Balanced Cut Approximation in Random Geometric Graphs ⋆

Josep Diaz[1], Fabrizio Grandoni[2], and Alberto Marchetti Spaccamela[3]

[1] Departament de Llenguatges i Sistemes Informatics, Universitat Politecnica de Catalunya, Campus Nord - Ed. Omega, 240 Jordi Girona Salgado, 1-3 E-08034 Barcelona, diaz@lsi.upc.edu

[2] Dipartimento di Informatica, Università di Roma "La Sapienza", via Salaria 113, 00198 Roma, Italy. grandoni@di.uniroma1.it

[3] Dipartimento di Informatica e Sistemistica, Università di Roma "La Sapienza", via Salaria 113, 00198, Roma, Italy, alberto.marchetti@dis.uniroma1.it

Abstract. A random geometric graph $\mathcal{G}(n, r)$ is obtained by spreading n points uniformly at random in a unit square, and by associating a vertex to each point and an edge to each pair of points at Euclidian distance at most r. Such graphs are extensively used to model wireless ad-hoc networks, and in particular sensor networks. It is well known that, over a critical value of r, the graph is connected with high probability. In this paper we study the robustness of the connectivity of random geometric graphs in the supercritical phase, under deletion of edges. In particular, we show that, for a sufficiently large r, any cut which separates two components of $\Theta(n)$ vertices each contains $\Omega(n^2r^3)$ edges with high probability. We also present a simple algorithm that, again with high probability, computes one such cut of size $O(n^2r^3)$. From these two results we derive a constant expected approximation algorithm for the β-balanced cut problem on random geometric graphs: find an edge cut of minimum size whose two sides contain at least $\beta\, n$ vertices each.

Keywords: ad-hoc networks, sensor networks, random geometric graphs, balanced cut, approximation algorithms.

1 Introduction

Let us consider a wireless network of sensors on a terrain, where the sensors communicate by radio frequency, using an omnidirectional antenna. Each sensor broadcasts with the same power to the same distance. Two sensors can communicate if and only if they are within the transmission radius of each other. Sensor

⋆ Partially supported by EU Integrated Project AEOLUS (FET-15964); the first author was partially supported by *La distinció de la Generalitat de Catalunya, 2002*, the second author by project ALGO-NEXT of the Italian Ministry of University and Research and the third author by project FIRB RBIN047MH9, Italy-Israel of the Ministry of University and Research.

networks, and more in general ad-hoc wireless networks, are often modelled via random geometric graphs [1,5]. A *random geometric graph* $\mathcal{G}(n,r)$ [9] is a graph resulting from placing a set V of n vertices uniformly at random on the unit square $[0,1]^2$, and connecting two vertices if and only if their Euclidean distance is at most the given radius r.

Random geometric graphs in general, and in particular their connectivity properties, have been intensively studied, both from the theoretical and from the empirical point of view. For the present paper, the most interesting result on random geometric graphs is the fact that, for $r = r(n) = \sqrt{(\ln n + c(n))/(\pi n)}$, for any $c(n)$ such that $c(n) \to \infty$ when $n \to \infty$, $\mathcal{G}(n,r)$ is connected whp [11,14,15]. (Throughout this paper, "whp" will abbreviate *with high probability*, that is with probability tending to 1 as n goes to ∞). Once the connectivity is achieved, it is natural to wonder how robust it is: how many edges one needs to remove in order to disconnect the graph? In most applications the disconnection of one vertex, or of a few vertices, does not affect significantly the behavior of the network. So we can reformulate the question above in the following more general way: given $\beta \in [0,1/2]$, how many edges one needs to remove in order to isolate two components (not necessarily connected) of $\beta\, n$ vertices each?

