Linear Orderings of Random Geometric Graphs (*Extended Abstract*) *

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

February 19, 1999

Abstract

In *random geometric graphs*, vertices are randomly distributed on $[0,1]^2$ and pairs of vertices are connected by edges whenever they are sufficiently close together. Layout problems seek a linear ordering of the vertices of a graph such that a certain measure is minimized. In this paper, we study several layout problems on random geometric graphs. The layout problems that we consider are: *Bandwidth*, *Minimum Linear Arrangement*, *Minimum Cut*, *Minimum Sum Cut*, *Vertex Separation* and *Bisection*. We first prove that some of these problem remain **NP**-complete even for geometric graphs. Afterwards, we compute lower bounds that hold with high probability on random geometric graphs. Finally, we characterize the probabilistic behavior of the lexicographic ordering for our layout problems on the class of random geometric graphs.

*This research was partially supported by ESPRIT LTR Project no. 20244 — ALCOM-IT, CICYT Project TIC97-1475-CE, and CIRIT project 1997SGR-00366.

†Departament de Llenguatges i Sistemes Informàtics. Universitat Politècnica de Catalunya. Campus Nord C6. c/ Jordi Girona 1-3. 08034 Barcelona (Spain). {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 Layout Problems. Their goal is to find a linear ordering (layout) of the nodes of an input graph such that a certain measure is minimized. Layout problems are an important class of problems with many different applications in Computer Science [4], Biology [15], Archaeology [3] and Linear Algebra [22]. Finding an optimal layout is **NP**-hard in general, and therefore it is natural to develop and analyze efficient methods that give good approximations in practice. However, evaluating heuristics as simulated annealing, greedy algorithms or spectral methods is a hard task [20, 21].

A standard way of analyzing the efficiency of an heuristic algorithm is to evaluate its performance on random instances. Two classes of random instances have been widely used in the literature to enable comparisons of algorithms for layout and partitioning problems: *random graphs* and *random geometric graphs*. We denote the former class by $\mathcal{G}_{n,p}$, where n represents the number of nodes and p is the probability of the existence of each possible edge. Random graphs $\mathcal{G}_{n,p}$ have received much attention and together with the probabilistic method have become a powerful tool in combinatorics (see e.g. [1]). The approximation properties of sparse random graphs for different layout problems are considered in [7, 22] and partitioning algorithms for random graphs are studied in [5, 6]. On the other hand, we denote the class of random geometric graphs as $\mathcal{G}_n(r)$, where n is the number of vertices and r is called the radius. The vertices of a random geometric graph correspond to n points randomly distributed on the unit square. Each of its possible edges appears if and only if the distance between their two end-points is at most r. Random geometric graphs are considered a relevant abstraction to model graphs that occur in practice in real applications, such as finite element graphs, VLSI circuits, and communication graphs [11, 12]. Moreover, since for many problems standard random graphs do not serve to differentiate good from bad heuristics [7, 6, 22], random geometric graphs offer a good alternative. Even though many empirical studies have used random models of geometric graphs [11, 21, 12], its theoretical study has mainly focussed on parameters as their clique number or chromatic number, or in their connectivity properties [19].

In this paper, we are concerned with bounds for several layout measures on random geometric graphs. The layout problems that we consider are: *Bandwidth*, *Minimum Linear Arrangement*, *Minimum Cut*, *Minimum Sum Cut*, and *Vertex Separation*. We also consider the *Bisection* problem, which is a partitioning problem, but can be also treated as a layout problem. All these problems, formally defined in Section 2, are **NP**-complete. Moreover, we prove that some of them remain **NP**-complete even for geometric instances. In Section 3, we compute lower bounds that hold with high probability on random geometric graphs. Afterwards, we obtain tight bounds on the cost of the *projection ordering* that is obtained by the projection of each node of a given random geometric graph into the x-axis. Section 4 analyzes this ordering. Our main result is the fact that the projection ordering is, with high probability, a constant approximation algorithm for our layout problems on the class of random geometric graphs considered here.

