Alcom-FT Technical Report Series
ALCOMFT-TR-03-84

Approximating Layout Problems on Random Geometric Graphs *

Josep Díaz[†] Mathew D. Penrose[‡] Jordi Petit[†] María Serna[†]

Abstract

In this paper, we study the approximability of several layout problems on a family of random geometric graphs. Vertices of *random geometric graphs* are randomly distributed on the unit square and are connected by edges whenever they are closer than some given parameter. The layout problems that we consider are: *Bandwidth*, *Minimum Linear Arrangement*, *Minimum Cut Width*, *Minimum Sum Cut*, *Vertex Separation* and *Edge Bisection*. We first prove that some of these problems remain **NP**-complete even for geometric graphs. Afterwards, we compute lower bounds that hold, almost surely, for random geometric graphs. Then, we present two heuristics that, almost surely, turn to be constant approximation algorithms for our layout problems on random geometric graphs. In fact, for the Bandwidth and Vertex Separation problems, these heuristics are asymptotically optimal. Finally, we use the theoretical results in order to empirically compare these and other well-known heuristics.

*This research was partially supported by the IST Programme of the EU under contract number IST-1999-14186 (ALCOM-FT). This research used the computing facilities of the Centre Europeu de Paral·lelisme de Barcelona.

†Departament de Llenguatges i Sistemes Informàtics. Universitat Politècnica de Catalunya. Campus Nord C6. c/ Jordi Girona 1-3. 08034 Barcelona (Catalonia). {diaz,jpetit,mjserna}@lsi.upc.es

‡Department of Mathematical Sciences, University of Durham, South Road, Durham DH1 3LE, England. Mathew.Penrose@durham.ac.uk

1 Introduction

Several well-known optimization problems on graphs can be formulated as *graph layout problems*. Given a graph with n vertices, a *layout* on the graph is a bijection between the vertex set and the set of naturals from 1 to n. Graph layout problems, also denoted in the literature as *linear ordering problems* or *linear arrangement problems*, seek for a layout that minimizes a measure associated with each problem. The particular layout problems that we consider include *Bandwidth*, *Minimum Linear Arrangement*, *Minimum Cut Width*, *Minimum Sum Cut* and *Vertex Separation*. We also consider the *Bisection* problem, which is a partitioning problem, but can also be treated as a layout problem. The definition of these problems is given in Section 2.

For general graphs, all these problems are **NP**-hard. All of them have a long history behind them, owing to their practical relevance in different applications. For instance, the Bandwidth problem is an important problem in matrix theory and is very much related to the dilation of edges in interconnection networks [16]. The Minimum Linear Arrangement problem, also called the *Edgesum* problem, is relevant in circuit and VLSI layout [72, 35], single machine job scheduling [1, 70] and as a simplified model for nervous system simulation [54]. The Minimum Cut Width problem was first used as a theoretical model for the number of channels in an optimal layout of a circuit [53, 2]. More recent applications of the problem include network reliability [42], automatic graph drawing [60], information retrieval [12] and as a subroutine for the cutting plane algorithm to solve the TSP [40]. The Minimum Sum Cut problem, also known as the *Profile* problem, is equivalent to the *Interval Graph Completion* problem, and has well-known applications in archaeology [47, 70] and clone fingerprinting [44]. Because of its importance in the design of divide-and-conquer algorithms for network problems and graph drawing, the Edge Bisection problem has received a lot of attention [8, 51, 61]. In the present paper we shall deal with the simpler non-weighted versions of those problems. An important characteristic of these applications is the specificity of their instances: for most applications (in particular, in routing and circuit design), input graphs tend to have some *geometrical structure* and are likely to be *sparse*.

Often, the literature has been concerned with algorithms for layout and partitioning problems using *random graphs*. The model of random graphs, denoted $\mathcal{G}_{n,p}$, has two parameters: n represents the number of nodes and p represents the probability of the existence of each possible edge. Random graphs have received much attention and together with the probabilistic method have become a powerful tool in combinatorics [3, 10, 38]. Partitioning properties for $\mathcal{G}_{n,p}$ graphs are studied in [11]. Layout problems for sparse random graphs are studied in [24], where it is shown that for each one of the considered layout measures, with high probability, the ratio between the maximum and minimum values is a constant.

Therefore, the $\mathcal{G}_{n,p}$ model does not provide an informative framework to compare heuristics for layout problems. *Geometric graphs* or *disk graphs* [13, 18] have been proposed as a possible model to take into account the structural characteristics of instances that appear in most of the practical applications. Disk graphs are intersection graphs of disks in the plane. *Random geometric graphs* are graphs whose n vertices are n points uniformly distributed in the unit square and whose edges between any pair of distinct nodes exist when their distance is smaller than some parameter r.

Many empirical studies have used random models of geometric graphs for layout or partitioning problems [39, 7, 49, 67]. However, the theoretical study of random geometric graphs has been mainly focused on parameters as their clique or chromatic number and their connectivity properties (see [22] for a survey). Therefore, the analysis of layout measures in random geometric graphs certainly seems worthwhile to study.

In this paper, we are concerned with the complexity and approximability of several layout problems on geometric graphs. Regarding complexity, Theorem 3.1 states that some of the layout problems remain **NP**-hard even for *unit disk graphs* or *grid graphs*. Regarding approximability, we consider a family of random geometric graphs and prove lower bounds for the layout problems on these graphs. Leading up to these, Propositions 4.1 and 4.2 contain some isoperimetric inequalities, which may be of some interest in their own right. Theorems 5.1 and 5.2 state the actual lower bounds. To obtain upper bounds, we introduce two simple heuristics and derive asymptotics for their costs. Combined with the lower bounds, these show that either of these heuristics turn to be constant approximation algorithms for layout problems on the considered family of random geometric graphs (Theorems 6.4 and 6.5). In the case of the Bandwidth and Vertex Separation problems, the solutions returned by either of our algorithms are asymptotically optimal. In these cases, our result is an analog of the seminal BHH theorem [6] on the TSP tour of a random set of points in the unit square (Theorem 6.3). For the remaining problems, the approximation factor of the values provided by the two algorithms are tight. We emphasize that all our approximability results hold for random geometric graphs in the sense of convergence with probability 1. Much work on optimization problems on random sets of points use geometric sub-additivity as a key property to prove convergence results [73, 76]. It is important to stress that geometric sub-additivity does not hold for our layout problems and this forces us to take another approach.

We use our new results on random geometric graphs to give empirical evidence of the goodness of several well-known heuristics for layout and partitioning problems. These heuristics include global methods (such as spectral, multilevel and greedy methods) and local methods (such as simulated annealing or Kernighan-Lin) as well as the algorithms presented in this paper. We close the paper with some concluding remarks.

2 Definitions and basic results

Given as input an undirected graph $G = (V, E)$ with $n = |V|$, a *layout* φ on G is a one-to-one function $\varphi : V \to \{1, \dots, n\}$. Given a layout φ on G, let us define the sets

$$L(i, \varphi, G) = \{u \in V(G) \ : \ \varphi(u) \leq i\} \quad \text{and} \quad R(i, \varphi, G) = \{u \in V(G) \ : \ \varphi(u) > i\}.$$

We are interested in the following measures:

$$\begin{aligned}
\theta(i, \varphi, G) &= |\{uv \in E(G) \ : \ u \in L(i, \varphi, G) \wedge v \in R(i, \varphi, G)\}| \\
\delta(i, \varphi, G) &= |\{u \in L(i, \varphi, G) \ : \ \exists v \in R(i, \varphi, G) : uv \in E(G)\}| \\
\lambda(uv, \varphi, G) &= |\varphi(u) - \varphi(v)| \qquad \text{where } uv \in E(G).
\end{aligned}$$

The problems we consider are the following:

- Bandwidth (Bandwidth): Given a graph $G = (V, E)$, find $\textsc{minbw}(G) = \min_\varphi \textsc{bw}(\varphi, G)$ where $\textsc{bw}(\varphi, G) = \max_{uv \in E} \lambda(uv, \varphi, G)$.
- Minimum Linear Arrangement (MinLA): Given a graph $G = (V, E)$, find $\textsc{minla}(G) = \min_\varphi \textsc{la}(\varphi, G)$ where $\textsc{la}(\varphi, G) = \sum_{uv \in E} \lambda(uv, \varphi, G) = \sum_{i=1}^n \theta(i, \varphi, G)$.
- Minimum Cut Width (Cutwidth): Given a graph $G = (V, E)$, find $\textsc{mincw}(G) = \min_\varphi \textsc{cw}(\varphi, G)$ where $\textsc{cw}(\varphi, G) = \max_{i=1}^n \theta(i, \varphi, G)$.
- Vertex Separation (VertSep): Given a graph $G = (V, E)$, find $\textsc{minvs}(G) = \min_\varphi \textsc{vs}(\varphi, G)$ where $\textsc{vs}(\varphi, G) = \max_{i=1}^n \delta(i, \varphi, G)$.
- Minimum Sum Cut (SumCut): Given a graph $G = (V, E)$, find $\textsc{minsc}(G) = \min_\varphi \textsc{sc}(\varphi, G)$ where $\textsc{sc}(\varphi, G) = \sum_{i=1}^n \delta(i, \varphi, G)$.
- Edge Bisection (EdgeBis): Given a graph $G = (V, E)$, find $\textsc{mineb}(G) = \min_\varphi \textsc{eb}(\varphi, G)$ where $\textsc{eb}(\varphi, G) = \theta(\lfloor n/2 \rfloor, \varphi, G)$.

The previous definitions are stated as non-constructive problems, where the answer is the value of the optimum, instead of the layout itself. Except in Sections 5 and 6, we deal with bounds of the optimal values. In Table 1 the reader can find a survey of known complexity results for each of these problems (the formal definitions of the classes can be found, for instance, in [63, 25]).

Before going further, let us review some basic definitions from probability theory (see e.g. [17]). Given a sequence of random variables $(X_n)_{n \geq 1}$ and a random variable X, we say that X_n *converges in probability to* X if $\lim_{n\to\infty} \mathbf{Pr}[|X_n - X| > \epsilon] = 0$ for any $\epsilon > 0$. We say that X_n *converges with probability* 1 *to* X if $\mathbf{Pr}[\lim_{n\to\infty} X_n = X] = 1$. This type of

Problem	Negative results	Positive results
MinLA	**NP-C** [31] **NP-C** for bipartite graphs [30]	**P** for trees, hypercubes, meshes [71, 34, 59, 54] **NC** for trees [19] $O(\log n)$ approximable [69] $O(\log \log n)$ approx. for planar graphs [69] **PTAS** for dense graphs [5]
SumCut	**NP-C** [30]	**P** for trees [23] **NC** for trees [23]
Bandwidth	**NP-C** [62] **NP-C** for trees $\Delta = 3$ [29], caterpillars with hair length 3 [55] no **PTAS** for trees [9] no **APX** in general [45]	**APX** for certain trees [33] **APX** for dense graphs [46]
Cutwidth	**NP-C** [32] **NP-C** for pl. graphs $\Delta = 3$ [57]	**P** for trees [75] **NC** for trees [19] **APX** for dense graphs [5]
VertSep	**NP-C** [52] **NP-C** for pl. graphs $\Delta = 3$ [57]	**P** for trees [28]
EdgeBis	**NP-C** [31]	**P** for trees (ref. in [15]) **P** for grid graphs without holes [64] **PTAS** for planar graphs [15] **NCAS** for planar graphs [26]

P Solvable in polynomial time.
NP Solvable in non-deterministic polynomial time.
NP-C **NP**-complete.
NC Solvable in polylog time with a polynomial number of processors.
APX Approximation algorithm in polynomial time.
PTAS Has an approximation scheme in polynomial time.
NCAS Has an **NC** approximation scheme.

Table 1: Layout problems: Review of complexity results.

convergence is also called *convergence almost surely*. Convergence almost surely is stronger than convergence in probability.

We now introduce several classes of geometric graphs on the plane. These graphs depend on which kind of norm is used to measure distances. Under the l_p norm ($p \geq 1$), the distance $\|x-y\|_p$ between two points $x = (x_1, y_1)$ and $y = (x_2, y_2)$ is $(|x_1 - x_2|^p + |y_1 - y_2|^p)^{1/p}$. Under the l_∞ norm, their distance is $\max\{|x_1 - x_2|, |y_1 - y_2|\}$. The l_2 metric is usually called the Euclidean metric, and the l_1 metric is sometimes called the Manhattan distance. For sake of simplicity, we use the l_∞ norm all through this paper. Whenever we use other norms, we shall make it explicit.

A graph is a *unit disk graph* if each vertex can be mapped to a closed, unit diameter disk in the plane such that two vertices are adjacent if and only if their corresponding disks intersect. A graph is a *grid graph* if it is a node-induced finite subgraph of the infinite grid. Observe that grid graphs are unit disk graphs in any l_p norm. See Figure 1 for the mapping of the grid nodes in l_1, l_2 and l_∞.

Let r be a positive number and let V be any set of n points in the unit square ($[0,1]^2$). A *geometric graph* $G(V, r)$ with vertex set V and radius r is the graph $G = (V, E)$ where $E = \{uv \mid u, v \in V \wedge 0 < \|u - v\| \leq r\}$. An appropriate scaling shows that a geometric graph is also a unit disk graph.

Let $(r_i)_{i \geq 1}$ be a sequence of positive numbers and let $X = (X_i)_{i \geq 1}$ be a sequence of independently and uniformly distributed (i.u.d.) random points in $[0,1]^2$. For any natural n, we write $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ and call $G(\mathcal{X}_n, r_n)$ a *random geometric graph* of n nodes on X.

In the remainder of this paper we restrict our attention to the particular case where the radius is of the form

$$r_n = \sqrt{\frac{a_n}{n}} \quad \text{where} \quad r_n \to 0 \quad \text{and} \quad a_n / \log n \to \infty.$$

It is important to remark that through this choice, the construction of sparse but connected graphs is guaranteed: Defining the connectivity distance $\rho_n = \rho_n(\mathcal{X}_n)$ as the smallest radius r such that the graph $G(\mathcal{X}_n, r)$ is connected, it is known [4] that $\rho_n \sqrt{n/\log n}$ converges to $\frac{1}{2}$ almost surely; that is, with probability 1 the sequence $(X_i)_{i \geq 1}$ yields a sequence $(\rho_n)_{n \geq 1}$ with this limit behavior.