Our results. We can formalize the question above in the following way. A *cut* of a graph is a partition of its vertices into two subsets W and B, the *sides* of the cut. The *size* of cut (W,B) is the number of edges $\delta(W,B)$ between W and B. Given $\beta \in [0,1/2]$, $\beta\, n \in \mathbb{N}$, a β-balanced cut is a cut where both sides contain at least $\beta\, n$ vertices. The *β-balanced cut problem* is to compute a β-balanced cut of minimum size. Here we prove that, if $r = r(n) = \sqrt{R \ln n/n}$ for $R \geq R^*$, with $R^* > 0$ a sufficiently large constant, with high probability any β-balanced cut of $\mathcal{G}(n,r)$ has size $\Omega(\min\{\beta n R \log n, \sqrt{\beta n R^3 \log^3 n}\})$.

We also present a simple algorithm that with high probability computes a cut of size $O(\min\{\beta\, n\, R\, \log n,\ \sqrt{\beta\, n\, R^3 \log^3 n}\})$, thus matching the lower bound. The two mentioned results imply a probabilistic constant expected approximation algorithm for the β-balanced cut problem. We eventually show how to extend such result to a constant expected approximation algorithm.

We remark that the above results hold also if R is a function of n, and that the hidden constants in the O and Ω notations do not depend on n, R and β.

Related Work. Nothing is known on β-balanced cut approximation in random geometric graphs, for arbitrary values of β. For $\beta = 1/2$, the β-balanced cut problem is the well-know *minimum edge bisection problem*. Minimum edge bisection is a difficult problem which has received a lot of attention due to its numerous applications (see e.g. [10]). The problem is known to be NP-Hard for general graphs [8], and in such case there is a $O(\log^{1.5} n)$ approximation [6]. In the same paper, the authors prove that if the graph is planar, the approximation can be reduced to $O(\log n)$. If the input graph is dense, i.e. each vertex has degree $\Theta(n)$, there is a polynomial time approximation scheme (PTAS) for the minimum bisection problem [2]. In the case of random geometric graphs, it is known how to obtain a constant approximation whp for the special case

$R = R(n) \to \infty$ for $n \to \infty$ [3]. Our approximation algorithm improves on the algorithm in [3] in several ways: (i) it holds for arbitrary values of β, including the case $\beta = o(1)$; (ii) it holds for constant values of R as well; (iii) the value of the approximation ratio is constant in expectation, not only with high probability. We remark that each of the mentioned improvements is achieved by introducing new, simple techniques (which do not trivially follow from [3,13]).

One of the first papers to introduce the general problem of the minimum β-balanced cut was [4]. In this paper, the authors also show that given an $\epsilon > 0$, it is NP-hard to approximate the minimum bisection within an additive term of $n^{2-\epsilon}$. The β-balanced cut problem admits a PTAS for $\beta \le 1/3$, if every vertex has degree $\Theta(n)$ [2]. For planar graphs there is a 2-approximation for the β-balanced cut, if $\beta \le 1/3$ [7]. However, it is still open whether bisection and β-balanced cut are NP-hard for planar graphs.

Preliminaries. Given a region Q of the unit square, $|Q|$ denotes the area of Q, and $\|Q\|$ the number of points falling in Q. Note that $\|Q\|$ is a Binomial random variable of parameters n and $|Q|$, for which the following standard Chernoff's Bounds hold [12]. Let $\mu = E[\|Q\|] = |Q|\, n$. Then:

$$Pr[\|Q\| < (1-\delta)\mu] \le e^{-\delta^2\mu/2} \qquad \text{for } \delta \in [0,1); \tag{1}$$

$$Pr[\|Q\| > (1+\delta)\mu] \le e^{-\delta^2\mu/3} \qquad \text{for } \delta \in [0,1); \tag{2}$$

$$Pr[\|Q\| > (1+\delta)\mu] \le e^{-\delta^2\mu/4} \qquad \text{for } \delta \in [1,2e-1); \tag{3}$$

$$Pr[\|Q\| > (1+\delta)\mu] \le e^{-\delta\mu\ln 2} \qquad \text{for } \delta \ge 2e-1. \tag{4}$$

From now on $r = r(n) = \sqrt{R\ln n/n}$. For the sake of simplicity, we will assume $R = o(n/\log n)$. For $R = \Omega(n/\log n)$, the problems considered here become trivial. In particular, for $R \ge 2n/\ln n$ the graph is a clique (deterministically).