2 Definitions and complexity results

We always consider undirected graphs without self loops. A *layout* φ on a graph $G = (V, E)$ is a one-to-one function $\varphi : V \to [n] = \{1, \ldots, n\}$ with $n = |V|$. Given a graph G and a layout φ on G, let us define:

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

The problems we consider and their associated measures are:

- Bandwidth (BANDWIDTH): Given a graph $G = (V, E)$, find $\text{MINBW}(G) = \min_\varphi \text{BW}(\varphi, G)$ where $\text{BW}(\varphi, G) = \max_{uv\in E} \lambda(uv, \varphi, G)$.
- Minimum Linear Arrangement (MINLA): Given a graph $G = (V, E)$, find $\text{MINLA}(G) = \min_\varphi \text{LA}(\varphi, G)$ where $\text{LA}(\varphi, G) = \sum_{uv\in E} \lambda(uv, \varphi, G) = \sum_{i=1}^n |\theta(i, \varphi, G)|$.
- Minimum Cut Width (MINCUT): Given a graph $G = (V, E)$, find $\text{MINCUT}(G) = \min_\varphi \text{CUT}(\varphi, G)$ where $\text{CUT}(\varphi, G) = \max_{i=1}^n |\theta(i, \varphi, G)|$.
- Vertex Separation (VERTSEP): Given a graph $G = (V, E)$, find $\text{MINVS}(G) = \min_\varphi \text{VS}(\varphi, G)$ where $\text{VS}(\varphi, G) = \max_{i=1}^n |\delta(i, \varphi, G)|$.
- Minimum Sum Cut (MINSUMCUT): Given a graph $G = (V, E)$, find $\text{MINSC}(G) = \min_\varphi \text{SC}(\varphi, G)$ where $\text{SC}(\varphi, G) = \sum_{i=1}^n |\delta(i, \varphi, G)|$.
- Bisection (BISECTION): Given a graph $G = (V, E)$, find $\text{MINBIS}(G) = \min_\varphi \text{BIS}(\varphi, G)$ where $\text{BIS}(\varphi, G) = |\theta(\lfloor n/2 \rfloor, \varphi, G)|$.

It is well known that all the above problems are **NP**-complete for general graphs [9, 10, 13].

We introduce now several classes of geometric graphs on the plane. These graphs depend on which kind of norm is used to measure distances. Under the l_2 norm (the Euclidean norm), the distance between two points (x_1, y_1) and (x_2, y_2) is $((x_1 - x_2)^2 + (y_1 - y_2)^2)^{1/2}$. Under the l_∞ norm, their distance is $\max\{|x_1 - x_2|, |y_1 - y_2|\}$.

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 (in the graph) if and only their corresponding disks intersect (in the plane). 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 both in l_2 and l_∞: it suffices to associate each node of the grid with a disk or a square (see Figure 1).

We define the class of *random geometric graphs* $\mathcal{G}_n(r_n)$ as the graphs of n nodes that can be obtained from the following experiment: Let the set $\mathcal{X}_n$ consist of n points sampled uniformly and independently at random from the unit square ($[0, 1]^2$); the nodes of the graph correspond to those points, and the edges of the graph connect pairs of distinct points whose distance is at most r_n. Random geometric graphs induce a probability distribution on unit disk graphs. Observe that, under this distribution, grid graphs have some positive probability.

All through this paper, we use the l_∞ norm. Furthermore, in the following we restrict our attention to the case

$$r_n = \sqrt{\frac{a_n}{n}} \quad \text{where} \quad a_n = b_n \log n \quad \text{with} \quad b_n \to \infty \quad \text{and} \quad b_n = \mathcal{O}\left(\sqrt{\log n}\right).$$

It is important to remark that through this choice, the construction of sparse but connected graphs is guaranteed: Define the connectivity distance ρ_n of a random geometric graph by

$\rho_n = \inf\{r \mid G \in \mathcal{G}_n(r) \text{ is connected}\}$. It is known [2] that as $n \to \infty$, $\left(\sqrt{n/\log n}\right)\rho_n$ converges to $\frac{1}{2}$ almost surely.

Complexity results. For the rest of this section, we will consider the decisional counterparts of the optimization problems previously defined. Let us show now that some of the layout problems we consider are still hard to solve efficiently, even when restricted to geometric instances.

Theorem 1. Bandwidth, MinCut and VertSep remain **NP**-complete even when restricted to grid graphs (and therefore, even when restricted to unit disk graphs).

We could not obtain similar results for MinSumCut, MinLA and Bisection. However, for the Bisection problem, we are able to give a weak result:

Theorem 2. If Bisection is **NP**-complete even when restricted to planar graphs with maximum vertex degree 4, then Bisection is **NP**-complete even when restricted to unit disk graphs.

