therefore are not incident with any bicolored edges.

The rest of our discussion, until considering the deprioritized algorithm, is mainly motivation but also includes some derivations of formulae used later. The difficulty of analysis is caused by the prioritization in phases 1 and 2. In phase 1 the algorithm performs one of two types of operations on a pair of vertices, in each case coloring them the same as their neighbors. If vertices of types $(0,2)$ and $(2,0)$ exist, these have priority (Op2), while if they don't, but at least one of each of types $(0,1)$ and $(1,0)$ exist, then these are treated (Op1). In practice, for a random graph there are never very many vertices of types $(0,2)$ and $(2,0)$; as soon as there are some of each, they are processed, and the number of new ones arising tends to be less than the number used up. Which leaves a few of one type or the other, waiting for the matching ones to be created. To proceed with the discussion, we assume that a given iteration in phase 1 performs Op1 with probability ϕ, and therefore Op2 with probability $1-\phi$.

Define
$$\theta = \frac{2Z_{01} - 2Z_{02}}{W}.$$
Due to Op1, there are *two* pairs exposed from each vertex. Thus the expected number of new vertices of type (0,2) arising from this is $2\bar{d}_{02} = \frac{4Z_{01}-4Z_{02}}{W} = 2\theta$. At the moment, we make the assumption that we have the *rb-symmetry*: for all i and j, $Z_{ij} = Z_{ji}$ (later we will see how to remove it). With this last assumption, there also are 2θ new vertices of type (2,0) it is also 2θ. So each iteration of phase 1 performing Op1 gives rise on average to 2θ steps performing Op2 in order to keep Z_{02} and Z_{20} constant.

However, these iterations with Op2 cause further vertices of types (0,2) and (2,0). In Op2, one pair is exposed from each vertex, which gives expected increase of θ to Z_{02} and to Z_{20}. Following from this, we expect θ extra steps of Op2. For $\theta < 1$, the expected number of iterations of Op2 executed by the algorithm for each Op1, is $\sum_{i=0}^{\infty} 2\theta(\theta)^i = \frac{2\theta}{1-\theta}$.

As the probability of Op2 will be $\frac{2\theta}{1-\theta}\phi = 1-\phi$, and we can conclude that

$$\phi = \frac{1-\theta}{1+\theta}. \tag{2}$$

Using this, we can compute the expected increments of the random variables Z_{ij} in each iteration in phase 1.

$$\mathbf{E}\left[\Delta(Z_{ij})\right] = \phi(2\bar{d}_{ij} - \delta_{01}) + (1-\phi)(\bar{d}_{ij} - \delta_{02}) = (1+\phi)\bar{d}_{ij} - \phi\delta_{01} - (1-\phi)\delta_{02} \tag{3}$$

for any i,j with $i \le j$ and $i+j = 3$, where $-\delta_{pq} = 1$ if $(p,q) = (i,j)$, and 0 otherwise. The equations for $i > j$ are the symmetric ones.

In (3) we may use $\bar{d}_{ij}$ as given in (1) but with $\bar{d}_{11} = (4Z_{01} - 2Z_{11})/W$ by rb-symmetry. Without justification at this point, 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} (all as functions of the number t of iterations). If we scale both time and the variables by dividing by n, and denote Z_{ij}/n by z_{ij} and W/n by w, then the equations are

$$z_{00}' = -(1+\phi)\frac{6z_{00}}{w},\ z_{01}' = (1+\phi)\frac{3z_{00}-4z_{01}}{w} - \phi,\ z_{02}' = (1+\phi)\frac{2z_{01}-2z_{02}}{w} + (\phi-1),$$
$$z_{11}' = (1+\phi)\frac{4z_{01}-2z_{11}}{w},\ z_{12}' = (1+\phi)\frac{z_{02}+z_{11}}{w},\ \text{where } w = 3z_{00}+4z_{01}+2z_{02}+z_{11},$$

by rb-symmetry, $\phi = \frac{1-\theta}{1+\theta}$ and $\theta = \frac{2z_{01}-2z_{02}}{w}$. The initial conditions at $x = t/n = 0$ are

$$z_{00}(0) = 1,\ z_{01}(0) = 0,\ z_{02}(0) = 0,\ z_{03}(0) = 0,\ z_{11}(0) = 0,\ z_{12}(0) = 0. \tag{4}$$

Note that by the way we defined ϕ in (2), $z_{02}' \equiv 0$ and hence $z_{02} \equiv 0$. We are interested in the point that z_{01} first goes negative, which by numerical solution occurs when

$$x = x_1 \approx 0.41178,\ z_{00} \approx 0.002405,\ z_{11} \approx 0.046633,\ z_{12} \approx 0.063700. \tag{5}$$