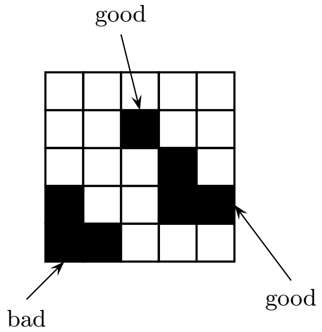
2 A Lower Bound

In this section we show that, for any $\beta \in [0,1/2]$, $\beta n \in \mathbb{N}$, and for $R \ge 240$, the size of any β-balanced cut is $\Omega(\min\{\beta nR\log n, \sqrt{\beta nR^3\log^3 n}\})$ with high probability.

In order to prove the mentioned lower bound, we consider a partition of the unit square into $5n/(R\ln n)$ non-overlapping square *cells* of the same size. Each cell is *adjacent* to the cells to its right, left, top, and bottom. Observe that, since the side of each cell has length $\sqrt{R\ln n/(5n)}$, a vertex is adjacent to all the vertices in the same cell and in all the adjacent cells. This property is crucial in the analysis. The number of points $\|C\|$ in each cell C satisfies the following probabilistic bounds.

Lemma 1. *For any $R \ge 240$, each cell C of the partition above contains $\|C\|$ vertices of $\mathcal{G}(n,r)$, $R\ln n/10 \le \|C\| \le 3\,R\ln n/10$, with probability $1 - o(1/n^2)$.*

Figure 1 Possible configuration of black and white cells. There are 3 black clusters and 1 white cluster.



Proof. Consider any cell C. Observe that $E[\|C\|] = R\ln n/5$. By Chernoff's Bounds (1) and (3),

$$Pr\left(\|C\| \notin \left[\frac{R\ln n}{10}, \frac{3\,R\ln n}{10}\right]\right) \leq e^{-(1/2)^2 R\ln n/10} + e^{-(1/2)^2 R\ln n/20} = O(1/n^3).$$

The claim follows by applying the union bound to the $O(n/(R\ln n))$ cells. ◻

Let (W, B) be any given cut, with $|W| = \beta\, n$. Let us call the vertices in W *white*, and the vertices in B *black*. A cell is *white* if at least one half of its points are *white*, otherwise the cell is *black*. We define a *cluster* $\mathcal{C}$ to be a maximal connected component of cells of the same color, with respect to the adjacency between cells defined above. The *frontier* $\partial\mathcal{C}$ of $\mathcal{C}$ is the subset of its cells which either touch the border of the unit square, or are adjacent to a cell of different color. We call *good* the cells of $\partial\mathcal{C}$ which are adjacent to a cell of different color, and *bad* the other cells of $\partial\mathcal{C}$. In particular, a cell is bad if it touches the border of the unit square and it is surrounded by cells of the same cluster (see Figure 1).

In order to prove the lower bound, we need the following two observations.

Lemma 2. *Given a cluster of k cells, its frontier contains at least $\sqrt{\pi\, k/4}$ cells.*

Proof. Suppose that the frontier contains $h < \sqrt{\pi\, k/4}$ cells. Thus the perimeter of the cluster has length at most $4hL$, where $L = \sqrt{R\ln n/(5n)}$ is the length of the side of one cell. Such perimeter can enclose an area of size at most $(4hL)^2/(4\pi)$ (case of a disk of radius $4hL/(2\pi)$), and thus at most $4h^2/\pi < k$ cells, which is a contradiction. ◻

Lemma 3. *Consider a cluster touching either* 0*, or* 1*, or* 2 *consecutive sides of the square. Then at least one third of the cells on its frontier are good.*

Proof. Consider any cluster $\mathcal{C}$. Without loss of generality, let $\mathcal{C}$ be white. If $\mathcal{C}$ does not touch any side of the square, all the cells of $\partial\mathcal{C}$ are good. Thus the claim is trivially true.