The proof of these results is given in the Appendix. Remark that Papadimitriou and Sideri [18] conjecture the hypothesis of Theorem 2, which is an important open problem.

3 Lower bounds

In this section we find asymptotic lower bounds for the optimum cost of our various layout problems. As said, we take $r_n = \sqrt{a_n/n}$ where $a_n = b_n \log n$ with $b_n \to \infty$ as $n \to \infty$ and $b_n = \mathcal{O}\left(\sqrt{\log n}\right)$. We consider a collection $\mathcal{X}_n$ of n points independently uniformly distributed in the unit square $[0,1]^2$. Consider $\lfloor 2/r_n \rfloor^2$ little boxes of size $\frac{1}{2}r_n \times \frac{1}{2}r_n$ placed packed in $[0,1]^2$ starting at $(0,0)$. Notice that, by construction, any two points of $\mathcal{X}_n$ in neighboring boxes will be connected by an edge in the geometric graph induced by $\mathcal{X}_n$.

Definition 1. Given $\epsilon \in \left(0, \frac{1}{2}\right)$, let us say that a configuration of n points in the unit square is *ϵ-nice* if every box has at least $\frac{1}{4}(1-\epsilon)a_n$ points and at most $\frac{1}{4}(1+\epsilon)a_n$ points.

Using Chernoff's bounds and Boole's inequality, it is possible to show that, given any $\epsilon \in \left(0, \frac{1}{2}\right)$, $\mathbf{Pr}[\mathcal{X}_n \text{ is } \epsilon\text{-nice}] \to 1$ as n tends to infinity.

Proposition 1. Let G be an ϵ-nice geometric graph with n nodes and $\epsilon \in \left(0, \frac{1}{2}\right)$. Let φ be any layout of G. Then for any integer $i \in \left\{\lceil \frac{1}{4}n \rceil, \ldots, \lfloor \frac{3}{4}n \rfloor\right\}$, we have $|\theta(i,\varphi,G)| \geq \frac{(1-\epsilon)^2}{32} n^{1/2} a_n^{3/2}$.

Proof. Consider n points in an ϵ-nice configuration, and take an arbitrary ordering φ of the n points. Let $\alpha \in \left[\frac{1}{4}, \frac{1}{2}\right]$ and let $i = \lfloor \alpha n \rfloor$. Call the first $\lfloor \alpha n \rfloor$ points in the ordering "red" and the last $\lceil (1-\alpha)n \rceil$ "green". Let R_n be the set of boxes containing at least $\frac{1}{4}\beta a_n$ red points (red boxes), and let G_n be the set of boxes containing at most $\frac{1}{4}\beta a_n$ red points (green boxes) for $\beta = \frac{1}{2}(1-\epsilon)$. Define u, v, w and t such that un is the number of green points in red boxes, vn is the number of red points in red boxes, wn is the number of red points in green boxes, and tn is the number of green points in green boxes. According to these definitions, $v + w = \alpha$ and $u + t = 1 - \alpha$. Moreover, as G is ϵ-nice, each box contains $\frac{1}{4}(1-\epsilon)a_n$ or more points. Therefore, each green box contains at least $\frac{1}{4}(1-\epsilon-\beta)a_n = \frac{1}{4}\beta a_n$ green points.

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.

Let us count the contribution C_1 to $|\theta(i,\varphi,G)|$ from within-box edges that are within green boxes. Each red point in a green box is connected, at least, to all the green points of its own box, and there are wn red points in all the green boxes. As a consequence, $C_1 \geq \frac{1}{4}wn\beta a_n$.

Let us count the contribution C_2 to $|\theta(i,\varphi,G)|$ from within-box edges that are within red boxes. Each green point in a red box is connected, at least, to all the red points of its own box, and there are un green points in all the red boxes. As a consequence, $C_2 \geq \frac{1}{4}un\beta a_n$.