For results on other cases, for example with a_n constant, see [20, 65]. The parameter $a_n = n r_n^2$ can be thought of as the local density of points on the scale at which they are interacting.

3 Complexity results for geometric instances

In this section, we shall consider the decisional counterparts of the optimization problems previously defined. We shall use K as the constant parameter in the definition of the decisional versions. We show that some of the layout problems under consideration remain hard to solve efficiently, even when restricted to geometric instances. All the proofs in this section are valid for any l_p norm ($1 \leq p \leq \infty$), but for readability purposes, the explanation and the figures use the Euclidean norm.

Definition 3.1 (Subdivision and homeomorphism). A graph H is a *subdivision* of a graph G if H can be constructed from G by subdividing some of its edges, that is, replacing an edge by a path of nodes with degree 2. Two graphs are *homeomorphic* if they are subdivisions of the same graph.

Lemma 3.1. Let H be a subdivision of a graph G. Then $\textsc{mincw}(G) = \textsc{mincw}(H)$.

Proof. We prove the statement of the lemma for the insertion of a new node w between an edge uv of G. Let φ be an optimal layout of G; without loss of generality assume that $\varphi(u) < \varphi(v)$. Let ψ be a layout of H that corresponds to inserting w just after u. Then, $\textsc{cw}(G, \varphi) = \textsc{cw}(H, \psi)$. To show that ψ is optimal for H, suppose the contrary; then there exists a layout ψ' of H such that $\textsc{cw}(H, \psi') < \textsc{cw}(H, \psi)$. Let φ' be the layout obtained by removing w from ψ'. Then $\textsc{cw}(G, \varphi') \leq \textsc{cw}(H, \psi') < \textsc{cw}(H, \psi) \leq \textsc{cw}(G, \varphi)$ which is a contradiction to the optimality of φ. $\square$

Recall that the Bandwidth problem remains **NP**-complete even when restricted to caterpillars with at most one hair attached to each vertex of the body [55]. Also, Cutwidth and VertSep remain **NP**-complete even when restricted to planar graphs with maximum vertex degree 3 [56].

Theorem 3.1. Bandwidth, Cutwidth and VertSep remain **NP**-complete even when restricted to grid graphs (and therefore, even when restricted to unit disk graphs).

Proof. Bandwidth remains **NP**-complete even when restricted to grid graphs because caterpillars with at most one hair attached to each vertex of the body are grid graphs. See Figure 2.

We present a reduction from the Cutwidth problem restricted to planar graphs with maximum vertex degree 3 to the Cutwidth problem restricted to grid graphs. Let $\langle G, K\rangle$ be an instance of Cutwidth restricted to planar graphs with maximum vertex degree 3. Using the algorithm of Valiant [74], we can draw G in such a way that its nodes are located at positions $(6x, 6y)$ for some $x, y \in \mathbb{N}$ and the edges only follow horizontal and vertical paths without crossing. This embedding uses an area polynomial in the size of G. Then, replace

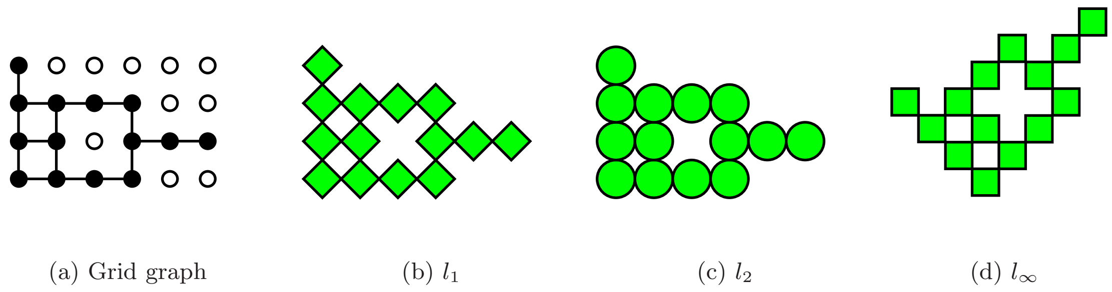

Figure 1: Any grid graph is a unit disk graph in any l_p norm.

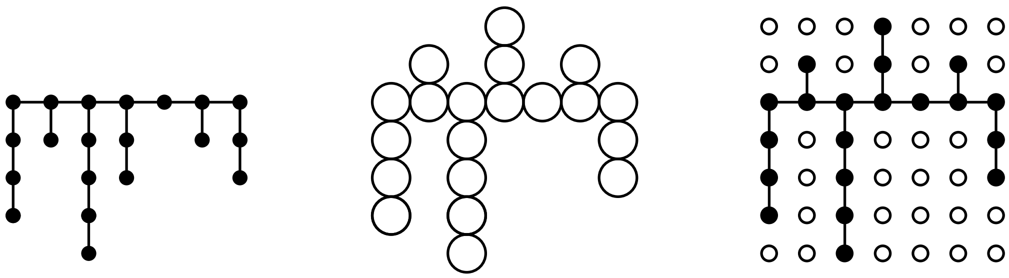

Figure 2: Any caterpillar with at most one hair attached to each vertex of the body is a unit disk graph and a grid graph.

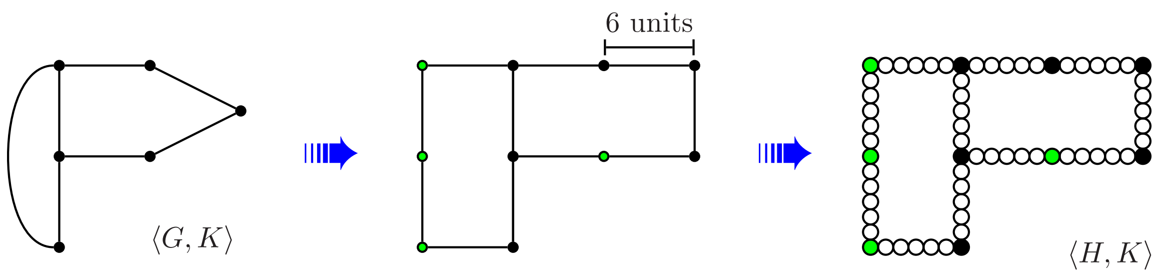

Figure 3: Reduction from CUTWIDTH restricted to planar graphs with maximum vertex degree 3 to CUTWIDTH restricted to grid graphs. Left, the input graph; center, the input graph embedded with Valiant's algorithm; right, substitution of the edges with paths of disks.

each edge by a string of unit disks to produce a grid graph H (see Figure 3). By construction H is a subdivision of G, therefore we have that $\text{MINCW}(G) \leq K \Leftrightarrow \text{MINCW}(H) \leq K$, which proves the claimed result.

Observe that the previous reduction creates graphs with maximum degree 3 and recall that for graphs with maximum degree 3, the SEARCHNB problem (whose measure is SN) is identical to the CUTWIDTH problem [53]. Therefore, we get as corollary that SEARCHNB remains **NP**-complete even when restricted to grid graphs.

For any graph G, the vertex separation of a homeomorphic image of G is identical to the search number of G [28]. Let us reduce SEARCHNB restricted to planar graphs with maximum vertex degree 3 to VERTSEP restricted to grid graphs using the same transformation that we used for CUTWIDTH. The resulting graph H is a grid graph homeomorphic to the input graph G, so we get $\text{MINVS}(H) \leq K \Leftrightarrow \text{MINSN}(G) \leq K$. □

We could not obtain similar results for SUMCUT, MINLA and EDGEBIS. However, for the EDGEBIS problem, we are able to give a weak result:

Theorem 3.2. If EDGEBIS is **NP**-complete even when restricted to planar graphs with maximum vertex degree 4, then EDGEBIS is **NP**-complete even when restricted to unit disk graphs.

Proof. Let $\langle G, K\rangle$ be an instance of EDGEBIS where G is a planar graph with an even number n of nodes and maximum vertex degree 4. Without loss of generality, we assume that n is even (otherwise, add a disconnected node to G). We will reduce $\langle G, K\rangle$ to another instance $\langle H, K\rangle$ of EDGEBIS where H is a unit disk graph such that $\text{MINEB}(G) = \text{MINEB}(H)$.

As in the proof of Theorem 3.1, using Valiant's algorithm [74], we embed G in the plane in such a way that its nodes are located at positions $(6x, 6y)$ for some $x, y \in \mathbb{N}$ and that the edges follow horizontal and vertical paths without crossing. Afterwards, we identify each "original" node of the embedding with a unit disk and replace each half edge of length l with a string of disks of length $\lfloor l/2 \rfloor$. As edges had an odd length, we must join the strings using two additional "extremal" disks as shown in Figure 4. Therefore, each edge has been replaced by an even number of disks. For each original node u in $V(G)$, define its "gadget" as the set of disks that represent its adjacent half edges. Notice that a gadget includes the extreme disks, where it ends. We give to each non extreme disk multiplicity n^2, while extreme disks retain multiplicity 1, and add multiplicity to the original nodes in such a way that every gadget receives the same number of disks. Let H be the resulting graph of this transformation, where disks with multiplicity m are m different disks on the same position forming a clique. The transformation can be computed in polynomial time

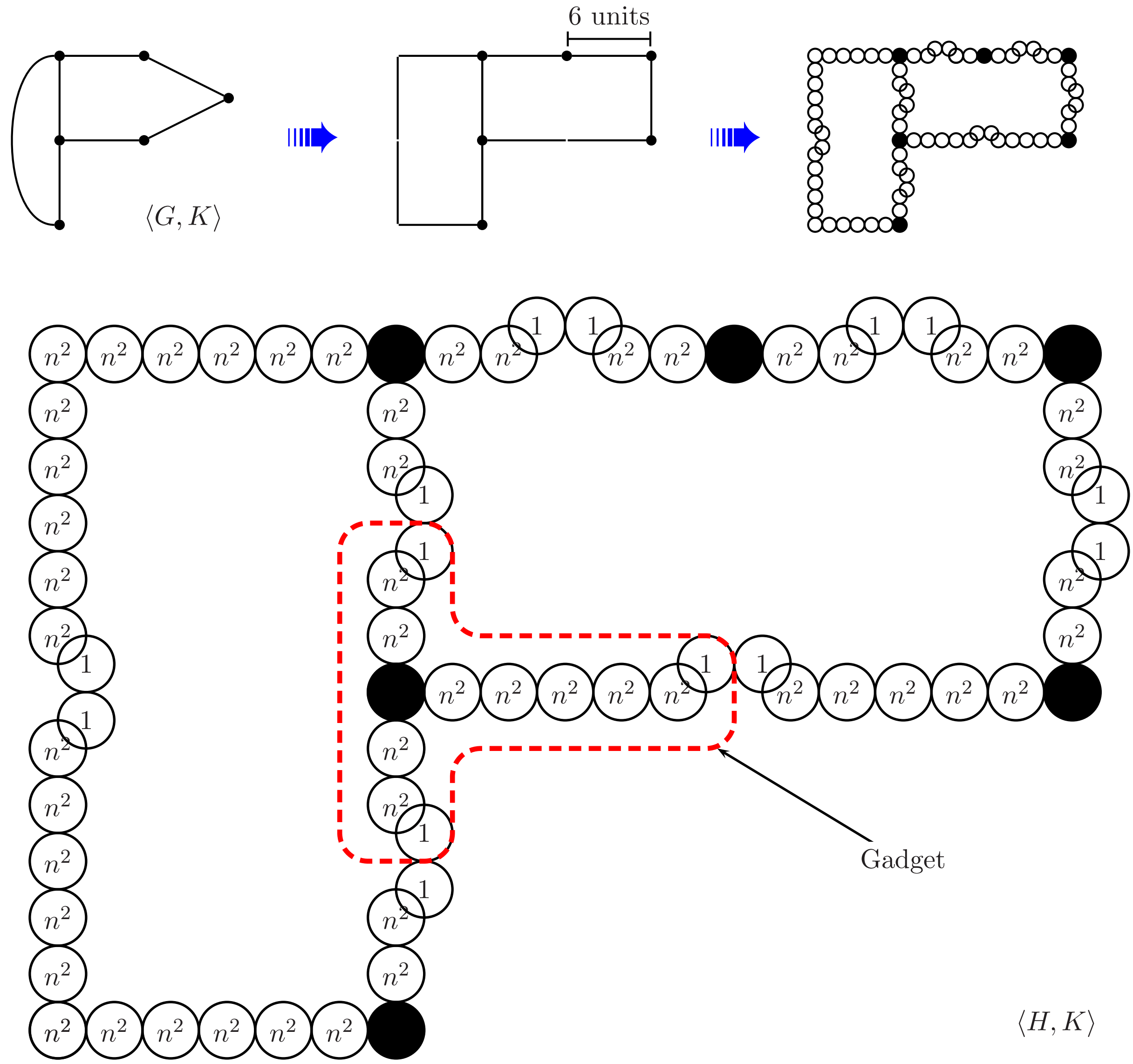

Figure 4: Reduction from EdgeBis restricted to planar graphs with maximum vertex degree 4 to EdgeBis restricted to unit disk graphs. Top left, the input graph with $n = 6$ nodes; top center, the input graph embedded with Valiant's algorithm; top right, substitution of the edges with paths of disks with even length. Bottom, we show how non extreme disks receive multiplicity n, extreme disks get multiplicity 1 and (not shown) original nodes receive the required multiplicity in order to ensure that all the gadgets contain the same number of disks.

We have to prove that $\langle G, K\rangle$ is a positive instance of EdgeBis if and only if $\langle H, K\rangle$ is also a positive instance of EdgeBis. We do so by showing that gadgets in H behave as the original nodes in G.

If $\langle G, K\rangle$ is a positive instance of EdgeBis then there exists a bisection B of G such that $\text{EB}(G, B) \leq K$. Coloring each gadget of H according to B, the bisection of H coincides with the bisection B and is a legal bisection (each gadget has the same number of nodes). Therefore $\langle H, K\rangle$ is a positive instance of EdgeBis.