Now suppose $\mathcal{C}$ touches one or two consecutive sides of the square, say the left side and possibly the top side. Let $\partial\mathcal{C}_{good}$ be the good cells of $\partial\mathcal{C}$, and $\partial\mathcal{C}_{bad} = \partial\mathcal{C} \setminus \partial\mathcal{C}_{good}$ the bad ones. Moreover, let $\partial\mathcal{C}^{out}$ be the cells of $\partial\mathcal{C}$ touching the border of the square, and $\partial\mathcal{C}^{in} = \partial\mathcal{C} \setminus \partial\mathcal{C}^{out}$. Note that $\partial\mathcal{C}^{in} \subseteq \partial\mathcal{C}_{good}$ since the cells in $\partial\mathcal{C}^{in}$ do not touch any side of the square.

At least one half $\partial\mathcal{C}'$ of the cells of $\partial\mathcal{C}^{out}$ touches one between the left and the top side of the square, say the left one. Consider any cell $C' \in \partial\mathcal{C}'$. If C' is bad, we can univocally associate to C' a good cell $C'' \in \partial\mathcal{C}^{in}$ in the following way. Consider the sequence of consecutive white cells at the right of C' (there must be at least one such cell, since C' is bad). We let C'' be the rightmost of such cells. As a consequence, the number of good cells is lower bounded by $|\partial\mathcal{C}'|$, and $|\partial\mathcal{C}_{good}| \geq |\partial\mathcal{C}'| \geq |\partial\mathcal{C}^{out}|/2$. Thus

$$|\partial\mathcal{C}| = |\partial\mathcal{C}^{in}| + |\partial\mathcal{C}^{out}| \leq |\partial\mathcal{C}_{good}| + |\partial\mathcal{C}^{out}| \leq 3|\partial\mathcal{C}_{good}|.$$

The claim follows. □

Theorem 1. *With probability* $1 - o(1/n^2)$, *for any* $\beta \in [0, 1/2]$, $\beta n \in \mathbb{N}$, *and for any* $R \geq 240$, *the size of any* β*-balanced cut of* $\mathcal{G}(n, r)$ *is*

$$\Omega(\min\{\beta\, n\, R \log n,\ \sqrt{\beta\, n\, R^3 \log^3 n}\}).$$

Proof. By Lemma 1, with probability $1 - o(1/n^2)$ for each cell C,

$$\|C\| \in \left[\frac{R \ln n}{10}, \frac{3R \ln n}{10}\right]. \tag{5}$$

Thus it is sufficient to show that, given (5), the lower bound holds (deterministically) for any $\beta \in [0, 1/2]$ and for any cut (W, B) with $|W| = \beta\, n$.

We need some notation. By $\mathcal{W}$ and $\mathcal{B}$ we denote the set of white and black cells respectively. Moreover, $W_{black} \subseteq W$ ($B_{white} \subseteq B$) is the subset of white (black) vertices in black (white) cells.

Since each vertex is adjacent to all the other vertices in the same cell, each vertex $w \in W_{black}$ contained into a (black) cell C contributes with at least $\|C\|/2 \geq R \ln n/20$ edges to the edges of the cut. It follows that, if $|W_{black}| \geq |W|/2 = \beta\, n/2$, the size of the cut is at least

$$|W_{black}| \frac{R \ln n}{20} \geq \frac{\beta\, n\, R \ln n}{40} = \Omega(\beta\, n\, R \log n).$$

Analogously, if $|B_{white}| \geq |B|/2 = (1 - \beta)\, n/2$, then the size of the cut is at least

$$|B_{white}| \frac{R \ln n}{20} \geq \frac{(1 - \beta)\, n\, R \ln n}{40} = \Omega(\beta\, n\, R \log n).$$