Let us count the contribution C_3 to $|\theta(i,\varphi,G)|$ of between-box edges. As, by ϵ-niceness, no box can contain more than $\frac{1}{4}(1+\epsilon)a_n$ points and there are vn red points in red boxes and tn green points in green boxes, we have

$$|R_n| \geq \frac{vn}{\frac{1}{4}(1+\epsilon)a_n} = \frac{4v}{(1+\epsilon)r_n^2} \quad \text{and} \quad |G_n| \geq \frac{tn}{\frac{1}{4}(1+\epsilon)a_n} = \frac{4t}{(1+\epsilon)r_n^2}.$$

Let ∂G denote the number of pairs of neighbor boxes of opposite colors in G. We have

$$C_3 \geq \partial G(\tfrac{1}{4}\beta a_n)(\tfrac{1}{4}\beta a_n) = \partial G\tfrac{(1-\epsilon)^2}{64}a_n^2.$$

As, in our case, the triangular and the rectangular shapes starting in a corner of the unit square are the ones that minimize the frontier while maximizing the area,

$$\begin{aligned}\partial G &\geq \min\left\{3\sqrt{1+8|R_n|}-1,\ 3\sqrt{1+8|G_n|}-1,\ \frac{2}{r_n}\right\}\\ &\geq \frac{2}{r_n}\min\left\{\sqrt{\frac{18v}{1+\epsilon}},\ \sqrt{\frac{18t}{1+\epsilon}},\ 1\right\}.\end{aligned}$$

Using the results obtained so far, we obtain

$$|\theta(i,\varphi,G)| \geq C_1+C_2+C_3 \geq A_nu + A_nw + B_n\min\left\{\sqrt{\frac{18v}{1+\epsilon}},\ \sqrt{\frac{18t}{1+\epsilon}},\ 1\right\}$$

where $A_n = \frac{1-\epsilon}{8}na_n$ and $B_n = \frac{(1-\epsilon)^2}{32}a_n^{3/2}n^{1/2}$.

Remember that $u+t = 1-\alpha$, $v+w=\alpha$ and $\alpha \in \left[\frac{1}{4},\frac{1}{2}\right]$. When $t<(1-\alpha)(1-\epsilon)$ we have $u \geq \frac{1}{2}\epsilon$. When $v<\alpha(1-\epsilon)$, we have $w\geq \frac{1}{4}\epsilon$. In both cases

$$|\theta(i,\varphi,G)| \geq A_nu + A_nw + B_n\min\left\{\sqrt{\frac{18v}{1+\epsilon}},\ \sqrt{\frac{18t}{1+\epsilon}},\ 1\right\} \geq A_n\tfrac{1}{4}\epsilon.$$

Finally, when $v \geq \alpha(1-\epsilon)$ and $t \geq (1-\alpha)(1-\epsilon)$, we have $\sqrt{18v/(1+\epsilon)} \geq \sqrt{6\alpha}$ and $\sqrt{18t/(1+\epsilon)} \geq \sqrt{6(1-\alpha)}$ because $\epsilon \in \left(0,\frac{1}{2}\right)$. As $\alpha \in \left[\frac{1}{4},\frac{1}{2}\right]$, in this case we obtain that

$$\begin{aligned}|\theta(i,\varphi,G)| &\geq A_nu + A_nw + B_n\min\left\{\sqrt{\frac{18v}{1+\epsilon}},\ \sqrt{\frac{18t}{1+\epsilon}},\ 1\right\}\\ &\geq B_n\min\left\{\sqrt{6\alpha},\ \sqrt{6(1-\alpha)},\ 1\right\} = B_n.\end{aligned}$$

As A_n grows faster than B_n since $a_n/n \to 0$, joining these three cases we get

$$|\theta(i,\varphi,G)| \geq \min\left\{A_n\tfrac{1}{4}\epsilon,\ A_n\tfrac{1}{2}\epsilon,\ B_n\right\} \geq B_n$$

for n big enough. The use of symmetry proves the claimed proposition for $\alpha \in \left[\frac{1}{2},\frac{3}{4}\right]$. $\square$

Theorem 3 (Lower bounds). Let G be an ϵ-nice geometric graph with $\epsilon \in (0,1)$ with n vertices. Then, the following lower bounds hold:

$$\frac{\text{MINBIS}(G)}{n^{1/2}a_n^{3/2}} \geq \frac{(1-\epsilon)^2}{32} \tag{lb1}$$

$$\frac{\text{MINCUT}(G)}{n^{1/2}a_n^{3/2}} \geq \frac{(1-\epsilon)^2}{32} \tag{lb2}$$

$$\lim_{n\to\infty} \frac{\text{MINLA}(G)}{n^{3/2}a_n^{3/2}} \geq \frac{(1-\epsilon)^2(4\epsilon-1)}{64(2\epsilon-1)} \tag{lb3}$$