On the other hand, if $\langle H, K\rangle$ is a positive instance of EdgeBis, we have two cases: When $K > 2n$, as G has maximum degree 4, the bisection width of G cannot exceed $2n$, thus $\langle G, K\rangle$ surely is a positive instance of EdgeBis. For the case $K \leq 2n$, let us consider any gadget. Each of the nodes of this gadget must be on the same side of the bisection (otherwise, the bisection width would be larger than $2n$ because of the cliques of size n^2 introduced in H). Taking a bisection of G that coincides with the one given to the gadgets of H, we get that $\langle G, K\rangle$ is a positive instance of EdgeBis. $\square$

4 Isoperimetric inequalities

In this section we prove several isoperimetric inequalities that will be used in the next sections. We start presenting an isoperimetric inequality on square meshes with additional diagonal connections. Let (A, B) be a partition of $\{1, \ldots, m\}^2$ for some integer m. Let $\partial_{A,B}$ be the number of elements of $A \times B$ that are neighbor pairs, including diagonal neighbors, that is, $\partial_{A,B} = |\{(x, y) \mid x \in A \wedge y \in B \wedge \|x - y\|_\infty = 1\}|$.

Proposition 4.1. For any integer $m \geq 3$ and any partition (A, B) of $\{1, \ldots, m\}^2$, it holds that $\partial_{A,B} \geq 3 \min\left\{\sqrt{|A|}, \sqrt{|B|}\right\}$.

Proof. If A includes an entire row of elements, and B includes an entire row of elements, then each column includes a neighbor pair of elements, one from A and the other from B, which contributes at least 3 to $\partial_{A,B}$ (see Figure 5) except for the pair in the right-most column which contributes 1, so that $\partial_{A,B} \geq 3m - 2$. If B contains no entire row or column, and at least as many rows as columns have non-empty intersection with B, then there are at least $\sqrt{B}$ such rows, and each contains a neighbor pair from different sets which contributes at least 3 to $\partial_{A,B}$, so that $\partial_{A,B} \geq 3\sqrt{B}$. Applying similar arguments to the other possible cases, we have

$$\partial_{A,B} \geq \min\left\{3\sqrt{|A|},\ 3\sqrt{|B|},\ 3m - 2\right\}$$

and if $m \geq 3$ this minimum is always achieved at $3\sqrt{|A|}$ or at $3\sqrt{|B|}$. $\square$

Figure 5: Illustration for the proof of Proposition 4.1

Let us present now an isoperimetric inequality for sets in $\mathbb{R}^2$. In the following, $|\cdot|$ denotes Lebesgue measure, $A+B$ denotes $\{\mathbf{x}+\mathbf{y} : \mathbf{x} \in A \wedge \mathbf{y} \in B\}$, and B_r denotes the l_∞ ball of radius r: $B_r = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\|_\infty \le r\}$. For $A \subset \mathbb{R}^2$, let $\partial_r A$ denote the set $(A+B_r)\setminus A$, and let $\partial_{-r} A$ denote the set $\partial_r(\mathbb{R}^2 \setminus A)$.

Lemma 4.1. Suppose A is a compact subset of $\mathbb{R}^2$, and $r > 0$. Then

$$|\partial_r A| \ge 4|A|^{1/2} r \quad \text{and} \quad |\partial_{-r} A| \ge 4|A|^{1/2} r - 16r^2.$$

Proof. By the Brunn-Minkowski inequality [50],

$$|A+B_r| \ge (|A|^{1/2} + |B_r|^{1/2})^2 = (|A|^{1/2} + 2r)^2 \ge |A| + 4|A|^{1/2} r,$$

and the first inequality follows.

Set $A^o = A \setminus \partial_{-r} A$. Then $\partial_r(A^o) \subset \partial_{-r} A$ and hence,

$$|\partial_{-r} A| \ge |\partial_r A^o| \ge 4|A^o|^{1/2} r = \left(4\sqrt{|A| - |\partial_{-r} A|}\right) r.$$

When $|\partial_{-r} A| \ge 4|A|^{1/2} r$, we have the second inequality at once. Otherwise, we have

$$\begin{aligned} |\partial_{-r} A| &\ge 4r\left(|A| - 4|A|^{1/2} r\right)^{1/2} \ge 4r|A|^{1/2}(1 - 4r|A|^{-1/2})^{1/2} \\ &\ge 4r|A|^{1/2}\left(1 - 4r|A|^{-1/2}\right) \ge 4r|A|^{1/2} - 16r^2 \end{aligned}$$

and the the second inequality follows. $\square$

We are interested here in subsets of the unit square $[0,1]^2$. For $A \subset [0,1]^2$, let A_r denote the set $(A+B_r) \cap [0,1]^2$. Also let A^c denote the set $[0,1]^2 \setminus A$.

Proposition 4.2. Suppose A is a compact subset of $[0,1]^2$, and $r \in (0,1)$. Then

$$|A_r \setminus A| \ge \min(2r|A|^{1/2}, 2r|A^c|^{1/2} - 4r^2, r)$$

Proof. Let A' be the set in $[0,1]^2$ obtained by "pushing each vertical section of A down as far as possible towards the x-axis"; formally, setting $S_x(A) = \{y : (x,y) \in A\}$ for $x \in [0,1]$, let

$$A' = \bigcup_{x \in [0,1],\, S_x(A) \neq \emptyset} \{\{x\} \times [0, |S_x(A)|] : x \in [0,1]\}.$$

By Fubini's theorem, $|A'| = |A|$.

We claim that $|A'_r| \leq |A_r|$. Indeed, consider a 1-dimensional setup. For any $W \subset [0,1]$, let W_r be the r-neighborhood of W in $[0,1]$ and let $|W|$ be the 1-dimensional measure of W. Then, if $W \neq \emptyset$, we have $|W_r| \geq \min\{|W| + r, 1\}$ because if $[0,1] \setminus W$ includes an interval of length at least r then $|W_r| \geq |W| + r$, and otherwise $W_r = [0,1]$ and so $|W_r| = 1$. Going back to $A \subset [0,1]^2$, for $x \in [0,1]$, $S_x(A_r) = \cup_{u\in[0,1],\ |u-x|\leq r}(S_u(A))_r$. Assume $x \in [0,1]$ is such that $S_u(A) \neq \emptyset$ for some $u \in [0,1]$ with $|u - x| \leq r$. Then,

$$|S_x(A_r)| \geq \sup_{u\in[0,1],\ |u-x|\leq r} |(S_u(A))_r| \geq \sup_{u\in[0,1],\ |u-x|\leq r} \min\{|(S_u(A))| + r,\ 1\}.$$

With $x \in [0,1]$, $S_x(A'_r) = \cup_{u\in[0,1],\ |u-x|\leq r}[0, \min\{|S_u(A)| + r, 1\}]$. Hence,

$$|S_x(A'_r)| \leq \sup_{u\in[0,1],\ |u-x|\leq r} \min\{|(S_u(A))| + r,\ 1\} \leq |S_x(A_r)|.$$

If $x \in [0,1]$ and $S_u(A) = \emptyset$ for all $u \in [0,1]$ with $|u - x| \leq r$, then $|S_x(A_r)| = 0$. Therefore, for all $x \in [0,1]$, $|S_x(A'_r)| \leq |S_x(A_r)|$. By Fubini's theorem

$$|A_r| = \int_0^1 |S_x(A_r)|\, dx \quad \text{and} \quad |A'_r| = \int_0^1 |S_x(A'_r)|\, dx$$

and thus $|A'_r| \leq |A|$.

Let A'' be the set in $[0,1]^2$, obtained by "pushing each horizontal section of A' sideways as far as possible towards the y-axis", in an analogous manner to the construction of A' from A. Then $|A''| = |A'| = |A|$, and $|A''_r| \leq |A'_r| \leq |A_r|$. Moreover, A'' is a *down-set*, that is, it has the property that for any $\mathbf{x} \in A''$, all points of $[0,1]^2$ lying directly below or directly to the left of $\mathbf{x}$ are in A'.

From now on we assume that A is a down-set. Otherwise, it could be converted to a down-set A''. We consider four different cases:

Case 1: $(1-r, 0) \in A$ and $(0, 1-r) \notin A$. Then $S_x(A_r \setminus A)$ contains an interval of length at least r, for each $x \in [0,1]$, so that by Fubini's theorem, $|A_r \setminus A| \geq r$.

Case 2: $(1-r, 0) \notin A$ and $(0, 1-r) \in A$. Clearly in this case, $|A_r \setminus A| \geq r$ by an analogous argument to that just given.

Case 3: $(1-r, 0) \notin A$ and $(0, 1-r) \notin A$. In this case, set $A_1 = A$; let A_2 be the reflection of A in the x-axis. Let A_3 (respectively A_4) be the reflection of A_1 (respectively A_2) in the y-axis, and let $A_5 = \cup_{i=1}^4 A_i$. Then, by Lemma 4.1,

$$|A_r \setminus A| = (1/4)|\partial_r A_5| \geq |A_5|^{1/2} r = 2|A|^{1/2} r.$$

Case 4: $(1-r, 0) \in A$ and $(0, 1-r) \in A$. In this case, set $A_1 = A^c$; let A_2 be the reflection of A_1 in the line $y = 1$. Let A_3 (respectively A_4) be the reflection of A_1 (respectively A_2) in the line $x = 1$, and let $A_5 = \cup_{i=1}^4 A_i$. Then, by Lemma 4.1,

$$|A_r \setminus A| = (1/4)|\partial_{-r} A_5| \geq |A_5|^{1/2} r - 4r^2 = 2|A^c|^{1/2} r - 4r^2.$$

Since one of the four cases considered above must occur, we obtain the lemma. $\square$

5 Lower bounds

In this section we find lower bounds for the optimum cost of our various layout problems. As said, from now on we consider only (random) geometric graphs whose radius is of the form $r_n = \sqrt{a_n/n}$ where $r_n \to 0$ and $a_n/\log n = b_n \to \infty$. We first deal with the "edge problems" (EdgeBis, Cutwidth and MinLA) and then with the "node problems" (VertSep, SumCut and Bandwidth).

5.1 Edge problems

The following definition captures the property that nodes of a geometric graph are "nicely spread" on the unit square. The subsequent lemma is the only probabilistic result of this subsection.

Definition 5.1. Consider any set $\mathcal{X}_n$ of n points in $[0,1]^2$, which together with a radius r_n, induce a geometric graph $G = G(\mathcal{X}_n, r_n)$. Dissect the unit square into $4\lceil 1/r_n \rceil^2$ boxes of size $1/(2\lceil 1/r_n \rceil) \times 1/(2\lceil 1/r_n \rceil)$ placed packed in $[0,1]^2$ starting at $(0,0)$. By construction, all the boxes exactly fit in the unit square, and any two points of $\mathcal{X}_n$ in neighboring boxes (including diagonals) will be connected by an edge in G because $1/(2\lceil 1/r_n \rceil) \leq r_n/2$. Given $\epsilon \in (0,1)$, let us say that G is *ϵ-nice* if every box of this dissection contains at least $(1-\epsilon)\frac{1}{4}a_n$ points and at most $(1+\epsilon)\frac{1}{4}a_n$ points.

Lemma 5.1. Let $\epsilon \in (0, \frac{1}{5})$. Then, with probability 1, for all large enough n, random geometric graphs $G(\mathcal{X}_n, r_n)$ are ϵ-nice.

Proof. Choose a box in the dissection and let Y be the random variable counting the number of points of $\mathcal{X}_n$ in this box. As the points in $\mathcal{X}_n$ are i.u.d.,

$$\mathbf{E}[Y] = \frac{n}{4\lceil 1/r_n \rceil^2} \sim \frac{nr_n^2}{4} = \tfrac{1}{4}a_n.$$

Using Chernoff's bounds [58], we obtain

$$\begin{aligned}\mathbf{Pr}\left[Y \geq (1+\epsilon)\tfrac{1}{4}a_n\right] &\leq \mathbf{Pr}\left[Y \geq (1+\tfrac{1}{2}\epsilon)\mathbf{E}[Y]\right] \leq \exp\left(-\left(\tfrac{1}{2}\epsilon\right)^2 \mathbf{E}[Y]/3\right) \\ &\leq \exp\left(-\tfrac{1}{13}\epsilon^2 \tfrac{1}{4}a_n\right) = n^{-\epsilon^2 b_n/52}\end{aligned}$$

and

$$\mathbf{Pr}\left[Y \le (1-\epsilon)\tfrac{1}{4}a_n\right] \le \mathbf{Pr}\left[Y \le (1-\tfrac{1}{2}\epsilon)\mathbf{E}\left[Y\right]\right] \le \exp\Big(-\left(\tfrac{1}{2}\epsilon\right)^2 \mathbf{E}\left[Y\right]/2\Big)$$
$$\le \exp\left(-\tfrac{1}{9}\epsilon^2\tfrac{1}{4}a_n\right) = n^{-\epsilon^2 b_n/36} \le n^{-\epsilon^2 b_n/52}.$$

The number of boxes is certainly smaller than n, so by Boole's inequality [17], which says that $\mathbf{Pr}\left[\cup_n \mathcal{E}_n\right] \le \sum_n \mathbf{Pr}\left[\mathcal{E}_n\right]$ for any finite or countable collection of events $\mathcal{E}_n$, the probability that for some box the number of points in the box is less than $(1-\epsilon)\frac{1}{4}a_n$ or bigger than $(1+\epsilon)\frac{1}{4}a_n$, is bounded by $2n^{1-b_n\epsilon^2/52}$, which is summable in n because $b_n \to \infty$. The result follows by the Borel-Cantelli lemma [17]. □

The following lemma is the basis of our lower bounds for nice graphs. In the next section we will show these lower bounds are sharp, since their order of magnitude matches the upper bounds that we will obtain with two heuristics.

Lemma 5.2. Let $\epsilon \in (0, \frac{1}{5})$ and n large enough. Let $G = G(\mathcal{X}_n, r_n)$ be any ϵ-nice geometric graph with n nodes and let φ be any layout of G. Then, for any integer i such that $\alpha = i/n \in (2\epsilon, 1-2\epsilon)$, it is the case that

$$\theta(i,\varphi,G) \ge \tfrac{3(1-5\epsilon)}{8} \min\left\{\sqrt{\alpha-2\epsilon}, \sqrt{1-\alpha-2\epsilon}\right\} n^2 r_n^3.$$

Proof. We assign colors to points (nodes) and boxes: color the first i points in the ordering "red" and the other points "green"; color the boxes containing at most $\frac{1}{5}\epsilon a_n$ green points "red", the boxes containing at most $\frac{1}{5}\epsilon a_n$ red points "green", and the other boxes "yellow". Let Y_n be the number of yellow boxes.