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 this point, since z_{01} and z_{02} are both 0, phase 2 is entered. The situation is similar to phase 1, but now the operation with highest priority is Op1. The other operation, Op3, is such that two vertices of type $(1,1)$ are randomly chosen and colored, one B and one R. Following the discussion above, let $\theta_2 = \frac{3Z_{00}-4Z_{01}}{W}$. Then, due to Op3, 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 $2\bar{d}_{01} = \frac{4Z_{01}-4Z_{02}}{W} = 2\theta_2$, where *two* vertices of type $(1,1)$ are used in this operation. In Op1, the expected number is $2\theta_2$. Letting ϕ_2 denote the probability that at a given time Op3 is performed in phase 2, the probability of Op1 will be $\frac{\theta_2}{1-2\theta_2}\phi_2 = 1-\phi_2$, giving

$$\phi_2 = \frac{1-2\theta_2}{1-\theta_2}.$$

In place of (3) we now have

$$\mathbf{E}\left[\Delta(Z_{ij})\right] = \phi_2(\bar{d}_{ij} - \delta_{11}) + (1-\phi_2)(2\bar{d}_{ij} - \delta_{01}) = (2-\phi_2)\bar{d}_{ij} - \phi_2\delta_{11} - (1-\phi_2)\delta_{01}. \tag{6}$$

To set the differential equations, let Y be a random variable that keeps track of the number of times Op3 is performed, and let $y = Y/n$. This is needed for record-keeping because each such operation causes two bicolored edges. Then

$$z_{00}' = -(2-\phi_2)\frac{6z_{00}}{w},\ z_{01}' = (2-\phi_2)\frac{3z_{00}-4z_{01}}{w} + (\phi_2-1),\ z_{02}' = (2-\phi_2)\frac{2z_{01}-2z_{02}}{w},$$
$$z_{11}' = (2-\phi_2)\frac{4z_{01}-2z_{11}}{w} - 2\phi_2,\ z_{12}' = (2-\phi_2)\frac{z_{02}+z_{11}}{w},\ y' = \phi_2,$$

with w as before and with initial conditions given by (5) and $z_{01} = z_{02} = 0$.

By the choice of ϕ_2, $z_{01} \equiv 0$, so $z_{02} \equiv 0$. The point of interest is

$$x_2 = \sup\{x \,:\, z_{11} > 0,\ w > 0,\ \theta_2 < 1\}. \tag{7}$$

Numerically, we find that θ_2 does not reach 1 before z_{11} reaches 0, which clearly must therefore hold at x_2. This corresponds to the beginning of phase 3. During phase 3 the number of bicolored 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 colored 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 $(6z_{00}+2z_{12}+2y)n$ where the variables are evaluated at x_2. Solving numerically, we find

$$z_{00}(x_2) + 2z_{12}(x_2) + 2y(x_2) < 0.1740381057, \tag{8}$$

where the constant is correct to ten decimal places.

Now we are in position to carry out the formal analysis via a deprioritized algorithm. For a given sufficiently small $\epsilon > 0$, consider the deprioritized algorithm in Figure 2. Notice that pre-phase 1 ensures a good supply of vertices of types $(0,1)$, $(1,0)$, $(0,2)$ and $(2,0)$.

Pre-phase 1: **do** the following $\lfloor \epsilon n \rfloor$ times:
  select two non-adjacent type (0,0) vertices u.a.r.,
  color one with R and the other with B;
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 perform Op1;
  otherwise perform Op2;
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 perform Op3;
  otherwise perform Op1;
Phase 3: as for Algorithm simple min greedy.

Figure 2: Algorithm deprioritized min greedy for Min Bisection

The expected changes in the variables Z_{ij} for each edge exposed are given in (1). In pre-phase 1, the derivation of (2) applies, but with ϕ redefined 1 at all times, and with different terms in the equations for z'_{00} and z'_{01} due to the fact that the vertex being processed is type $(0,0)$ rather than $(0,1)$ or $(0,2)$. At this stage, we entirely avoid using the rb-symmetry assumption. Referring back to (1), this requires only the adjustment of the formulae for z'_{11} and w. The result is

$$z'_{00} = 1 - \frac{12z_{00}}{w},\ z'_{01} = 2\frac{3z_{00} - 4z_{01}}{w},\ z'_{02} = 2\frac{2z_{01} - 2z_{02}}{w},$$
$$z'_{11} = 2\frac{2z_{01} + 2z_{10} - 2z_{11}}{w},\ z'_{12} = 2\frac{z_{02} + z_{11}}{w},\ w = 3z_{00} + 2z_{01} + 2z_{10} + z_{02} + z_{20} + z_{11}.$$

Other derivatives z'_{ji} are the symmetric versions of z'_{ji} (with indices swapped); z'_{03} and z'_{30} are not needed. It follows that the unique solution must be the symmetric one, which satisfies $z_{ij}(t) = z_{ji}(t)$ for all i, j and t, as well as the stated equations.