Thus, let us assume $|W_{black}| < |W|/2$ and $|B_{white}| < |B|/2$. Note that, since all the vertices in adjacent cells are adjacent, each pair of adjacent (good) cells (C', C''), with $C' \in \mathcal{W}$ and $C'' \in \mathcal{B}$ contributes with at least

$$\frac{\|C'\|}{2} \frac{\|C''\|}{2} \geq \frac{R^2 \ln^2 n}{400} = \Omega(R^2 \log^2 n).$$

distinct edges to the total number of edges in the cut. Since there must be at least one such pair (C', C''), if $\beta = O(R \log n/n)$, trivially the size of the cut is $\Omega(R^2 \log^2 n) = \Omega(\beta\, n\, R\, \log n)$.

For $\beta = \Omega(R \log n/n)$ we need to bound the number of distinct pairs (C', C'') in a more sophisticated way. In particular, we will show that the number of good cells, either white or black, is $\Omega(\sqrt{\beta\, n/(R \log n)})$, from which it follows that the size of the cut is at least

$$\Omega(R^2 \log^2 n)\, \Omega(\sqrt{\beta\, n/(R \log n)}) = \Omega(\sqrt{\beta\, n\, R^3 \log^3 n}).$$

Observe that, from Equation (5) and from the assumption $|W_{black}| < |W|/2$ and $|B_{white}| < |B|/2$,

$$|\mathcal{W}| \geq \frac{\beta\, n/2}{3R \ln n/10} = \frac{5\beta\, n}{3R \ln n} \quad \text{and} \quad |\mathcal{B}| \geq \frac{(1-\beta)n/2}{3R \ln n/10} = \frac{5(1-\beta)n}{3R \ln n} \tag{6}$$

We distinguish three sub-cases, depending on the existence of white clusters with some properties.

(B.1) There is a white cluster $\mathcal{C}$ touching either 3 or 2 opposite sides of the square (but not 4). Without loss of generality, let the right side of the square be untouched. Consider all the cells of $\mathcal{C}$ which have no cell of the same cluster to their right. Note that such cells belong to the frontier $\partial\mathcal{C}$ of the cluster. Moreover, they are all good (they have a black cell to their right). The number of such cells is exactly $\sqrt{5n/(R \ln n)} = \Omega(\sqrt{\beta\, n/(R \log n)})$.

(B.2) Every white cluster touches 0, 1, or 2 consecutive sides of the square. Recall that the white cells are $|\mathcal{W}| \geq 5\beta\, n/(3R \ln n)$ by (6). Let $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_p$ be the p white clusters. It follows by Lemmas 2 and 3, that the total number of white good cells is at least

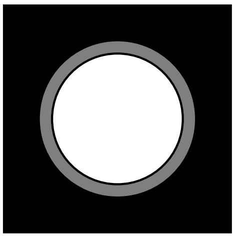
$$\sum_{i=1}^{p} \frac{1}{3}\sqrt{\frac{\pi\, |\mathcal{C}_i|}{4}} \geq \frac{1}{3}\sqrt{\frac{\pi\, |\mathcal{W}|}{4}} \geq \frac{1}{3}\sqrt{\frac{\pi\, 5\, \beta\, n}{12R \ln n}} = \Omega(\sqrt{\beta\, n/(R \log n)}).$$

(B.3) There is a white cluster touching the 4 sides of the square. It follows that each black cluster touches 0, 1, or 2 consecutive sides of the square. Thus, by basically the same argument as in case (B.2), the number of black good cells is at least

$$\frac{1}{3}\sqrt{\frac{\pi\, |\mathcal{B}|}{4}} \geq \frac{1}{3}\sqrt{\frac{\pi\, 5\, (1-\beta)\, n}{12R \ln n}} = \Omega(\sqrt{\beta\, n/(R \log n)}).$$

□

Figure 2 The white disk D contains one side W of the cut, $\|W\| = \beta n$. The annulus A of D, of width $\sqrt{R \ln n / n}$, is drawn in gray.