$$\frac{\text{MINVS}(G)}{n^{1/2}a_n^{1/2}} \geq \frac{(1-\epsilon)^2}{72(\epsilon+1)} \tag{lb4}$$

$$\lim_{n\to\infty} \frac{\text{MINSC}(G)}{n^{3/2}a_n^{1/2}} \geq \frac{(\epsilon-1)^2(4\epsilon-1)}{144(1+\epsilon)(2\epsilon-1)} \tag{lb5}$$

$$\frac{\text{MINBW}(G)}{n^{1/2}a_n^{1/2}} \geq \frac{(1-\epsilon)^2}{72(\epsilon+1)} \tag{lb6}$$

Proof. The proof of (lb1) and (lb2) is directly obtained from Proposition 1. To prove (lb3), take any layout φ; using Proposition 1, we have:

$$\begin{aligned}
\text{LA}(\varphi, G) &= \sum_{i=1}^{n} |\theta(i,\varphi,G)| \geq \sum_{i=\lceil n/4 \rceil}^{\lfloor 3n/4 \rfloor} |\theta(i,\varphi,G)| \\
&\geq \left(\tfrac{3}{4}n - \tfrac{1}{4}n - 2\right) \tfrac{(1-\epsilon)^2}{32} n^{1/2} a_n^{3/2} = \tfrac{(1-\epsilon)^2(1-\epsilon)^2}{32} n^{3/2} a_n^{3/2} (1 - o(1)).
\end{aligned}$$

To prove (lb4), let Δ be the degree of the graph (i.e. the maximum degree of its vertices). Then, for any layout φ and any $i \in [n]$, we have $|\delta(i,\varphi,G)| \geq |\theta(i,\varphi,G)|/\Delta$. For any ϵ-nice graph, $\Delta \leq \frac{9}{4}(1+\epsilon)a_n$. Therefore, by Proposition 1, for any layout φ and any i with $\frac{1}{4}n < i < \frac{3}{4}n$, we have $|\delta(i,\varphi,G)| \geq \frac{(1-\epsilon)^2}{72(1+\epsilon)} n^{1/2} a_n^{1/2}$.

We can now prove (lb5):

$$\begin{aligned}
\text{SC}(\varphi, G) &= \sum_{i=1}^{n} |\delta(i,\varphi,G)| \geq \sum_{i=\lceil n/4 \rceil}^{\lfloor 3n/4 \rfloor} |\delta(i,\varphi,G)| \geq \sum_{i=\lceil n/4 \rceil}^{\lfloor 3n/4 \rfloor} |\theta(i,\varphi,G)|/\Delta \\
&\geq \left(\tfrac{3}{4}n - \tfrac{1}{4}n - 2\right) \tfrac{(1-\epsilon)^2}{32} n^{1/2} a_n^{3/2} \tfrac{4}{9(1+\epsilon)a_n} = \tfrac{(1-\epsilon)^2(1-\epsilon)^2}{72(1+\epsilon)} n^{3/2} a_n^{1/2} (1 - o(1)).
\end{aligned}$$

Finally, let us prove that (lb6) holds. We have already shown that for any layout φ and any i with $\frac{1}{4}n < i < \frac{3}{4}n$, we have $|\delta(i,\varphi,G)| \geq \frac{(1-\epsilon)^2}{72(\epsilon+1)} n^{1/2} a_n^{1/2}$. Thus, before the node at position i, $\frac{1}{4}n < i < \frac{3}{4}n$, there have to be at least $|\delta(i,\varphi,G)|$ nodes, all of them connected with some other nodes located after the position i. So, the first of these $|\delta(i,\varphi,G)|$ nodes must have an edge that jumps at least $|\delta(i,\varphi,G)|$ nodes. In other words, for any layout φ there is an edge $uv \in E(G)$ with $\lambda(uv,\varphi,G) \geq |\delta(i,\varphi,G)|$. □

4 The projection ordering

In this section, we characterize the behavior of the projection ordering. Recall that the projection ordering is obtained through the projection of the nodes on the x-axis. Another way to see this ordering is to sweep a vertical line starting from $x = 0$ to $x = 1$, numbering vertices in the order the line touches them. As in the previous section, we consider a collection $\mathcal{X}_n$ of n points independently and uniformly distributed in the unit square $[0,1]^2$ and work in the case $r_n = \sqrt{a_n/n}$ where $a_n = b_n \log n$ with $b_n \to \infty$ and $b_n = \mathcal{O}\left(\sqrt{\log n}\right)$. The coordinates of a point u are denoted $x(u)$ and $y(u)$. We dissect the unit square in boxes of size $\gamma r_n \times \gamma r_n$ with $\gamma = 1/k$ for some large enough integer k. Let $t = n/(\gamma^2 a_n)$ denote the total number of boxes. Without loss of generality, we will suppose that $1/(\gamma r_n)$ and t are integers.