Let $z_{ij}(t/n)$ denote Z_{ij}/n after t steps. Thus the previous equations give the expected one-step change in the variables Z_{ij} 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. This applies with initial conditions given in (4). We write $\tilde{z}_{ij}(x)$ for the (unique) solutions of this initial value problem, $0 \le x \le \epsilon$. We may now apply the differential equation method (using, for example, [Wor95, Theorem 1] or [Wor99b, Theorem 5.1]) to deduce that during pre-phase 1, we have a.a.s.

$$Z_{ij}(t) = n\tilde{z}_{ij}(t/n) + o(n) \tag{9}$$

for each i and j, where $Z_{ij}(t)$ is the value of Z_{ij} after t steps. 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 $\theta > -1 + \epsilon$ and $w > \epsilon$, 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 $x < \epsilon$ assuming that $\epsilon > 0$ is sufficiently small. We conclude that a.a.s.

$$Z_{ij}(t_0) = n\tilde{z}_{ij}(t_0/n) + o(n), \qquad t_0 := \lfloor \epsilon n \rfloor. \tag{10}$$

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

$$\tilde{z}_{01}(\epsilon) \geq \epsilon_1, \quad \tilde{z}_{02}(\epsilon) \geq \epsilon_1. \tag{11}$$

Now consider phase 1. Arguing as above, the expected changes in the Z_{ij} are given, with error $O(1/W)$, by the right hand sides of the equations in (2), with

$$w = 3z_{00} + 2z_{01} + 2z_{10} + z_{02} + z_{20} + z_{11}, \ \phi = \frac{1-\theta}{1+\theta}, \ \theta = \frac{2z_{01} - 2z_{02}}{w},$$

and the replacement equation

$$z'_{11} = (1+\phi)\frac{2z_{01} + 2z_{10} - 2z_{11}}{w}$$

to avoid the rb-symmetry assumption. Again the symmetrically reversed functions have symmetrically reversed equations (except that θ stays the same). Continue the definition of the functions $\tilde{z}_{ij}(x)$ for $x > \epsilon$ by the solution of these equations with initial conditions given by the values of these functions at $x = \epsilon$ as determined above.

Note that setting $z_{ij} = z_{ji}$ for all i and j in the equations, except in the definition of θ, again makes the formulae for z'_{ij} and z'_{ji} identical, despite the asymmetrical definition of θ. It follows that again the unique solution is symmetric, with $\tilde{z}_{ij} = \tilde{z}_{ji}$. We deduce that the equations (2) are satisfied, with the symmetric definitions of w and z_{11}, and we may restrict attention to the variables appearing there.

Again applying the differential equation method, we deduce that (9) 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} and Z_{02} stay positive. We may select the domain D satisfying $\tilde{z}_{01} \geq \epsilon_1$, $\tilde{z}_{02} \geq \epsilon_1$, $w > \epsilon$ and $\theta > 1 - \epsilon$. By (11), the first two of these inequalities hold at $x = \epsilon$, and the other two also hold for ϵ sufficiently small by boundedness of derivatives. Arguing by continuity, z'_{01} as given in (2) is strictly positive for $x < \delta$, where $\delta > 0$ is an absolute constant independent of ϵ. By definition of θ, z'_{02} is identical to 0, and so $\tilde{z}_{02} \equiv \tilde{z}_{02}(\epsilon)$. Thus, for ϵ sufficiently small, the solution set $\tilde{z}_{ij}$ stays within D for $x < \delta$, and can only leave D when, for $x > \delta$,

$$\tilde{z}_{01} = \epsilon_1, \ w = \epsilon \text{ or } \theta = 1 - \epsilon. \tag{12}$$

Note that for ϵ and ϵ' sufficiently small, the initial conditions for $\tilde{z}_{ij}$ are arbitrarily close to (4). Let us denote the solutions with initial conditions (4) 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 D, by taking ϵ and ϵ_1 sufficiently small. By numerical computation, the conditions corresponding to (12) are not reached by the solution $\bar{z}_{ij}$, until x approaches x_1 given in (5), at which point $\bar{z}_{01}$ reaches 0. It follows that, as ϵ and $\epsilon_1 \to 0$, the exit point of the $\tilde{z}_{ij}$ from D also tends towards x_1, and the values are given in the limit by the values of $\bar{z}_{ij}$ in (5).