3 A Simple Cutting Algorithm

In this section we describe a simple algorithm `simpleCut` which, for a given input $\beta \in [0, 1/2]$, $\beta n \in \mathbb{N}$, computes a β-balanced cut. We will show that, for $R \geq 3/\pi$, the size of the cut computed is $O(\min\{\beta n R \log n, \sqrt{\beta n R^3 \log^3 n}\})$ with high probability. This, together with Theorem 1, implies that `simpleCut` is a probabilistic constant approximation algorithm for the β-balanced cut problem for $R \geq 240$. We later show how to convert such result into a constant expected approximation algorithm.

Algorithm 1 (`simpleCut`) *Take the βn vertices which are closest to* $(1/2, 1/2)$ *(breaking ties arbitrarily). Such vertices form one side W of the cut.*

Observe that `simpleCut` can be easily implemented in polynomial time.

In order to bound the size of the cut produced by `simpleCut`, we need the following simple probabilistic bound on the degree of the vertices.

Lemma 4. *For $R > 3/\pi$, the degree of each vertex of $\mathcal{G}(n, r)$ is upper bounded by* $(3\,\pi R \ln n)$ *with probability* $1 - o(1/n^2)$.

Proof. Consider the ball of radius $\sqrt{R \ln n / n}$ centered at vertex v, and let X_v be the number of vertices it contains. Clearly, the degree of v is $X_v - 1$. By denoting $\mu_v = E[X_v]$ we have

$$\pi\, R\, \ln n/4 \leq \mu_v \leq \pi\, R\, \ln n,$$

where the upper and lower bounds correspond to the case v is in the middle of the unit square and in one corner, respectively. By Chernoff's Bounds (2)-(4),

$$Pr(X_v > 3\pi R \ln n) \leq e^{-\ln 2 (3\pi R \ln n/\mu_v - 1)\mu_v} \leq e^{-\ln 2 (2\,\pi\, R \ln n)} = o(1/n^3).$$

Hence, from the union bound,

$$Pr(\exists v \in V : X_v > 3\pi R \ln n) \leq \sum_{v \in V} Pr(X_v > 3\pi R \ln n) = o(1/n^2).$$

$\square$

Theorem 2. *For any* $\beta \in [0, 1/2]$, $\beta\, n \in \mathbb{N}$, *and for* $R > 3/\pi$, *the size of the cut of* $\mathcal{G}(n, r)$ *computed by* `simpleCut` *is* $O(\min\{\beta\, n\, R \log n,\ \sqrt{\beta\, n\, R^3 \log^3 n}\,\})$ *with probability* $1 - o(1/n^2)$.

Proof. The upper bound $O(\beta\, n\, R \log n)$ trivially follows from Lemma 4. So, it is sufficient to show that for $\beta = \Omega(R \ln n/n)$, the size of the cut is $O(\sqrt{\beta\, n\, R^3 \log^3 n}\,)$. In particular, $\beta \geq 8\, \pi\, R\, \ln n/n$ is sufficient for our purposes.

Recall that, for a given region Q of the unit square, $|Q|$ denotes the area of Q, and $\|Q\|$ the number of points inside Q. Let us denote by D the disk centered in $(1/2, 1/2)$, of minimum possible radius ρ, which contains all the vertices in W (see Figure 2). In the following we will assume $\|D\| = \beta\, n$, which happens with probability one by standard probabilistic techniques.

Let A denote the annulus of width $\sqrt{R \ln n/n}$ surrounding D. The edges of the cut are a subset of the edges incident to the vertices in A. Hence, from Lemma 4, it is sufficient to show that the number $\|A\|$ of vertices in A is $O(\sqrt{\beta\, n\, R \log n})$ with probability $1 - o(1/n^2)$.

Consider the disk D' centered in $(1/2, 1/2)$ of radius $\rho' = \sqrt{(3/2)\beta/\pi}$, and let A' be the annulus of width $\sqrt{R \ln n/n}$ surrounding D'. Since $\rho' \leq \sqrt{3/(4\pi)} < 1/2$, for n large enough, D' and A' are entirely contained in the unit square.