Definition 2. A set $\mathcal{X}_n$ of n points in $[0,1]^2$ is said to be γ-good if every box contains no less than $p_- = (1-\gamma)\gamma^2 a_n$ points and no more than $p_+ = (1+\gamma)\gamma^2 a_n$ points. In this case, the random geometric graph induced by $\mathcal{X}_n$ is also said to be γ-good.

Later we will prove that with high probability, random geometric graphs are γ-good for any $\gamma \in (0, \frac{1}{2})$. The behavior of the Projection ordering on γ-good graphs is characterized by the following result:

Theorem 4. Let G_n be a sequence of γ-good graphs with n vertices, and let π be the projection layout on G_n. Then, for any $\epsilon \in (0, 1/2)$ and for any measure $f \in \{\text{BW}, \text{VS}, \text{SC}, \text{CUT}, \text{BIS}, \text{LA}\}$, we have

$$1 - \epsilon \leq \left(\lim_{n\to\infty} \frac{f(\pi, G_n)}{A_f} \right) \leq 1 + \epsilon$$

where

$$A_{\text{BW}} = n^{1/2} a_n^{1/2}, \qquad A_{\text{VS}} = n^{1/2} a_n^{1/2}, \qquad A_{\text{SC}} = n^{3/2} a_n^{1/2},$$
$$A_{\text{CUT}} = n^{1/2} a_n^{3/2}, \qquad A_{\text{BIS}} = n^{1/2} a_n^{3/2}, \qquad A_{\text{LA}} = n^{3/2} a_n^{3/2}.$$

4.1 Upper bounds on the projection ordering

Definition 3. Consider the geometric graph G induced by $\mathcal{X}_n$. Given a node u from 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 the π on uv, that is, the number of nodes w such that $x(u) < x(w)$ and $x(w) < x(v)$.

Lemma 1. For any node u and for any edge uv of any γ-good graph, we have that

$$\theta(u) \leq c_\theta(\gamma) \cdot n^{1/2} a_n^{3/2} \quad \text{and} \quad \lambda(uv) \leq c_\lambda(\gamma) \cdot n^{1/2} a_n^{1/2}$$

where $\lim_{\gamma\to 0} c_\theta(\gamma) = 1$ and $\lim_{\gamma\to 0} c_\lambda(\gamma) = 1$.

Proof. Every possible edge is between boxes with centers at distance at most r_n. Thus,

$$\theta(u) \leq \sum_{i=0}^{1/\gamma} \left(\frac{1}{\gamma} - i + 1 \right) \left(\frac{2}{\gamma} + 1 \right) \frac{1}{\gamma r_n} \cdot p_+^2 \leq c_\theta(\gamma) \cdot n^{1/2} a_n^{3/2}.$$

On the other hand, $\lambda(uv)$ 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,

$$\lambda(uv) \leq p_+ \left(\frac{1}{\gamma} + 2 \right) \frac{1}{\gamma r_n} \leq c_\lambda(\gamma) \cdot n^{1/2} a_n^{1/2}.$$

$\square$

Corollary 1. For any γ-good graph G with n nodes, the following upper bounds on the cost of the projected layout π of G hold:

$$\mathrm{CUT}(\pi, G) \leq c_\theta(\gamma) \cdot n^{1/2} a_n^{3/2} \tag{ub1}$$
$$\mathrm{BIS}(\pi, G) \leq c_\theta(\gamma) \cdot n^{1/2} a_n^{3/2} \tag{ub2}$$
$$\mathrm{BW}(\pi, G) \leq c_\lambda(\gamma) \cdot n^{1/2} a_n^{1/2} \tag{ub3}$$
$$\mathrm{LA}(\pi, G) \leq c_\theta(\gamma) \cdot n^{3/2} a_n^{3/2} \tag{ub4}$$
$$\mathrm{VS}(\pi, G) \leq c_\lambda(\gamma) \cdot n^{1/2} a_n^{1/2} \tag{ub5}$$
$$\mathrm{SC}(\pi, G) \leq c_\lambda(\gamma) \cdot n^{3/2} a_n^{3/2} \tag{ub6}$$