Observe that $\theta(i,\varphi,G)$ is the total number of edges between opposite-color points. Let us refer to such edges as "within-box" if the points in question lie in the same box, or "between-box" if not. We consider two cases:

Case 1: $Y_n \ge 25\epsilon^{-2}n^{1/2}a_n^{-1/2}$. For each yellow box, the cost of within-box edges is at least $\epsilon^2 a_n^2/25$. Hence,

$$\theta(i,\varphi,G) \ge (\tfrac{1}{5}\epsilon a_n)^2 \cdot Y_n \ge n^{1/2}a_n^{3/2} = n^2 r_n^3.$$

Case 2: $Y_n < 25\epsilon^{-2}n^{1/2}a_n^{-1/2}$. In this case, we consider only between-box edges (between opposite colored points) which are between neighboring boxes (including diagonal neighbors). Consider a particular box containing a total of t points, r of them red. Suppose that the total number of red points in neighboring boxes is r' and the total number of green points in neighboring boxes is g'. Then, the total number of between-box edges of the type we are considering, involving points in that particular box, is $rg' + (t-r)r' = r(g'-r') + tr'$. Given t, r', and g', this is a linear function of r and so attains its minimum over the range $[0,t]$

either at $r = 0$ or at $r = t$ (or both). Hence, it is possible to change the points in that box to either all red or all green without increasing the total number of between-box edges of the type we are considering.

Let us modify the coloring of points by going through the yellow boxes in turn, successively changing the color of points in each box either to all red or to all green, whichever does not increase the total number of between-box edges of the type we are considering. When done, there will no longer be any yellow boxes! Let R_n be the number of red boxes and G_n the number of green boxes based on this modified coloring.

By niceness, the number of points whose color has been changed is at most

$$Y_n \cdot (1+\epsilon)\tfrac{1}{4}a_n \le 25\epsilon^{-2}n^{1/2}a_n^{1/2} = 25\epsilon^{-2}nr_n.$$

Thus, for n so large that $25\epsilon^{-2}r_n \le \epsilon$, the number of red points in the modified coloring is at least $(\alpha - \epsilon)n$.

By definition of "green box", the number of red points in green boxes is at most $\frac{1}{5}\epsilon a_n 4\lceil 1/r_n \rceil^2 \le \epsilon n$. Thus the total number of red points in red boxes (in the modified coloring) is at least $(\alpha - 2\epsilon)n$ for n large enough. By a similar argument, the number of green points in green boxes in the modified coloring is at least $(1-\alpha-2\epsilon)n$ for n large enough.

As, by ϵ-niceness, no box can contain more than $\frac{1}{4}(1+\epsilon)a_n$ points and there are at least $(\alpha - 2\epsilon)n$ red points in red boxes and $(1-\alpha-2\epsilon)n$ green points in green boxes, we have

$$R_n \ge \frac{(\alpha-2\epsilon)n}{\frac{1}{4}(1+\epsilon)a_n} = \frac{4(\alpha-2\epsilon)}{(1+\epsilon)r_n^2} \quad \text{and} \quad G_n \ge \frac{(1-\alpha-2\epsilon)n}{\frac{1}{4}(1+\epsilon)a_n} = \frac{4(1-\alpha-2\epsilon)}{(1+\epsilon)r_n^2}.$$

Let ∂G denote the number of pairs of neighbor boxes of opposite colors in G (with the modified coloring). By Proposition 4.1,

$$\partial G \ge 3\min\left\{\sqrt{R_n},\ \sqrt{G_n}\right\} \ge \frac{6}{r_n}\min\left\{\sqrt{\frac{\alpha-2\epsilon}{1+\epsilon}},\ \sqrt{\frac{1-\alpha-2\epsilon}{1+\epsilon}}\right\}.$$

By niceness and the definition of box coloring, each red box contains at least $(1-2\epsilon)\frac{1}{4}a_n$ red points, and each green box contains at least $(1-2\epsilon)\frac{1}{4}a_n$ green points. As a consequence,

$$\begin{aligned}\theta(i,\varphi,G) &\ge \partial G \tfrac{(1-2\epsilon)^2}{16}a_n^2 \ge \tfrac{3(1-2\epsilon)^2}{8\sqrt{1+\epsilon}}a_n^{3/2}n^{1/2}\min\left\{\sqrt{\alpha-2\epsilon},\ \sqrt{1-\alpha-2\epsilon}\right\} \\ &\ge \tfrac{3(1-5\epsilon)}{8}n^2r_n^3\min\left\{\sqrt{\alpha-2\epsilon},\ \sqrt{1-\alpha-2\epsilon}\right\}.\end{aligned}$$

This lower bound is smaller than the one for Case 1 and thus holds for both cases. □

The following result presents our lower bounds for the "edge problems" on nice graphs:

Theorem 5.1 (Lower bounds I). Let $\epsilon \in (0, \frac{1}{5})$ and n large enough. Let $G = G(\mathcal{X}_n, r_n)$ be an ϵ-nice geometric graph with n nodes. Then, the following lower bounds hold:

$$\text{MINEB}(G) \geq \tfrac{3(1-8\epsilon)}{8\sqrt{2}} \cdot n^2 r_n^3 \tag{lb1}$$

$$\text{MINCW}(G) \geq \tfrac{3(1-8\epsilon)}{8\sqrt{2}} \cdot n^2 r_n^3 \tag{lb2}$$

$$\text{MINLA}(G) \geq \tfrac{(1-42\sqrt{\epsilon})}{4\sqrt{2}} \cdot n^3 r_n^3 \tag{lb3}$$

Proof. The proofs of (lb1) and (lb2) are obtained from Lemma 5.2 by setting $i = \lfloor n/2 \rfloor$. To prove (lb3), take any layout φ of G. Then by Lemma 5.2

$$\text{LA}(\varphi, G) = \sum_{i=1}^{n} \theta(i, \varphi, G) \geq \sum_{2\epsilon n < i < (1-2\epsilon)n} \theta(i, \varphi, G) \geq \tfrac{3(1-5\epsilon)}{4} n^2 r_n^3 \sum_{2\epsilon n < i < n/2} \sqrt{i/n - 2\epsilon}.$$

Using the facts that $a > b$ implies $\sqrt{a-b} \geq \sqrt{a} - \sqrt{b}$ and that $\sum_{k=1}^{m} \sqrt{k} \geq \frac{2}{3} m^{3/2} + O\left(m^{1/2}\right)$ we obtain (lb3) by successive minorizations. □

5.2 Node problems

For "node problems", we shall obtain even tighter lower bounds. We take a finer dissection for these problems.

Definition 5.2. Again, we consider any set $\mathcal{X}_n$ of n points in $[0,1]^2$, which together with a radius r_n induce a random geometric graph $G = G(\mathcal{X}_n, r_n)$. Given a constant $\epsilon \in (0, 1)$, G is said to be ϵ-*good* if setting

$$\gamma = \left(\left\lfloor \frac{2}{\epsilon r_n} \right\rfloor r_n \right)^{-1}$$

and dissecting the unit square into $(\gamma r_n)^{-2}$ boxes, each of size $\gamma r_n \times \gamma r_n$, every box contains at most $p_+ = (1+\gamma)\gamma^2 a_n$ points and at least $p_- = (1-\gamma)\gamma^2 a_n$ points.

Observe that $1/(\gamma r_n)$ is an integer and that $\gamma \to \epsilon/2$ so that $\epsilon/4 \leq \gamma \leq \epsilon$ for n large enough.

Lemma 5.3. Let $\epsilon \in (0, 1)$. Then, with probability 1, for all large enough n, random geometric graphs $G(\mathcal{X}_n, r_n)$ are ϵ-good.

Proof. Choose a box in the dissection and let Y be the random variable counting the number of points of $\mathcal{X}_n$ in this box. As the points in $\mathcal{X}_n$ are i.u.d., we have $\mathbf{E}[Y] = \gamma^2 a_n$ where $\gamma = 1/\left(\left\lfloor \frac{2}{\epsilon r_n} \right\rfloor r_n \right)$. By Chernoff's bounds and Boole's inequality, used as in the proof of Lemma 5.1, the probability that some box has more than p_+ points or fewer than p_- points is bounded by

$$2 \left(\frac{1}{\gamma r_n} \right)^2 \exp\left(-\gamma^2(\gamma^2 a_n)/3\right) = \frac{2n}{\gamma^2 a_n} n^{-\gamma^4 b_n/3} \leq \frac{2}{\gamma^2} n^{1-\gamma^4 b_n/3},$$

which is summable in n because $b_n \to \infty$ as $n \to \infty$. The result follows by the Borel-Cantelli lemma. $\square$

The following lemma is the basis of our lower bounds for good graphs.

Lemma 5.4. Let $\epsilon \in (0, 1)$ and n large enough. Let $G = G(\mathcal{X}_n, r_n)$ be an ϵ-good geometric graph with n nodes. Let $i \in \{1, \ldots, n\}$ and consider any ordering φ on G. Then,

$$\delta(i, \varphi, G) \geq (1 - 3\epsilon) \left(h(i/n) - 2\sqrt{\epsilon} - 4r_n\right) \cdot nr_n \tag{1}$$

where for $x \in [0, 1]$ we set $h(x) = \min\left\{2\sqrt{x},\ 2\sqrt{1-x},\ 1\right\}$.

Proof. With $[0, 1)^2$ divided into boxes of side γr_n where $\gamma = 1/\left(\left\lfloor \frac{2}{\epsilon r_n} \right\rfloor r_n\right)$, we say that two boxes are *adjacent* if the l_∞ distance between their centers is at most $(1 - \gamma)r_n$. Note that under the l_∞ norm, any two points in adjacent boxes are at distance at most r_n from each other.

Given a layout φ, let the first i points be denoted "red" and the others "green". Then $\delta(i, \varphi, G)$ is the number of red points of $\mathcal{X}_n$ having one or more green points within a distance r_n. Let $\delta'(i, \varphi, G)$ be number of red points X such that there is at least one green point lying either in the box containing X or in a box adjacent to the box containing X. Then $\delta'(i, \varphi, G) \leq \delta(i, \varphi, G)$. We shall show that the right side of (1) is a lower bound for $\delta'(i, \varphi, G)$.

Given φ, let boxes containing only red points be denoted "red", let boxes containing only green points be denoted "green", and let the other boxes be denoted "yellow". Note that $\delta'(i, \varphi, G)$ is the number of red points X for which the box containing X is either itself yellow, or has some non-red box adjacent to it.

We claim that there is an ordering φ on $\mathcal{X}_n$ minimizing $\delta'(i, \cdot, G)$ such that φ induces at most one yellow box. Indeed, given an ordering φ inducing more than one yellow box, choose an ordering on yellow boxes. It is then possible to modify φ to an ordering φ' on points which respects the chosen ordering on yellow boxes of φ, and which satisfies $\delta'(i, \varphi', G) \leq \delta'(i, \varphi, G)$. This can be done by successively swapping red and green points, with each swap not increasing δ'.

Thus, without loss of generality, we can assume that φ induces at most one yellow box. Set $\alpha = i/n$ and let N_R be the number of red boxes. Then, by goodness and the fact that there are αn red points,

$$\frac{\alpha}{(1 + \gamma)(\gamma r_n)^2} - 1 \leq N_R \leq \frac{\alpha}{(1 - \gamma)(\gamma r_n)^2}. \tag{2}$$

Let A_R be union of the red boxes and let $A_G = [0, 1]^2 \setminus A_R$, the union of green and yellow boxes. Since each box has area $(\gamma r_n)^2$, by (2) we have

$$|A_R| \geq \frac{\alpha}{1 + \gamma} - (\gamma r_n)^2 \quad \text{and} \quad |A_G| \geq 1 - \frac{\alpha}{1 - \gamma}.$$

Let B be the union of red boxes that are adjacent to green or yellow boxes. Then $B = (A_G)_{(1-\gamma)r_n} \setminus A_G$, and therefore, by Proposition 4.2,

$$
\begin{aligned}
|B| &\geq \min\left\{2(1-\gamma)r_n|A_G|^{1/2},\ 2(1-\gamma)r_n|A_R|^{1/2} - 4((1-\gamma)r_n)^2,\ (1-\gamma)r_n\right\} \\
&\geq (1-\gamma)r_n \min\left\{2\left(1-\frac{\alpha}{1-\gamma}\right)^{1/2},\ 2\left(\frac{\alpha}{1+\gamma}-(\gamma r_n)^2\right)^{1/2} - 4r_n,\ 1\right\} \\
&\geq (1-\gamma)r_n \left(\min\left\{2\left(1-\frac{\alpha}{1-\gamma}\right)^{1/2},\ 2\left(\frac{\alpha}{1+\gamma}-(\gamma r_n)^2\right)^{1/2},\ 1\right\} - 4r_n\right).
\end{aligned}
$$

Using the fact that $a > b$ implies $\sqrt{a-b} \geq \sqrt{a} - \sqrt{b}$, we have

$$
\sqrt{1-\frac{\alpha}{1-\gamma}} \geq \frac{\sqrt{1-\alpha}-\sqrt{\gamma}}{\sqrt{1-\gamma}} \geq \sqrt{1-\alpha}-\sqrt{\gamma} \geq (1-\gamma)\left(\sqrt{1-\alpha}-\sqrt{\gamma}\right).
$$