A similar argument applies for phase 2, and at this point we also introduce the variable y, to keep track of the number Y of times Op3 is performed. The conclusion is that there is a deprioritized algorithm in which the values of the variables Z_{ij} are a.a.s. $\tilde{z}_{ij}n + o(n)$, where the functions $\tilde{z}_{ij}$ and $\tilde{y}$ solve (2). They can be made arbitrarily close to $\bar{z}_{ij}$ for all $x < x_2 - \delta'$, where x_2 is given in (7) and δ' is an arbitrary positive quantity. Additionally, $Y = yn + o(n)$ a.a.s. Note that $|\bar{z}_{ij}(x_2 - \delta') - \bar{z}_{ij}(x_2)| = O(\delta')$ since the derivatives are all bounded. Examining phase 3 as for Algorithm simple min greedy, it follows that the size of the bisection produced is a.a.s.

$$(z_{00}(x_2) + 2z_{12}(x_2) + 2y(x_2))n + O(\delta' n),$$

where δ' can be chosen arbitrarily small. By (8), this completes the proof of Theorem 1.

3 Maximum Bisection

Let us consider the maximization version of the bisection problem, Max Bisection, i.e. given a connected cubic graph with n vertices, for n even, find a bisection which maximizes the number of crossing edges. In general, the problem is also known to be NP-complete, even for cubic graphs, moreover for the particular case of cubic graphs, the problem can be approximated with an approximation ratio of 0.847 [KKL00]. For motivation on the problem, see [FR97].

Let us consider the variation Simple max greedy of the Algorithm 1 Simple min greedy, obtained by changing the meaning of Op1 and Op2. In the Simple max greedy algorithm, Op1 consists of choosing one vertex of type $(1, 0)$ and one of type $(0, 1)$ u.a.r., and coloring each opposite to its colored neighbor. Op2 consists of choosing one vertex of type $(2, 0)$ and one of type $(0, 2)$ u.a.r., and coloring each opposite to its colored neighbors. Op3, as before, consists of choosing two non-adjacent vertices of type $(1, 1)$ u.a.r., and coloring one with R and one with B. Also, in phase 3, change equal and majority, by different and minority.

Let us say that an edge is *fully colored* when both its ends are finally colored. A fully colored edge is *monocolored* if both ends have the same color and *bicolored* if both ends have different color. So the monocolored edges by Min greedy get bicolored 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 cubic graph is $1.5n$, we have proved Therorem 2.

4 The comb swapping algorithm

In this section we include a different approach that provides an independent verification of approximately the same bound as in the previous section. Its proof is just sketched, since it gives a slightly weaker result.

Let us recall that Robinson and Wormald proved that asymptotically almost all random cubic graphs are Hamiltonian [RW01]. The proof gave a contiguity result which permits us, for the purpose of proving any statement to be true a.a.s., to assume that a random cubic graph with n vertices (n even) is given as the union of a Hamiltonian cycle and perfect matching chosen u.a.r. (See [Wor99a].) Given such a graph G we can construct an initial bisection, by placing $n/2$ consecutive vertices in the cycle on one side, say the R side, and the remaining $n/2$ on the B side.

Given an arbitrary bisection (R, B) of G, define an *R-comb* to be a maximal path of vertices in R such that each is adjacent to at least one vertex in B, and such that this property also

Given a random cubic graph as a hamiltonian cycle + a perfect matching.
Construct an initial bisection (R, B), by breaking the Hamiltonian cycle in two disjoint paths of length $n-1$. Set $k = K$.
while $k \geq 2$
  while there is a compatible R-comb and B-comb of length k
    swap the two combs.
  endwhile
  $k = k - 1$
endwhile

Figure 3: The comb swapping algorithm for Min Bisection

holds in the initial bisection. In a symmetric way we can define a *B-comb*. The *length* of a comb is the length of the corresponding path. An R-comb and a B-comb of length k are *compatible* if there is no edge joining a vertex in one to a vertex in the other. To *swap* a compatible pair of combs, swap each vertex in the B-comb into R, and each vertex in the R-comb into B. Notice that by swapping two compatible combs of length k, we decrease the bisection size by $k-2$. We will consider the algorithm in Fig. 3, where $K > 2$ is a constant to be fixed later. At any step of the algorithm, let Y_i be a random variable counting the number of B-combs of length i ($i \leq k$). We will consider only the B point of view, as R has the same equations.