Proof. Bounds (ub1), (ub2) and (ub3) follow directly from Lemma 1. Bounds (ub4), (ub5) and (ub6) hold because for any layout φ, we have $\mathrm{LA}(\varphi, G) \leq n \cdot \mathrm{CUT}(\varphi, G)$, $\mathrm{VS}(\varphi, G) \leq \mathrm{BW}(\varphi, G)$ and $\mathrm{SC}(\varphi, G) \leq n \cdot \mathrm{VS}(\varphi, G)$. □

4.2 Lower bounds on the projection ordering

Lemma 2. For any sequence G_n of γ-good graphs with n vertices, the following lower bounds on the cost of the projected layout π of G_n hold:

$$\frac{\mathrm{BW}(\pi, G_n)}{n^{1/2} a_n^{1/2}} \geq c_1(\gamma) \tag{lb1}$$
$$\frac{\mathrm{VS}(\pi, G_n)}{n^{1/2} a_n^{1/2}} \geq c_2(\gamma) \tag{lb2}$$
$$\liminf_{n\to\infty} \frac{\mathrm{SC}(\pi, G_n)}{n^{3/2} a_n^{1/2}} \geq c_3(\gamma) \tag{lb3}$$
$$\liminf_{n\to\infty} \frac{\mathrm{CUT}(\pi, G_n)}{n^{1/2} a_n^{3/2}} \geq c_4(\gamma) \tag{lb4}$$
$$\liminf_{n\to\infty} \frac{\mathrm{BIS}(\pi, G_n)}{n^{1/2} a_n^{3/2}} \geq c_5(\gamma) \tag{lb5}$$
$$\liminf_{n\to\infty} \frac{\mathrm{LA}(\pi, G_n)}{n^{3/2} a_n^{3/2}} \geq c_6(\gamma) \tag{lb6}$$

where $c_i(\gamma)$ are functions that only depend on γ and such that $\lim_{\gamma\to 0} c_i(\gamma) = 1$.

Proof. Let G be a γ-good graph with n vertices. Let us prove (lb1). Consider a node u far enough from the square boundaries (u exists because of goodness). This node will be connected with some other node v which is located $k-1$ columns away from the column of u (v also exists because of goodness). The length of the edge uv in the projected layout π is certainly larger than the total number of nodes located at columns between the column of u and the column of v:

$$\lambda(\pi, uv, G) \geq p_- \cdot (k-2) \cdot \frac{1}{\gamma r_n} \geq (\gamma - 1)(2\gamma - 1) \cdot a_n^{1/2} n^{1/2}.$$

As $\mathrm{BW}(\pi, G)$ is the maximal edge length, we obtain the claimed bound.

Let us give a proof of (lb2). Consider any node u far enough of the square boundaries. All the nodes in the $k-2$ columns preceding the columns of u must be connected to some node in the next column after the column of u. Therefore,

$$\textsc{vs}(\pi, G) \geq p_{-} \cdot (k-2) \cdot \frac{1}{\gamma r_n} \geq (\gamma - 1)(2\gamma - 1) \cdot a_n^{1/2} n^{1/2}.$$

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:

$$\textsc{sc}(\pi, G) \geq p_{-} \left(\frac{1}{\gamma r_n}\right) \left(\frac{1}{\gamma r_n} - 2k\right) \left(p_{-} \cdot (k-2) \cdot \frac{1}{\gamma r_n}\right).$$

In this case,

$$\lim_{n\to\infty} \frac{\textsc{sc}(\pi, G)}{n^{3/2} a_n^{1/2}} \geq 1 - 4\gamma + 5\gamma^2 - 2\gamma^3.$$

We prove now (lb4) and (lb5). Take any node u in the central part of $[0,1]^2$. We have

$$\textsc{cut}(\pi, G) \geq \sum_{i=1}^{k-2} p_{-}^2 (k-i-1) \left(\frac{1}{\gamma r_n} - 2k\right) 2k.$$

Therefore,

$$\lim_{n\to\infty} \frac{\textsc{cut}(\pi, G)}{n^{1/2} a_n^{3/2}} \geq 1 - 5\gamma + 9\gamma^2 - 7\gamma^3 + 2\gamma^4.$$

As the $\lceil n/2 \rceil$-th node of the projected layout must be in the central part of $[0,1]^2$, we have also that $\lim_{n\to\infty} \frac{\textsc{bis}(\pi,G)}{n^{1/2} a_n^{3/2}} \geq 1 - 5\gamma + 9\gamma^2 - 7\gamma^3 + 2\gamma^4$.