Observe that, given $\rho \leq \rho'$, the density of points in both A and A' is the same, that is $(n - \beta\, n)/(1 - |D|)$. The density is maximized when $\rho = \rho'$. Note that whp, $\rho \neq \rho'$, but, for our purposes of getting and upper bound to the size of $||A'||$, the argument below is valid. Thus, for any $c > 0$,

$$Pr[\|A\| > c \,|\, \rho \leq \rho'] \ \leq\ Pr[\|A'\| > c \,|\, \rho \leq \rho'] \ \leq\ Pr[\|A'\| > c \,|\, \rho = \rho'].$$

For $\rho = \rho'$,

$$\frac{n - \beta\, n}{1 - |D|} = \frac{n - \beta\, n}{1 - 3\beta/2} \quad \text{and} \quad |A'| = \pi \sqrt{\frac{R \ln n}{n}} \left(2\sqrt{\frac{3\beta}{2\pi}} + \sqrt{\frac{R \ln n}{n}} \right).$$

Therefore

$$\mu = E[\|A'\| \,|\, \rho = \rho'] = \frac{n - \beta\, n}{1 - 3\beta/2}\, \pi \sqrt{\frac{R \ln n}{n}} \left(2\sqrt{\frac{3\beta}{2\pi}} + \sqrt{\frac{R \ln n}{n}} \right).$$

In particular

$$\sqrt{108}\, \ln n \leq \sqrt{(3/2)\pi\beta R n \ln n} \leq \mu \leq 12\sqrt{(3/2)\pi\beta R n \ln n}.$$

It follows from Chernoff's Bound (3) that

$$Pr[\|A'\| > 2\mu \,|\, \rho = \rho'] \leq e^{-\mu/4} \leq e^{-\sqrt{108/16} \ln n} = o(1/n^2).$$

Moreover, being $E[\|D'\|] = (3/2)\, \beta\, n$, from Chernoff's Bound (1),

$$Pr[\,\rho > \rho'\,] = Pr[\|D'\| < \beta\, n] \leq e^{-(1/3)^2 (3/2)\beta\, n/2} \leq e^{-\beta\, n/12} = o(1/n^2).$$

Altogether

$$\begin{aligned} Pr[\|A\| > 2\mu] &\leq Pr[\,\rho > \rho'\,] + Pr[\|A\| > 2\mu \,|\, \rho \leq \rho']\, Pr[\,\rho \leq \rho'\,] \\ &\leq o(1/n^2) + Pr[\|A'\| > 2\mu \,|\, \rho = \rho'] \\ &= o(1/n^2). \end{aligned}$$

It follows that $\|A\| \leq 2\mu = O(\sqrt{\beta\, n\, R \log n})$ with probability $1 - o(1/n^2)$. □

Theorems 1 and 2 imply that `simpleCut` is a probabilistic constant approximation algorithm for the β-balanced cut problem. We next show how to extend this result to a constant expected approximation algorithm for the same problem. Consider the following algorithm `zeroCut` to compute a cut of size zero, if any. Compute the connected components of $\mathcal{G}(n, r)$. For any integer m, $\beta n \leq m \leq n/2$, check whether there is a subset of components whose total size is m. If yes, return such subset of components as one side of the partition. Note that for each of the $O(n)$ possible values of m, we have to solve an instance of the *subset sum* problem. Since the sum of the sizes of the connected components is n, it follows that dynamic programming allows to solve all such instances in total time $O(n^2)$ and space $O(n)$ [8]. Combining `zeroCut` and `simpleCut`, one obtains the desired constant expected approximation algorithm.