As $r_n \to 0$, for n large enough, $r_n \leq \gamma^{-1/2}$. Since $(1+\gamma)^{-1/2} \geq 1-\gamma$, we have

$$
\begin{aligned}
\sqrt{\frac{\alpha}{1+\gamma}-(\gamma r_n)^2} &\geq \frac{\sqrt{\alpha-\gamma^2 r_n^2(1+\gamma)}}{\sqrt{1+\gamma}} \geq (1-\gamma)\sqrt{\alpha-\gamma^2 r_n^2(1+\gamma)} \\
&\geq (1-\gamma)\left(\sqrt{\alpha}-\gamma r_n\sqrt{1+\gamma}\right) \geq (1-\gamma)\left(\sqrt{\alpha}-\sqrt{\gamma}\right).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
|B| &\geq (1-\gamma)r_n\left(\min\left\{2(1-\gamma)\left(\sqrt{1-\alpha}-\sqrt{\gamma}\right),\ 2(1-\gamma)\left(\sqrt{\alpha}-\sqrt{\gamma}\right),\ 1\right\} - 4r_n\right) \\
&\geq (1-\gamma)^2 r_n\left(\min\left\{2\sqrt{1-\alpha}-2\sqrt{\gamma},\ 2\sqrt{\alpha}-2\sqrt{\gamma},\ 1-2\sqrt{\gamma}\right\} - 4r_n\right) \\
&\geq (1-\gamma)^2 r_n\left(\min\left\{2\sqrt{1-\alpha},\ 2\sqrt{\alpha},\ 1\right\} - 2\sqrt{\gamma} - 4r_n\right). \qquad (3)
\end{aligned}
$$

Since B is a union of disjoint boxes, each of area $\gamma^2 r_n^2$, the number of such boxes is their total area divided by $\gamma^2 r_n^2$. By goodness the number of points lying in the region B is bounded below by $(1-\gamma)n$ times its area. Hence by (3) we obtain

$$
\delta'(i,\varphi,G) \geq (1-3\gamma)nr_n(h(\alpha)-2\gamma^{1/2}-4r_n) \geq (1-3\epsilon)nr_n(h(\alpha)-2\epsilon^{1/2}-4r_n),
$$

which proves the lemma. □

The following result presents our lower bounds for the "node problems" on good graphs:

Theorem 5.2 (Lower bounds II). Let $\epsilon \in (0,1)$ and n large enough. Let $G = G(\mathcal{X}_n, r_n)$ be an ϵ-good geometric graph with n nodes. Then, the following lower bounds hold:

$$
\textsc{minvs}(G) \geq (1-6\sqrt{\epsilon}) \cdot nr_n \qquad \text{(lb4)}
$$

$$
\textsc{minbw}(G) \geq (1-6\sqrt{\epsilon}) \cdot nr_n \qquad \text{(lb5)}
$$

$$
\textsc{minsc}(G) \geq \left(\tfrac{5}{6}-6\sqrt{\epsilon}\right) \cdot n^2 r_n \qquad \text{(lb6)}
$$

Proof. The proof of (lb4) is obtained directly from Lemma 5.4 by taking i with $\frac{1}{4} \leq \frac{i}{n} \leq \frac{3}{4}$, so that $h(i/n) = 1$.

To prove (lb5), consider any layout φ and any $i \in [n]$ such that $\delta(i, \varphi, G) = \textsc{vs}(\varphi, G)$. Then, there must be $\textsc{vs}(\varphi, G)$ nodes before the node at position i, all of them connected with some other nodes located after position i. The first of these $\textsc{vs}(\varphi, G)$ nodes must have an edge that jumps at least $\textsc{vs}(\varphi, G)$ nodes. In other words, there is an edge $uv \in E(G)$ with $\lambda(uv, \varphi, G) \geq \textsc{vs}(\varphi, G)$ so that $\textsc{bw}(\varphi, G) \geq \textsc{vs}(\varphi, G)$ for any layout φ. Hence $\textsc{minbw}(G) \geq \textsc{minvs}(G)$ and (lb5) follows from (lb4).

To prove (lb6), consider any layout φ of G. Then, using again Lemma 5.4, we have that for large enough n,

$$\begin{aligned}
\textsc{sc}(\varphi, G) &= \sum_{i=1}^{n} \delta(i, \varphi, G) \\
&\geq (1 - 3\epsilon) n r_n \sum_{i=1}^{n} \left(h(i/n) - 2\sqrt{\epsilon} - 4r_n\right) \\
&\geq (1 - 3\epsilon) n r_n \left(\left(4 \sum_{1 \leq i \leq n/4} \sqrt{\tfrac{i}{n}} \right) + \left(\sum_{n/4 < i < 3n/4} 1 \right) - \left(\sum_{i=1}^{n} 2\sqrt{\epsilon} + 4r_n \right) \right) \\
&\geq \left(\tfrac{5}{6} - 6\sqrt{\epsilon}\right) \cdot n^2 r_n .
\end{aligned}$$

Here, we used again that $\sum_{k=1}^{m} \sqrt{k} \geq \frac{2}{3} m^{3/2} + O\left(m^{1/2}\right)$. □

6 Approximation algorithms

In this section we present two algorithms that compute feasible solutions that are within a constant of the previous lower bounds. As in the previous section, we work with (random) geometric graphs whose radius is of the form $r_n = \sqrt{a_n/n}$ where $r_n \to 0$ and $a_n/\log n = b_n \to \infty$. The coordinates of a point $u \in [0,1]^2$ are denoted $x(u)$ and $y(u)$. We use again the notion of ϵ-good graphs introduced in Definition 5.2.

The Projection algorithm. The Projection algorithm creates a layout by ordering the nodes according to their projection onto the x-axis. Another way to see this algorithm is to sweep a vertical line starting from $x = 0$ to $x = 1$, numbering vertices in the order the line touches them. Figure 6 illustrates this algorithm.

The expected running time of the Projection algorithm to compute the projected layout of a random geometric graph is linear, as it requires only the ranking of n numbers distributed uniformly.

Theorem 6.1 (Upper bounds). Let $\epsilon \in (0, 1)$ and n large enough. Let $G = G(\mathcal{X}_n, r_n)$ be an ϵ-good geometric graph with n nodes. Then, the following upper bounds on the cost of

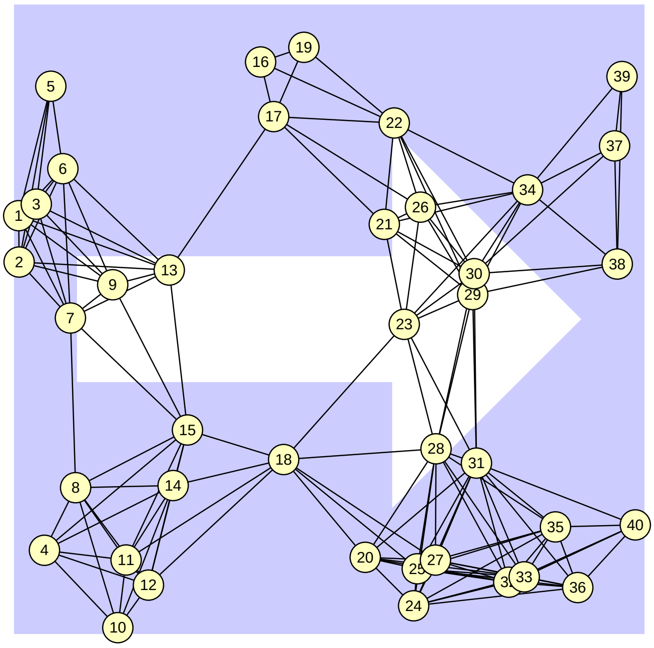

Figure 6: Illustration of the projection algorithm.

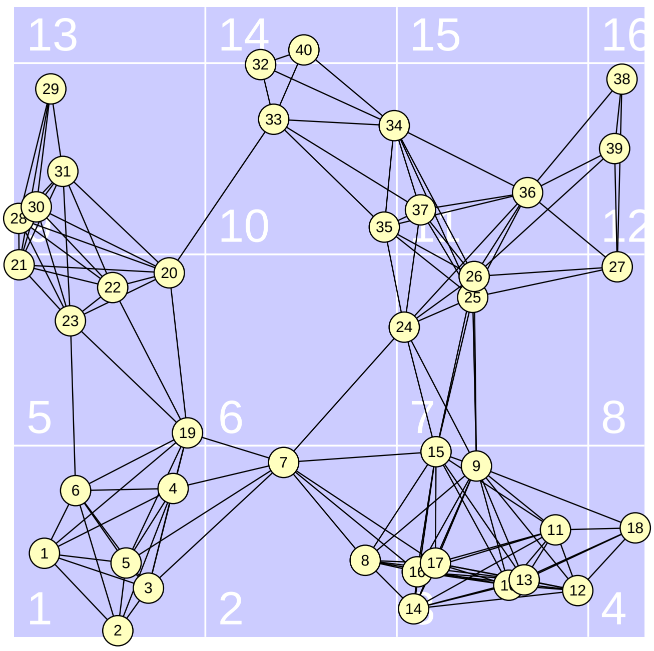

Figure 7: Illustration of the dissection algorithm.

the projected layout π of G hold:

$$\text{CW}(\pi, G) \leq (1+3\epsilon)^5 \cdot n^2 r_n^3 \tag{ub1}$$

$$\text{EB}(\pi, G) \leq (1+3\epsilon)^5 \cdot n^2 r_n^3 \tag{ub2}$$

$$\text{BW}(\pi, G) \leq (1+3\epsilon)^2 \cdot n r_n \tag{ub3}$$

$$\text{LA}(\pi, G) \leq (1+3\epsilon)^5 \cdot n^3 r_n^3 \tag{ub4}$$

$$\text{VS}(\pi, G) \leq (1+3\epsilon)^2 \cdot n r_n \tag{ub5}$$

$$\text{SC}(\pi, G) \leq (1+3\epsilon)^2 \cdot n^2 r_n \tag{ub6}$$

Proof. Given a node u of G, let $\theta(u)$ denote the cut induced by the projected layout π on u, that is, the number of edges vw such that $x(v) \leq x(u)$ and $x(u) < x(w)$. Given an edge uv from G, let $\lambda(uv)$ denote the length induced by π on uv, that is, the number of nodes w such that $x(u) < x(w)$ and $x(w) < x(v)$. Set $\gamma = 1/\left(\left\lfloor \frac{2}{\epsilon r_n} \right\rfloor r_n\right)$ and $k = \lceil 1/\gamma \rceil \leq 1/\gamma + 1$. Recall that $\frac{1}{4}\epsilon \leq \gamma \leq \epsilon$. Every possible edge is between boxes (of side γr_n) with centers at distance at most r_n. Thus,

$$\theta(u) \leq \sum_{0 \leq i \leq k} (k+1-i)\,(2k+1)\,\frac{1}{\gamma r_n} p_+^2 \leq (1+3\epsilon)^5 n^2 r_n^3. \tag{4}$$

On the other hand, observe that $\lambda(uv)$ is bounded above by the number of possible nodes in the columns of boxes between the column of u and the column of v. Thus,

$$\lambda(uv) \leq p_+\,(k+2)\,\frac{1}{\gamma r_n} \leq (1+3\epsilon)^2 n r_n. \tag{5}$$

Bounds (ub1), (ub2) and (ub3) follow directly from Equations (4) and (5). Bounds (ub4), (ub5) and (ub6) hold because for any layout φ, we have $\text{LA}(\varphi, G) \leq n\text{CW}(\varphi, G)$, $\text{VS}(\varphi, G) \leq \text{BW}(\varphi, G)$ and $\text{SC}(\varphi, G) \leq n\text{VS}(\varphi, G)$. ☐

Next we give lower bounds on the costs of the layouts delivered by the Projection algorithm on good geometric graphs:

Theorem 6.2 (Lower bounds of Projection). Let $\epsilon \in (0,1)$ and n large enough. Let $G = G(\mathcal{X}_n, r_n)$ be an ϵ-good geometric graph with n nodes. Then, the following lower bounds on the cost of the projected layout π of G hold:

$$\text{VS}(\pi, G) \geq (1-3\epsilon) \cdot n r_n \tag{lb1'}$$

$$\text{BW}(\pi, G) \geq (1-3\epsilon) \cdot n r_n \tag{lb2'}$$

$$\text{SC}(\pi, G) \geq (1-5\epsilon) \cdot n^2 r_n \tag{lb3'}$$

$$\text{CW}(\pi, G) \geq (1-8\epsilon) \cdot n^2 r_n^3 \tag{lb4'}$$

$$\text{EB}(\pi, G) \geq (1-8\epsilon) \cdot n^2 r_n^3 \tag{lb5'}$$

$$\text{LA}(\pi, G) \geq (1-10\epsilon) \cdot n^3 r_n^3 \tag{lb6'}$$

Proof. Set $\gamma = 1/\left(\left\lfloor \frac{2}{\epsilon r_n} \right\rfloor r_n\right)$ and $k = \lceil 1/\gamma \rceil$. Let us prove (lb1'). Consider any node u far enough from the square boundaries. All the nodes in the $k-2$ columns preceding the column of u must be connected to some node in the next column after the column of u. Therefore,

$$\textsc{vs}(\pi, G) \geq p_-(k-2)\frac{1}{\gamma r_n} \geq (1-3\epsilon) n r_n.$$

Bound (lb2') follows because $\textsc{bw}(\varphi, G) \geq \textsc{vs}(\varphi, G)$ for any layout φ (see the proof of Theorem 5.2).

Next, let us prove (lb3'). We can extend the previous proof to all the points which are away from the left and the right borders of the unit square:

$$\begin{aligned}\textsc{sc}(\pi, G) &\geq p_-\frac{1}{\gamma r_n}\left(\frac{1}{\gamma r_n} - 2k\right)\left(p_-(k-2)\frac{1}{\gamma r_n}\right) \\ &\geq (a_n^2/r_n^3)(1-\gamma)^3(1-2\gamma) \geq n^2 r_n (1-5\epsilon).\end{aligned}$$

We prove now (lb4') and (lb5'). Take any node u away from the left and right borders of $[0,1]^2$. We have

$$\begin{aligned}\textsc{cw}(\pi, G) &\geq \sum_{i=1}^{k-2} p_-^2 (k-i-1)\left(\frac{1}{\gamma r_n} - 2k\right)(2k-3) \\ &\geq (1-2\gamma)^2(1-\gamma)^4 a_n^2/r_n \geq (1-8\epsilon) n^2 r_n^3.\end{aligned}$$

As the $\lfloor n/2 \rfloor$-th node of the projected layout must be away from the left and right borders of $[0,1]^2$, we have the same result for $\textsc{eb}(\pi, G)$.