Initially, the expected bisection size is $n/4$, and the expected number of B-combs of length i is asymptotically $\mathbf{E}[Y_i] \sim 2^{-i-3}n$. So, at this point, the expected contribution to the bisection of combs with length bigger than d is $(1+o(1))\sum_{i=d+1}^{n} i2^{-i-3}n$. This tends quickly to 0 as $d \to \infty$, showing that our bounds will not improve much by considering very long combs.

Let k be fixed. During phase k, the maximum comb length is k. Let $s = \sum_{i=1}^{k} i\,Y_i$. This quantity corresponds (roughly) to the number of vertices in combs of one color. We compute the expected increment in Y_i by first considering the contribution due to one edge from a vertex in the red comb. This edge, e say, may subdivide, or split, a B-comb. Then the probability that e hits a B-comb of length i is $\frac{iY_i}{s}$ (arguing about randomness without justification in this sketch). For each B-comb of length $j > i$ there are two positions giving a splitting with length $i < j$, except when $j = 2i+1$, but in such a case we have two equally sized split portions. So, due to e, the expected increase in the number of combs with length i is

$$d_i = -\frac{iY_i}{s} + \sum_{j=i+1}^{k} \frac{2Y_j}{s}.$$

The total increase due to the $2k$ edges involved in one swap is therefore

$$\mathbf{E}[\Delta(Y_i)] = -\delta_{ik} + kd_i,$$

where δ denotes the Kronecker delta function, and the expected bisection decrease is $4-2k$. Phase k finishes when all pairs of combs of different colors are incompatible. It turns out that at this point, we may assume that the number of combs of length k is negligible.

We may write down an associated differential equation (as with the other method) and solve it numerically (we use a Runge-Kutta method), with initial conditions $y_i = 2^{-i-3}$, $1 \leq i \leq K$, $z = 1/4$ for the bisection width, and $k = K$ initially. When y_k hits 0, we can move to the next phase and make the initial values of phase $k-1$ equal to the final values of phase k. We select the value K such that the possible gain by treating the longer combs (as calculated above) is not significant. Taking $K = 24$ the solution of this system of equations, provides a bisection with size less than $0.17451n$ (and treating longer combs cannot reduce to $0.174507n$). So we have

Theorem 3 *The bisection width of a random cubic graph on n vertices is a.a.s. less than* $0.17451n$.

The same argument can be used to derive a lower bound for the maximum bisection. In this case we start with a partition obtained from the Hamiltonian cycle, but this time putting consecutive vertices in different parts of the bipartition. This provides an initial bisection of expected size $1.25n$. Define an equivalent notion of *anti-comb*, a maximal alternating path so that for every vertex v its third neighbor is in the same side as v. Then by swapping anti-combs with the same length we get the same gain (now positive) as when swapping combs. Furthermore the equations for both systems are the same, so we improve the initial bisection by the same amount, which gives the following.

Theorem 4 *The maximum bisection of a random cubic graph on n vertices is a.a.s. greater than* $1.32549n$.

References

[BCLS87] T. Bui, S. Chaudhuri, T. Leighton, and M. Sipser. Graph bisection algorithms with good average case behavior. *Combinatorica*, 7:171–191, 1987.

[Bol85] B. Bollobas. *Random Graphs.* Academic Press, London, 1985.

[DPPS01] J. Díaz, M. D. Penrose, J. Petit, and M. Serna. Approximating layout problems on random geometric graphs. *Journal of Algorithms*, 39:78–116, 2001.

[FR97] Alan Frieze and Mark R.Jerrum. Improved approximation algorithms for max k-cut and max bisection. *Algorithmica*, 18:61–67, 1997.

[GJ79] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* Freeman, San Francisco, 1979.

[JŁR00] S. Janson, T. Łuczak and A. Ruciński, *Random graphs*, Wiley, New York, 2000.

[KKL00] Marek Karpinski, Miroslaw Kowaluk, and Andrzej Lingas. Approximation algorithms for max-bisection on low degree regular graphs and planar graphs. Technical report, Department of Computer Science, University of Bonn, 2000.

[KM92] 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.

[RW01] R.W. Robinson and N.C. Wormald. Almost all cubic graphs are hamiltonian. *Random Structures and Algorithms*, 19:128–147, 2001.

[Wor95] N.C. Wormald, Differential equations for random processes and random graphs, *Ann. Appl. Probab.* **5** (1995), 1217–1235.

[Wor99a] N.C. Wormald. Models of random regular graphs. In *Surveys in Combinatorics*, pages 239–298. Cambridge University Press, 1999.

[Wor99b] 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.

[Wor02] N.C. Wormald, Analysis of greedy algorithms on graphs with bounded degrees, *Discrete Math.* (submitted).