Algorithm 2 (`refinedCut`) *If* `zeroCut` *returns a solution, return it. Otherwise, return the solution computed by* `simpleCut`*.*

Theorem 3. *For any $\beta \in [0, 1/2]$, $\beta\, n \in \mathbb{N}$, and for any $R \geq 240$,* `refinedCut` *is a constant expected approximation algorithm for the β-balanced cut problem on $\mathcal{G}(n, r)$.*

Proof. Let z^H and z^* denote the size of the solution found by `refinedCut` and the size of the optimum cut, respectively. Let moreover $\mathcal{A}$ denote the event that

$$z^* \geq c \min\{\beta\, n\, R\, \log n,\ \sqrt{\beta\, n\, R^3 \log n^3}\}$$

and

$$z^H \leq C \min\{\beta\, n\, R\, \log n,\ \sqrt{\beta\, n\, R^3 \log n^3}\},$$

where the constants c and C are as in the proofs of Theorems 1 and 2. Note that $Pr[\mathcal{A}] = 1 - o(1/n^2)$. Given $\mathcal{A}$, the approximation ratio of `refinedCut` is at most $C/c = O(1)$. Given $\overline{\mathcal{A}}$, if the size of the optimum cut is zero, `zeroCut` computes the optimum solution and the approximation ratio is 1 by definition. Otherwise, any cut, and hence also the cut computed by `simpleCut`, is a $O(n^2)$ approximation. Altogether the expected approximation ratio is

$$E(z^H/z^*) = Pr[\mathcal{A}]\, O(1) + Pr[\overline{\mathcal{A}}] O(n^2) = O(1).$$

□

Remark 1. The threshold 240 can be reduced to a value arbitrarily close to 30 by adapting the constants in Lemma 1. However, this would increase the approximation ratio. If we only desire a probabilistic constant approximation, such threshold can be made arbitrarily close to 10, with the same drawback as above.

Acknowledgments We thank an anonymous referee for the comments that have improved the presentation of the paper.

References

1. I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: a survey. *Computer Networks*, 38:393–422, 2002.
2. S. Arora, D. Karger, and M. Karpinski. Polynomial time approximation schemes for dense instances of NP-hard problems. In *ACM Symposium on the Theory of Computing (STOC)*, pages 284–293, 1995.
3. J. Díaz, M. Penrose, J. Petit, and M. Serna. Approximating layout problems on random geometric graphs. *Journal of Algorithms*, 39:78–116, 2001.
4. T.N. Bui, and C. Jones. Finding good approximate vertex and edge partitions is NP-hard. *Information Processing Letters*, 42:153–159, 1992.
5. J. Díaz, J. Petit, and M. Serna. Evaluation of basic protocols for optical smart dust networks. *IEEE Transactions on Mobile Networks*, 2:189–196, 2003.
6. U. Feige, and R. Krauthgamer. A polylogarithmic approximation of the minimum bisection. *SIAM Journal on Computing*, 31(3): 1090–1119, 2002.
7. G. Garg, H. Saran, and V. Vazirani. Finding separator cuts in planar graphs within twice the optimal. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 14–23, 1994.
8. M. Garey, and D. Johnson. *Computers and Intractability.* Freeman. N.Y., 1979.
9. E. Gilbert. Random plane networks. *Journal of the Society for Industrial and Applied Mathematics*, 9:533–543, 1961.
10. S.H. Gerez. *Algorithms for VLSI design automation.* Wiley, 2003.
11. A. Goel, S. Rai, and V. Krishnamachari. Sharp thresholds for monotone properties in random geometric graphs. In *ACM Symposium on Foundations of Computer Science (FOCS)*, pages 13–23, 2004.
12. R. Motwani and P. Raghavan. *Randomized Algorithms.* Cambridge University Press, 1995.
13. S. Muthukrishnan and G. Pandurangan. The Bin-covering technique for thresholding random geometric graph properties graphs. In *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 989–998, 2005.
14. M. Penrose. The longest edge of the random minimal spanning tree. *The Annals of Applied Probability*, 7(2):340–361, 1997.
15. M. Penrose. *Random Geometric Graphs.* Oxford Studies in Probability. Oxford U.P., 2003.