Finally, let us prove (lb6'). By the argument for the cut width,

$$\begin{aligned}\textsc{la}(\pi, G) &\geq p_-\left(\frac{1}{\gamma r_n}\right)\left(\frac{1}{\gamma r_n} - 2k\right)\sum_{i=1}^{k-2} p_-^2(k-i-1)\left(\frac{1}{\gamma r_n} - 2k\right)(2k-3) \\ &\geq (1/r_n^2)(1-2\gamma)^2(1-6\epsilon)a_n^2/r_n \geq (1-10\epsilon) n^3 r_n^3.\end{aligned}$$

□

The behavior of the Projection algorithm on good graphs is therefore characterized by the following consequence of Theorems 6.1 and 6.2:

Corollary 6.1. Let $\epsilon \in (0,1)$ and n large enough. Let $G = G(\mathcal{X}_n, r_n)$ be an ϵ-good geometric graph with n nodes. Then, for any measure $f \in \{\textsc{bw}, \textsc{vs}, \textsc{sc}, \textsc{cw}, \textsc{eb}, \textsc{la}\}$, we have

$$(1-10\epsilon) \leq \frac{f(\pi, G)}{A_f} \leq (1+3\epsilon)^5$$

where

$$\begin{aligned}
A_{\textsc{bw}} &= n r_n, & A_{\textsc{vs}} &= n r_n, & A_{\textsc{sc}} &= n^2 r_n, \\
A_{\textsc{cw}} &= n^2 r_n^3, & A_{\textsc{eb}} &= n^2 r_n^3, & A_{\textsc{la}} &= n^3 r_n^3.
\end{aligned}$$

As a consequence of Lemmata 5.1 and 5.3, Theorems 5.1 and 5.2, and Corollary 6.1, we obtain the main result of our paper:

Theorem 6.3. Let $(r_i)_{i\geq 1}$ be a sequence of positive numbers with $r_n \to 0$ and $nr_n^2/\log n \to \infty$; let $(X_i)_{i\geq 1}$ be a sequence of i.u.d. random points in $[0,1]^2$. Let $G_n = G(\mathcal{X}_n, r_n)$; then, with probability 1,

$$\begin{array}{ccccccccc}
 & & \lim_{n\to\infty} \frac{\mathrm{MINBW}(G_n)}{nr_n} & = 1 & & & & & \\
 & & \lim_{n\to\infty} \frac{\mathrm{MINVS}(G_n)}{nr_n} & = 1 & & & & & \\
1 & \geq & \limsup_{n\to\infty} \frac{\mathrm{MINSC}(G_n)}{n^2 r_n} & \geq & \liminf_{n\to\infty} \frac{\mathrm{MINSC}(G_n)}{n^2 r_n} & \geq & \frac{5}{6} & \approx & 0.833 \\
1 & \geq & \limsup_{n\to\infty} \frac{\mathrm{MINCW}(G_n)}{n^2 r_n^3} & \geq & \liminf_{n\to\infty} \frac{\mathrm{MINCW}(G_n)}{n^2 r_n^3} & \geq & \frac{3}{8\sqrt{2}} & \approx & 0.265 \\
1 & \geq & \limsup_{n\to\infty} \frac{\mathrm{MINEB}(G_n)}{n^2 r_n^3} & \geq & \liminf_{n\to\infty} \frac{\mathrm{MINEB}(G_n)}{n^2 r_n^3} & \geq & \frac{3}{8\sqrt{2}} & \approx & 0.265 \\
1 & \geq & \limsup_{n\to\infty} \frac{\mathrm{MINLA}(G_n)}{n^3 r_n^3} & \geq & \liminf_{n\to\infty} \frac{\mathrm{MINLA}(G_n)}{n^3 r_n^3} & \geq & \frac{1}{4\sqrt{2}} & \approx & 0.176
\end{array}$$

The consequence is immediate:

Theorem 6.4. Let $(r_i)_{i\geq 1}$ be a sequence of positive numbers with $r_n \to 0$ and $nr_n^2/\log n \to \infty$; let $(X_i)_{i\geq 1}$ be a sequence of i.u.d. random points in $[0,1]^2$. Then, with probability 1, and for all large enough n, the Projection algorithm is a constant approximation algorithm for the Bandwidth, Minimum Linear Arrangement, Minimum Cut, Minimum Sum Cut, Vertex Separation and Bisection problems on random geometric graphs $G(\mathcal{X}_n, r_n)$. Moreover, for the Bandwidth and Vertex Separation problems, the Projection algorithm is asymptotically optimal.

The Dissection algorithm. An important idea in combinatorial optimization on the plane was Karp's analysis on the dissection algorithm for the TSP problem [43]. Let us adapt this algorithm to layout problems. For a geometric graph with n nodes and radius r_n, our Dissection algorithm, parameterized by a constant κ, is defined as follows:

1. Dissect $[0,1]^2$ into boxes of size $r_n/\kappa \times r_n/\kappa$.

2. Enumerate the points, following the order of the boxes in lexicographic order (from bottom to top and from left to right). For points in the same box, enumerate them arbitrarily. Return the layout computed in this way.

Figure 7 shows an example of the application of this algorithm.

Using the same kind of arguments used for the Projection algorithm, we can prove the following result:

Theorem 6.5. Let $(r_i)_{i\geq 1}$ be a sequence of positive numbers with $r_n \to 0$ and $nr_n^2/\log n \to \infty$; let $(X_i)_{i\geq 1}$ be a sequence of i.u.d. random points in $[0,1]^2$. Then, with probability 1, for all large enough n, the Dissection algorithm is a constant approximation algorithm for the Bandwidth, Minimum Linear Arrangement, Minimum Cut, Minimum Sum Cut, Vertex Separation and Bisection problems random geometric graphs $G(\mathcal{X}_n, r_n)$. Moreover, for the Bandwidth and Vertex Separation problems, the solutions computed by the Dissection algorithm are asymptotic to the optimal ones.

7 Experimental considerations

This section describes several computational experiments we have performed in order to obtain a better knowledge of layout problems on random geometric graphs. The goal is to complement and expand the previous theoretical results. In order to help to reproduce and verify the measurements and the code mentioned in this research, its code and raw data are available at http://www.lsi.upc.es/∼jpetit/RandomGeometricGraphs on the World Wide Web. The pictures, in color, are also available at this address.

Behavior of the Projection and Dissection algorithms. According to Corollary 6.1, the Projection and Dissection algorithms exhibit a behavior such that for any measure $f \in \{\text{BW}, \text{VS}, \text{SC}, \text{CW}, \text{EB}, \text{LA}\}$, we have that $f(\pi, G)/A_f$ and $f(\psi, G)/A_f$ are as close to 1 as we want. However, this result is asymptotic in nature. The following computational experiment was designed in order to compare the behavior of these two algorithms for the different measures. The aim was to compare between the algorithms, and to compare the experimental results with the predicted ones.

Random geometric graphs with $r_n = \sqrt{(\log n)(\log\log n)/n}$ with up to 200,000 nodes were generated. For each each value of n, we computed the average of all the considered measures f using the Projection algorithm and the Dissection algorithm (with $\kappa \in \{1, 2, 4, 8, 16\}$) and normalized by their respective orders of magnitude A_f (as defined in Corollary 6.1). Figures 8 and 9 summarize the results. Note that the plots have neither the same scaling nor the same origin. The standard deviation was very low in all cases (refer to the raw files).

With regard to the competitive analysis, we observe for all the measures except EB that the Projection algorithm is the one that obtains better approximations in the considered range of values of n. We also observe that the approximations of the Dissection algorithm improve as κ increases. In fact, Dissection with $\kappa = 16$, performs only slightly worse than Projection. For EB both algorithms seem to perform similarly. However, we have observed that the execution time increases when κ increases. On the other hand, with regard to the predicted asymptotic behavior, we can observe that the only measure which achieves it for n <200,000 is EB. For the rest of measures, LA seems to tend to 1 quickly, SC and VS remain still far from 1 (about 5%), and BW and CW remain quite far from 1 (10% and 15% respectively).

The conclusions induced from that experiment are twofold: On one hand, the Projection algorithm behaves better than the Disection algorithm, both in quality and time. On the other hand, the predicted convergence to 1 is far from being observed for the considered number of nodes for the SC, CW, VS and BW measures. This may be due to the fact that $b_n = \log \log n$ goes to infinity very slowly, but experiments with denser graph exhibit the same phenomenon, albeit reduced.

Bisection of random geometric graphs. The problem of partitioning a graph into a number of pieces is a fundamental task in computer science. Many heuristics have been proposed for these problems and many libraries implement them. On the other hand, the bisection of geometric random graphs has already been considered [39, 7, 49]. In the following, we analyze the behavior of the Projection algorithm and of several heuristics included in the Chaco and Party libraries [36, 68] for the EdgeBis problem on random geometric graphs with $r = \sqrt{\log^2 n/n}$. These libraries offer global and local heuristics which can be combined.

We have compared the following global heuristics. pro: the Projection algorithm we have presented in Section 5; mul: the Multilevel method of Hendrickson and Leland, which uses edge contraction schemes; ine: the Inertial method requires geometric information of the vertices in order to assimilate mass values to vertices; gai: the Gain method is based on a greedy strategy; far: the Farhat method is based on another greedy strategy; spl: an Spectral bisection algorithm using Lanczos eigen solver, spm: a combination of Multilevel and Spectral methods. The solutions obtained with these methods have been refined with local search based on the Kernighan–Lin heuristic (kl) or the Helpful-Set heuristic (hs). More details on all these heuristics are available in [36, 68] .

The measurement results are summarized in Figure 10 for graphs with up to 210,000 nodes and $r_n = \sqrt{(\log n)(\log \log n)/n}$. These were produced by generating 25 different random geometric graphs for each data point and taking the average of their bisection normalized by $n^{1/2}((\log n)(\log \log n))^{3/2}$. Again, the standard deviation was very low. Table 2 show the run time measurements needed by the different combinations of local and global heuristics to bisect a random geometric graph with $100,000$ nodes on a DEC Alpha Server 8400 machine.

From these experimental results and the knowledge we have on the projection algorithm for EdgeBis on random geometric graphs, we can classify the global heuristics into two groups according to whether they offer better approximations than the projection algorithm on the considered range of nodes. The global heuristics that perform worse than Projection are spl,

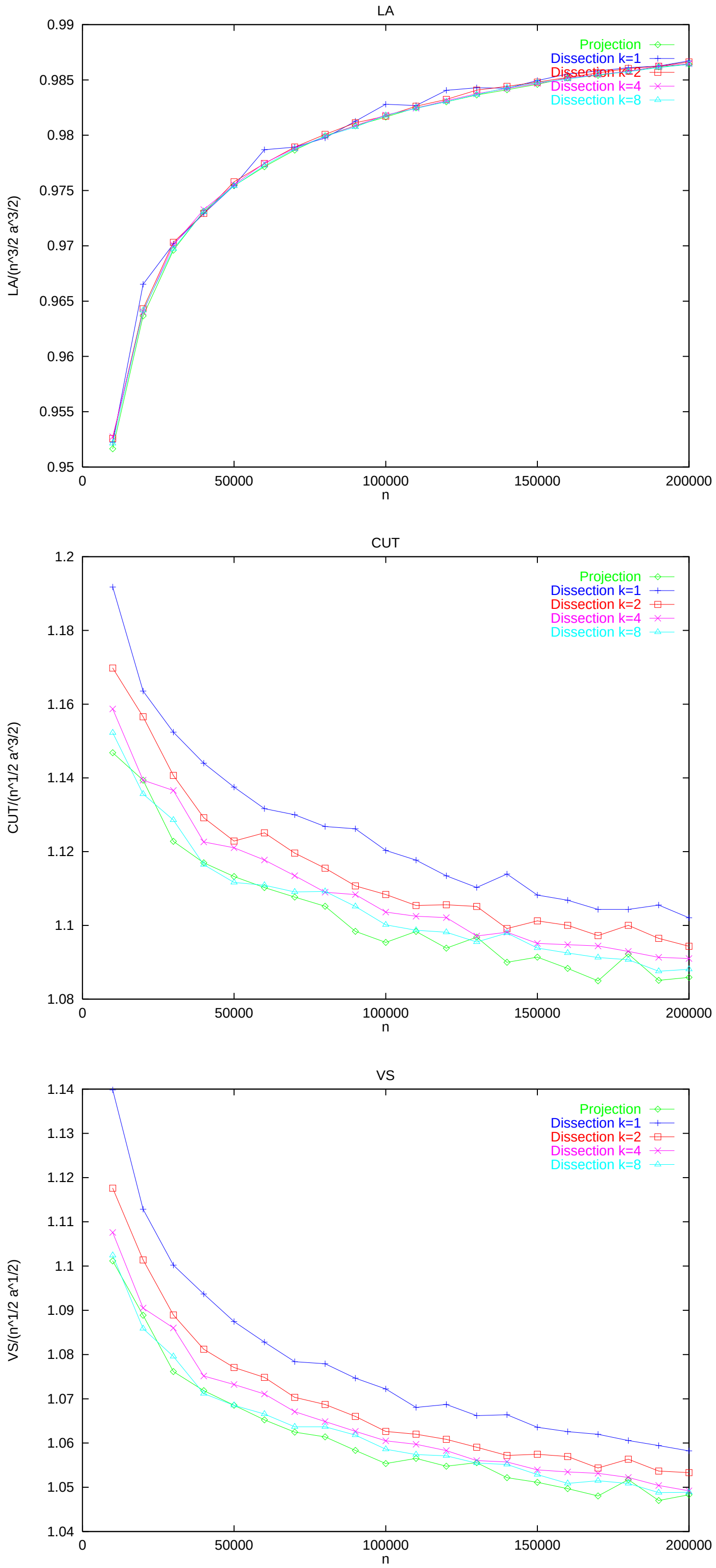

Figure 8: Normalized cost of each problem for the Projection and Dissection algorithms (Part 1).

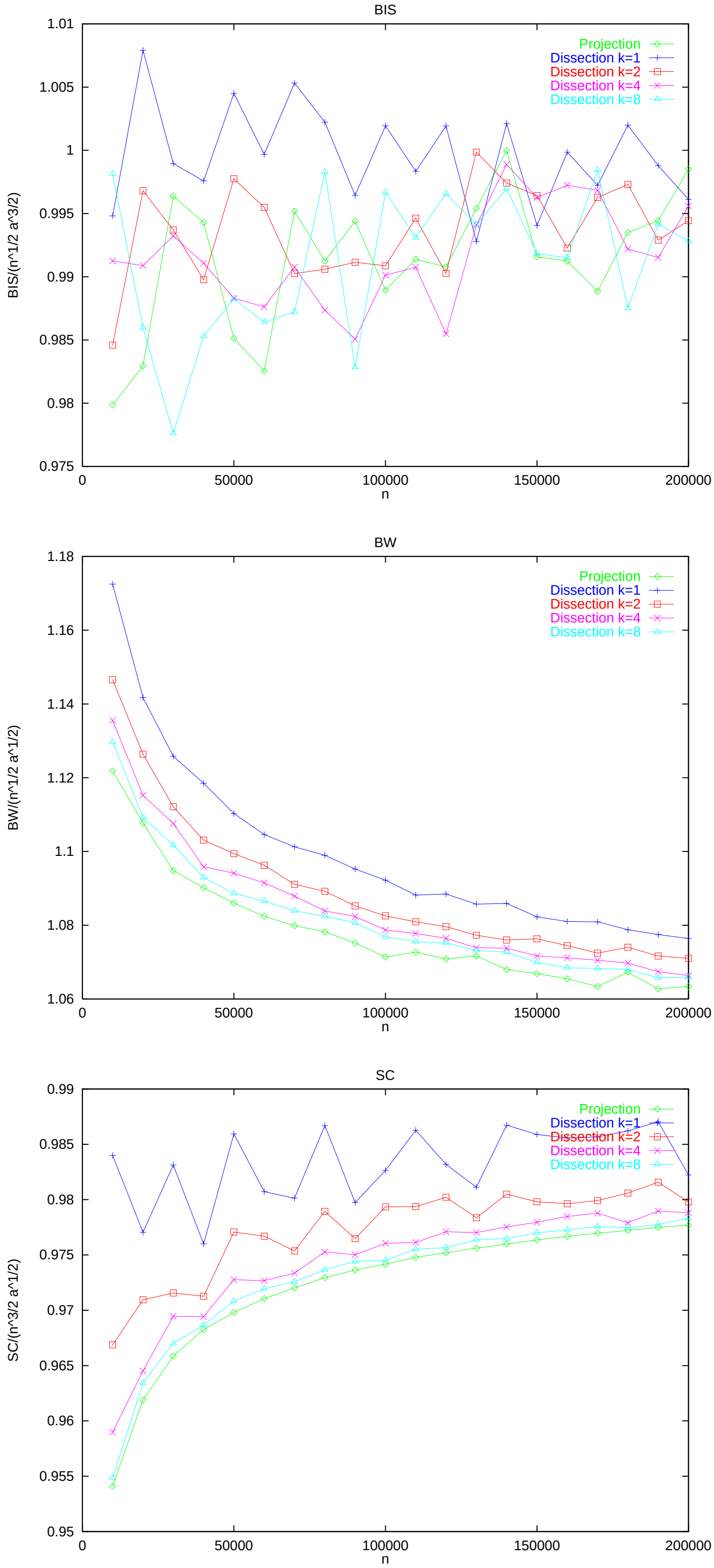

Figure 9: Normalized cost of each problem for the Projection and Dissection algorithms (Part 2).

far, gai and spm. The ones that perform better than pro are mul and spm. It is remarkable that spl performs better than pro for $n \leq 100,000$ but abruptly gets worse afterwards.

If we consider the application of local search after the global heuristics, we can remark that all the methods except mul can be substantially improved (the improvement factor is around 20%). However, the quality of the improved solution depends directly on the quality of the global solution. Moreover, we can observe that kl always returns slightly better solutions than hs when fixing the global heuristic. It is worth to remark that global solutions obtained by mul cannot be improved more than 0.25%.

Since by Theorem 6.4 we know that the Projection algorithm is a constant approximation algorithm, it seems reasonable to conjecture that all the considered global heuristics except spl are also constant approximation algorithms for the kind of graphs we are dealing with. Moreover, the local search heuristic of choice to improve them is Kernighan–Lin. However, the better solutions are always delivered by the Multilevel method, and these are difficult to improve even when using Kernighan–Lin. Referring to our computational experiment, we infer that the Multilevel method is the best method to bisect random geometric graphs (moreover, this method has the advantage of not using geometric information).

Minimum Linear Arrangement. In [41], Juvan and Mohar propose a technique to find approximative solutions to the MinLA problem based on eigenvectors associated to the second smallest Laplace eigenvalue of the input graph. We call their heuristic Spectral Sequencing (ss). In [66], Petit presented and analyzed sequential and parallel versions of an heuristic to approximate the MinLA problem on sparse graphs embedded in some geometry. The sequential heuristic (ss+sa) consists in obtaining a first global solution using Spectral Sequencing and improving it locally through Simulated Annealing starting at a low temperature using an special neighborhood. Petit presented also in [67] a lower bounding technique for the MinLA based only on the degree of the nodes (lb-degree).

Figure 11 compares the Projection algorithm (pro) with the ss and ss+sa heuristics for the MinLA problem. The lower bound computed by lb-degree is also shown. These results were again obtained by taking the average of 100 geometric random graphs for each n and normalizing by $n^{3/2}a_n^{3/2}$. We had to content ourselves with graphs up to 20,000 nodes because simulated annealing consumes a lot of time.

The obtained results show that ss+sa performs better than ss, which performs better than pro. We can also observe that the $0.176n^3r_n^3$ lower bound presented in Theorem 5.1 is better than the lb-degree lower bounding technique.

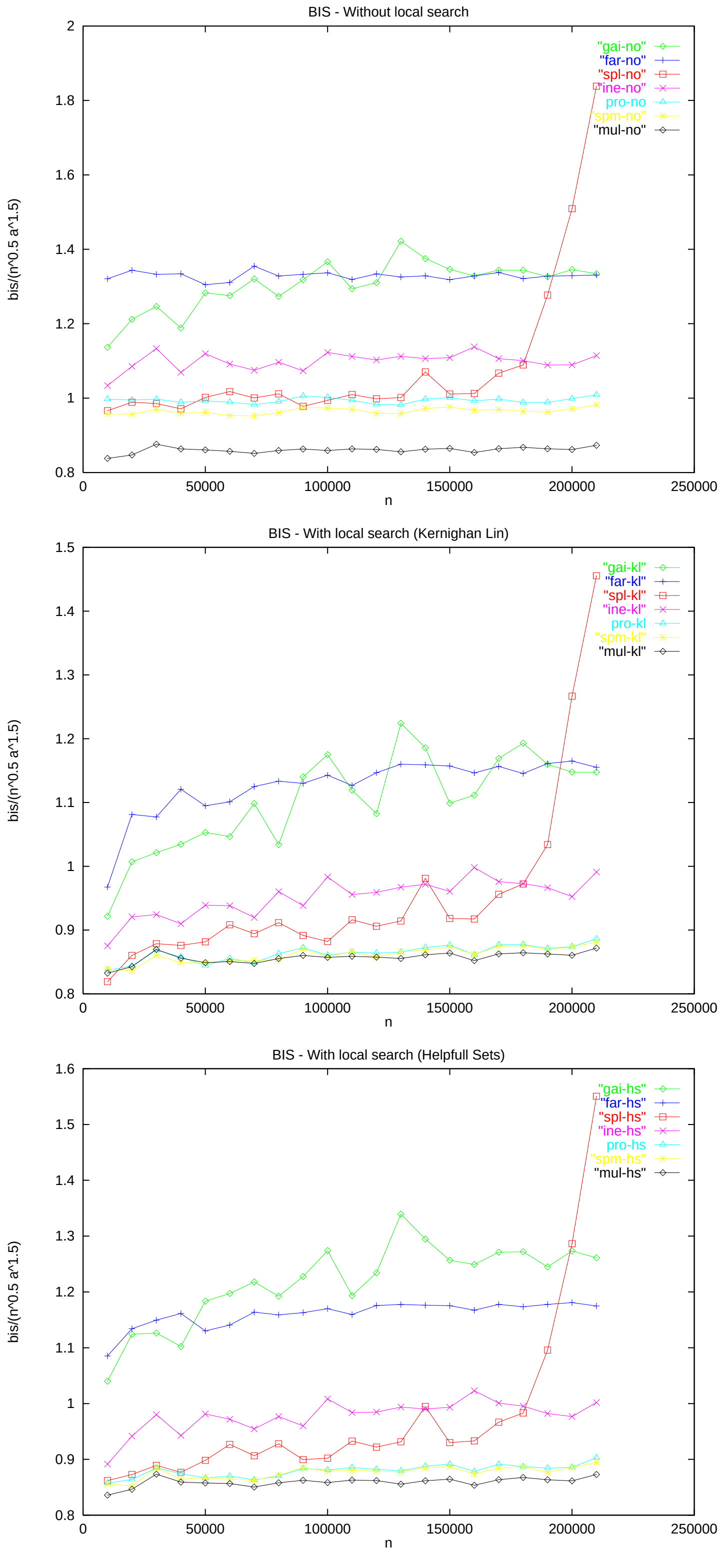

Figure 10: Bisection of random geometric graphs with different combinations of global and local heuristics (averages over 25 runs).

	gai	far	spl	ine	spm	mul	pro
no	2.36	3.51	164.51	0.48	23.01	18.44	0.44
kl	11.58	10.28	170.71	8.26	29.12	21.97	6.49
hs	6.36	6.78	167.47	3.42	26.18	21.73	3.60

Table 2: Running times (in seconds) for bisecting a random geometric graph with 10000 nodes and 5534888 edges.

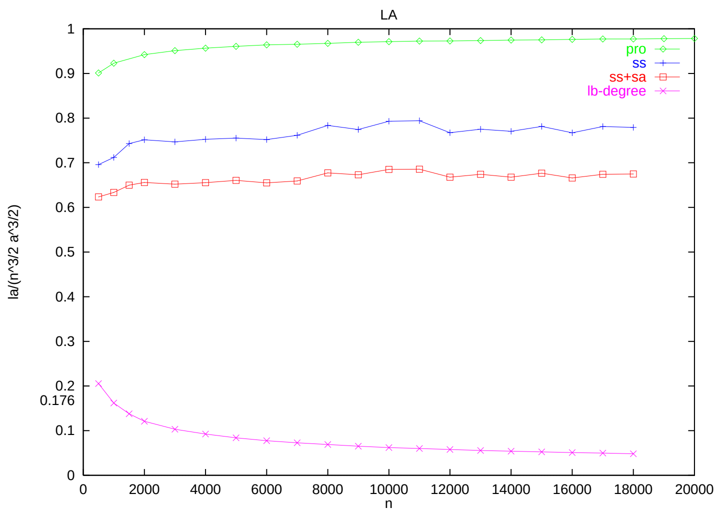

Figure 11: Comparison of upper and lower bounding techniques for the MinLA problem.

8 Conclusion

The behavior of heuristics is very difficult to characterize analytically. A way to increase the knowledge on these methods is to study them on a restricted setting. In this paper we have analyzed the approximation properties of several heuristics for layout problems on a particular class of random geometric graphs. Layout problems have importance in very different fields, and random geometric graphs may be a relevant abstraction to model instances that occur in practice.

First of all we have shown that for some of the problems, restricting their inputs to geometric instances does not lower their complexity. Our results reaffirm the close relation between disk graphs and planar graphs as it was already noticed by Clark, Colbourn and Johnson [18]. In regard to Theorem 3.2, we remark that Papadimitriou and Sideri conjecture that bisection of planar graphs is **NP**-complete in [64]. We conjecture that SumCut and MinLA have the same behavior, but in this case, the link with planar graphs with restricted degree is not clear.

Afterwards, we have provided lower bounds on the cost of all our considered problems on random geometric graphs whose radius ensures connectivity while preserving sparsity. Much effort has been invested in order to obtain tight lower bounds. Owing to the probabilistic nature of random geometric graphs, these lower bounds hold, almost surely, for large enough graphs. Two simple and fast heuristics have been proposed to approximate layout problems. We have proved that these heuristics, while naive, are in fact constant approximation algorithms. In fact, for the Bandwidth and the Vertex Separation problems, the algorithms are asymptotically optimal. It must be remarked, that our algorithms use graphical information (the node coordinates) in order to build a layout. This simplification is an accepted common practice in the literature [39, 49], motivated by the fact that recognizing geometric graphs is **NP**-hard [14] (refer to [37] for a discussion on that topic). In any case, this is not a nuisance for our results: On one hand, the characterization of the algorithms in Corollary 6.1 is so tight, that we already know which solution they will deliver with high probability. On the other hand, now that we have analytical information on these algorithms, we do not view them as a practical way to find approximate solutions, but rather as a valuable tool with which to compare other heuristics.

Indeed, while several heuristic methods have been acclaimed as an effective way to find good approximate solutions, the assessment of their quality can only be shown empirically. We claim that our new analytical results on random geometric graphs offer more insight and give more informative ways to analyze and compare heuristics in an experimental setting. For instance, many heuristics to approximate the Bisection problem are local search methods that accept an initial partitioning of the input graph and then attempt to improve it (i.e.

the Kernighan–Lin heuristic [48], Simulated Annealing [39], Path Optimization [7] or Helpful Sets [27]). In [39] and [49] it is shown that initial partitionings created by the *Line* heuristic dramatically improve the performance of the Kernighan–Lin and Simulated Annealing heuristics on geometric random graphs. The Line heuristic uses geometric information to split the vertex set of a geometric graph into two equal sized halves with a line of randomly chosen slope. Thanks to our results on the Projection algorithm (which is close to the Line heuristic), we can now affirm that all these heuristics which perform so well in practice are, in fact, constant approximation algorithms when applied to geometric random graphs (with high probability).

Our experimental results show there still remains a gap between theory and practice. The problem is that, for several problems, the asymptotic behavior of Corollary 6.1 is far from being achieved even with huge graphs with 200,000 vertices. On the other hand, our experimental comparison of several well-known heuristics for the Bisection problem with our constant approximation algorithms gives an informative benchmarking of all these techniques.

In the proofs of our results, we have concentrated only on the l_∞ norm. However, all the bounds also hold, with different values of constants for the remaining of the norms. We have considered only the two-dimensional geometric graphs as most real instances belong to that case, but we expect that similar results will also hold on d-dimensional spaces. Our results might be generalizable in several ways. For instance, it would be interesting to understand how the optimal costs of our problems behave for different radii; see [20, 65]. Bounds on the cost of layout problems for lattice graphs and random lattice graphs can be found in [21]. Finally, we leave as an open problem to determine whether or not the Projection and Dissection algorithms are, in some probabilistic and asymptotic sense, optimal for layout problems other than Bandwidth and Vertex Separation.

References

[1] D. Adolphson. Single machine job sequencing with precedence constraints. *SIAM Journal on Computing*, 6:40–54, 1977.

[2] D. Adolphson and T. C. Hu. Optimal linear ordering. *SIAM Journal on Applied Mathematics*, 25(3):403–423, November 1973.

[3] N. Alon, J. H. Spencer, and P. Erdős. *The probabilistic method*. Wiley-Interscience, New York, 1992.

[4] M. J. Appel and R. P. Russo. The connectivity of a graph on uniform points in $[0,1]^d$. Technical Report #275, Department of Statistics and Actuarial Science, University of Iowa, 1996.

[5] S. Arora, A. Frieze, and H. Kaplan. A new rounding procedure for the assignment problem with applications to dense graphs arrangements. In *37th IEEE Symposium on Foundations of Computer Science*, 1996.

[6] J. Beardwood, J. Halton, and J. M. Hammersley. The shortest path through many points. *Proceedings of the Cambridge Philos. Society*, 55:299–327, 1959.

[7] J. W. Berry and M. K. Goldberg. Path optimization for graph partitioning problems. *Discrete Applied Mathematics. Combinatorial Algorithms, Optimization and Computer Science*, 90(1-3):27–50, 1999.

[8] S. N. Bhatt and F. T. Leighton. A framework for solving VLSI graph layout problems. *Journal of Computer and System Sciences*, 28:300–343, 1984.

[9] G. Blache, M. Karpinski, and J. Wirtgen. On approximation intractability of the bandwidth problem. Technical Report TR98-014, Electronic Colloquium on Computational Complexity, 1998.

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

[11] R. Boppana. Eigenvalues and graph bisection: An average case analysis. In *Proc. on Foundations of Computer Science*, pages 280–285, 1987.

[12] R. A. Botafogo. Cluster analysis for hypertext systems. In *16th Annual ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 116–125, 1993.

[13] H. Breu. *Algorithmic aspects of constrained unit disk graphs*. PhD thesis, University of British Columbia, 1996.

[14] H. Breu and D. G. Kirkpatrick. Unit disk graph recognition is NP-hard. *Computational Geometry. Theory and Applications*, 9(1-2):3–24, 1998.

[15] T. N. Bui and A. Peck. Partitioning planar graphs. *SIAM Journal on Computing*, 21(2):203–215, 1992.

[16] P. Chin, J. Chvatalova, A. Dewdney, and N. Gibbs. The bandwidth problem for graphs and matrices: A survey. *Journal of Graph Theory*, 6:223–254, 1982.

[17] K. L. Chung. *A Course in Probability Theory.* Academic Press, New York, 1974.

[18] B. N. Clark, C. J. Colbourn, and D. S. Johnson. Unit Disk Graphs. *Discrete Mathematics*, 86:165–177, 1990.

[19] J. Díaz, A. Gibbons, G. Pantziou, M. Serna, P. Spirakis, and J. Torán. Parallel algorithms for the minimum cut and the minimum length tree layout problems. *Theoretical Computer Science*, (181):267–288, 1997.

[20] J. Díaz, M. D. Penrose, J. Petit, and M. Serna. Convergence theorems for some layout measures on random lattice and random geometric graphs. *Combinatorics, Probability and Computing.* To appear.

[21] J. Díaz, M. D. Penrose, J. Petit, and M Serna. Layout problems on lattice graphs. In T. Asano, H. Imai, D.T. Lee, S. Nakano, and T. Tokuyama, editors, *Computing and Combinatorics — 5th Annual International Conference, COCOON'99, Tokyo, Japan, July 26-28, 1999*, number 1627. Springer, Lecture Notes in Computer Science, July 1999.

[22] J. Díaz, J. Petit, and M. Serna. Random geometric problems on $[0,1]^2$. In M. Luby, J. Rolim, and M. Serna, editors, *Randomization and Approximation Techniques in Computer Science*, number 1518 in Lecture Notes in Computer Science. Springer, Lecture Notes in Computer Science, October 1998.

[23] J. Díaz, A. M. Gibbons, M. S. Paterson, and J. Torán. The minsumcut problem. In F. Dehen, R. J. Sack, and N. Santoro, editors, *Algorithms and Datastructure*, volume 519 of *Lecture Notes in Computer Science*, pages 65–79. Springer-Verlag, 1991.

[24] J. Díaz, J. Petit, M. Serna, and L. Trevisan. Approximating layout problems on random sparse graphs. *Discrete Mathematics.* To appear.

[25] J. Díaz, M. Serna, P. Spirakis, and J. Torán. *Paradigms for fast parallel approximability.* Cambridge University Press, Cambridge, 1997.

[26] J. Díaz, M. Serna, and J. Torán. Parallel approximation schemes for problems on planar graphs. *Acta Informatica*, 33:387–408, 1996.

[27] R. Diekmann, R. Lüling, B. Monien, and C. Spräner. Combining Helpful Sets and Parallel Simulated Annealing for the Graph-Partitioning Problem. *Parallel Algorithms and Applications*, 8:61–84, 1996.

[28] J. Ellis, I. H. Sudborough, and J. Turner. The vertex separation and search number of a graph. *Information and Computation*, (113):50–79, 1979.

[29] M. R. Garey, R. L. Graham, D. S. Johnson, and D. Knuth. Complexity results for bandwidth minimization. *SIAM Journal on Applied Mathematics*, 34:477–495, September 1978.

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

[31] M. R. Garey, D. S. Johnson, and L. Stockmeyer. Some simplified NP-complete graph problems. *Theoretical Computer Science*, 1:237–267, 1976.

[32] F. Gavril. Some NP-complete problems on graphs. In *Proc. 11th. Conference on Information Sciences and Systems*, pages 91–95, John Hopkins Univ., Baltimore, 1977.

[33] J. Haralambides and F. Makedon. Approximation algorithms for the bandwidth minimization problem for a large class of trees. *Theory of Computing Systems*, (30):67–90, 1997.

[34] L. H. Harper. Optimal numberings and isoperimetric problems on graphs. *Journal of Combinatorial Theory*, 1(3):385–393, 1966.

[35] L. H. Harper. Stabilization and the edgesum problem. *Ars Combinatoria*, 4:225–270, December 1977.

[36] B. Hendrickson and R. Leland. The Chaco User's Guide, 1995.

[37] H. B. Hunt III, M. V. Marathe, V. Radhakrishnan, S. S. Ravi, D. J. Rosenkrantz, and R. E. Stearns. NC-approximation schemes for NP- and PSPACE-hard problems for geometric graphs. *Journal of Algorithms*, 26(2):238–274, 1998.

[38] S. Janson, T. Łuczak, and A. Rucinski. *Random Graphs*. John Wiley and Sons, New York, 2000.

[39] D. S. Johnson, C. R. Aragon, L. A. McGeoch, and C. Schevon. Optimization by simulated annealing: an experimental evaluation; part I, graph partitioning. *Operations Research*, 37(6):865–892, November 1989.

[40] M. Junguer, G. Reinelt, and G. Rinaldi. The travelling salesman problem. In C. L. Monma M. Ball, T. Magnanti and G. L. Nemhauser, editors, *Handbook on Operations Research and Management Sciences*, volume 7, pages 225–330. North–Holland, 1995.

[41] M. Juvan and B. Mohar. Optimal linear labelings and eigenvalues of graphs. *Discrete Applied Mathematics*, (36):153–168, 1992.

[42] D. R. Karger. A randomized fully polynomial approximation scheme for all terminal network realiability problem. In *27th ACM Symposium on Theory of Computing*, pages 11–17, 1996.

[43] R. M. Karp. Probabilistic analysis of partitioning algorithms for the travelling-salesman problem in the plane. *Mathematics of Operation Research*, 2:209–224, 1977.

[44] R. M. Karp. Mapping the genome: some combinatorial problems arising in molecular biology. In *24th ACM Symposium on the Theory of Computing*, pages 278–285, 1993.

[45] M. Karpinski and J. Wirtgen. On approximation hardness of the bandwidth problem. Technical Report TR97-041, Electronic Colloquium on Computational Complexity, 1997.

[46] M. Karpinski, J. Wirtgen, and A. Zelikovsky. An approximating algorithm for the bandwidth problem on dense graphs. Technical Report TR 97-017, ECCC, 1997.

[47] D. G. Kendall. Incidence matrices, interval graphs, and seriation in archeology. *Pacific Journal of mathematics*, 28:565–570, December 1969.

[48] B. W. Kernighan and S. Lin. An efficient heuristic procedure for partitioning graphs. *The Bell System Tech. Journal*, pages 291–307, February 1970.

[49] K. Lang and S. Rao. Finding Near-Optimal Cuts: An Empirical Evaluation. In *Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 212–221, 1993.

[50] M. Ledoux. Isoperimetry and gaussian analysis. In P. Groeneboom R. Dobrushin and M. Ledoux, editors, *Lectures in Probability Theory. Ecole d'Ete de Probabilites de Saint-Flour XXIV*, volume 1648 of *Lectures Notes in Mathematics*, pages 165–294. Springer-Verlag, Berlin, 1994.

[51] F. T. Leighton. *Introduction to Parallel Algorithms and Architectures: Arrays, Trees, Hypercubes.* Morgan Kaufmann, San Mateo, CA, 1993.

[52] T. Lengauer. Black-white pebbles and graph separation. *Acta Informatica*, 16:465–475, 1981.

[53] F. Makedon and I. H. Sudborough. Minimizing width in linear layouts. In J. Díaz, editor, *Proc. 10th. International Colloquium on Automata, Languages and Programming*, pages 478–490. Springer-Verlag, Lectures Notes in Computer Science, 1983.

[54] G. Mitchison and R. Durbin. Optimal numberings of an $n \times n$ array. *SIAM Journal on Discrete Mathematics*, 7(4):571–582, 1986.

[55] B. Monien. The bandwidth minimization problem for caterpillars with hair length 3 is NP-complete. *SIAM Journal on Algebraic and Discrete Methods*, 7(4):505–512, 1986.

[56] B. Monien and I. H. Sudborough. Min cut is NP-complete for edge weighted trees. In L. Kott, editor, *Proc. 13th Internatinal Colloquium on Automata, Languages and Programming*, number 226 in Lecture Notes in Computer Science, pages 265–274. Springer-Verlag, Lectures Notes in Computer Science, 1986.

[57] B. Monien and I. H. Sudborough. Min cut is NP-complete for edge weighted trees. *Theoretical Computer Science*, 58(1–3):209–229, June 1988.

[58] R. Motwani and P. Raghavan. *Randomized Algorithms.* Cambridge University Press, 1995.

[59] D. O. Muradyan and T. E. Piliposjan. Minimal numberings of vertices of a rectangular lattice. *Akad. Nauk. Armjan. SRR*, 1(70):21–27, 1980. In Russian.

[60] P. Mutzel. A polyedral approach to planar augmentation and related problems. In P. Spirakis, editor, *European Symposium on Algorithms*, volume 979 of *Lecture Notes in Computer Science*, pages 497–507. Springer-Verlag, 1995.

[61] J. Pach, F. Shahrokhi, and M. Szegedy. Applications of the crossing number. *Algoritmica*, 16:111–117, 1996.

[62] C. Papadimitriou. The NP-completeness of the bandwidth minimization problem. *Computing*, 16:263–270, 1976.

[63] C. Papadimitriou. *Computational Complexity.* Addison-Wesley, Reading, Mass., 1994.

[64] C. H. Papadimitriou and M. Sideri. The bisection width of grid graphs. In *First ACM-SIAM Symposium on Discrete Algorithms*, pages 405–410, San Francisco, 1990.

[65] M. D. Penrose. Vertex ordering and partitioning problems for random spatial graphs. *Ann. Appl. Probab.*, 10(2):517–538, 2000.

[66] J. Petit. Combining Spectral Sequencing with Simulated Annealing for the MinLA Problem: Sequential and Parallel Heuristics. Technical Report LSI-97-46-R, Departament de Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, 1997.

[67] J. Petit. Approximation Heuristics and Benchmarkings for the MinLA Problem. In R. Battiti and A. Bertossi, editors, *Alex '98 — Building bridges between theory and applications*, http://www.lsi.upc.es/~jpetit/MinLA/Benchmark, February 1998. Università di Trento.

[68] R. Preis and R. Diekmann. *The PARTY Partitioning Library User Guide.* University of Paderborn, 1996. http://www.uni-paderborn.de/fachbereich/AG/monien/RESEARCH/PART/party.html.

[69] S. Rao and A. W. Richa. New Approximation Techniques for Some Ordering Problems. In *9th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 211–218, 1998.

[70] R. Ravi, A. Agrawal, and P. Klein. Ordering problems approximated: single-processor scheduling and interval graph completition. In M. Rodriguez J. Leach, B. Monien, editor, *18th. International Colloquium on Automata, Languages and Programming*, volume 510 of *Lecture Notes in Computer Science*, pages 751–762. Springer-Verlag, 1991.

[71] Y. Shiloach. A minimum linear arrangement algorithm for undirected trees. *SIAM Journal on Computing*, 8(1):15–31, February 1979.

[72] M. T. Shing and T. C. Hu. Computational complexity of layout problems. In T. Ohtsuki, editor, *Layout design and verification*, pages 267–294, Amsterdam, 1986. North-Holland.

[73] J. M. Steele. *Probability theory and Combinatorial Optimization.* SIAM CBMS-NSF Regional Conference Series in Applied Mathematics, 1997.

[74] L. G. Valiant. Universality considerations in VLSI circuits. *Institute of Electrical and Electronics Engineers. Transactions on Computers*, 30(2):135–140, 1981.

[75] M. Yannakakis. A polynomial algorithm for the min cut linear arrangement of trees. In *24th IEEE Symposium on Foundations of Computer Science*, pages 274–281, Providence RI, November 1983.

[76] J. E. Yukich. *Probability Theory of classical Euclidian Optimization Problems.* Springer–Verlag, Heidelberg, 1998.