Finally, let us prove (lb6). We can extend the cut width proof to all the points which are away from the left and the right borders of the unit square:

$$\textsc{la}(\pi, G) \geq p_{-} \left(\frac{1}{\gamma r_n}\right) \left(\frac{1}{\gamma r_n} - 2k\right) \cdot \sum_{i=1}^{k-2} p_{-}^2 (k-i-1) \left(\frac{1}{\gamma r_n} - 2k\right) 2k.$$

In this case,

$$\lim_{n\to\infty} \frac{\textsc{la}(\pi, G)}{n^{3/2} a_n^{3/2}} \geq 1 - 6\gamma + 14\gamma^2 - 16\gamma^3 + 9\gamma^4 - 2\gamma^5$$

$\square$

4.3 Approximability of the projection ordering

In Section 3, we have given lower bounds that hold with high probability for all the considered problems on ϵ-nice graphs. Theorem 4 characterizes the behavior of the Projection ordering on γ-good graphs. Using Chernoff's bounds and Boole's inequality, one can show that the probability of a random geometric graph to be both ϵ-nice and γ-good tends to one as n tends to infinity. Therefore, we have the following result:

Theorem 5. The projection ordering is an approximation algorithm with high probability for the Bandwidth, Minimum Linear Arrangement, Minimum Cut, Minimum Sum Cut, Vertex Separation and Bisection problems on the class $\mathcal{G}_n(r)$ with $r_n = \sqrt{a_n/n}$ where $a_n = b_n \log a_n$ with $b_n \to \infty$ and $b_n = \mathcal{O}\left(\sqrt{\log n}\right)$.

5 Conclusion

In this paper we have presented upper and lower bounds for different measures of vertex orderings. We have also shown that the projection ordering is able to deliver with high probability solutions whose cost is not more than a constant times bigger the optimum on a particular class of random geometric graphs for several layout problems. Given the importance of the considered problems and the intensive use of these graphs in experimental papers, our result fills an important gap that existed between theory and practice.

We have considered only the two-dimensional geometric graphs as most real instances belong to that case, but we think that similar results will also hold on d-dimensional spaces. Our current work is trying to generalize the results on other models of random geometric graphs. For instance, it would be interesting to understand how the optimal costs of our problems change for different radii.

References

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

[2] M.J. Appel and R.P. Russo. The connectivity of a graph on uniform points in $[0, 1]^d$. Technical Report (preprint), University of Iowa, 1996.

[3] J.E. Atkins, E.G. Boman, and B. Hendrickson. A spectral algorithm for seriation and the consecutive ones problem. *SIAM J. Comput.*, 28(1):297–310, 1999.

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

[5] R. Boppana. Eigenvalues and graph bisection: An average case analysis. In *FOCS*, pages 280–285, 1987.

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

[7] J. Díaz, J. Petit, M. Serna, and L. Trevisan. Approximating layout problems on random sparse graphs. Report de recerca LSI-98-44-R, Departament de Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, http://www.lsi.upc.es/~jpetit/Publications, 1998.

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

[9] M.R. Garey and D.S. Johnson. *Computers and Intractability. A guide to the theory of NP-completeness.* Freeman and Company, 1979.

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

[11] 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.

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

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

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

[15] G. Mitchison and R. Durbin. Optimal numberings of an $n \times n$ array. *SIAM J. Alg. Disc. Math.*, 7(4):571–582, 1986.

[16] 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.

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

[18] C.H. Papadimitriou and M. Sideri. The Bisection Width of Grids Graphs. In *ACM-SIAM Symp. on Discrete Algorithms*, volume 1, pages 405–410, San Francisco, 1990.

[19] M.D. Penrose. The longest edge of the random minimal spanning tree. *The Annals of Applied Probability*, 7(2):340–361, 1997.

[20] J. Petit. Combining Spectral Sequencing with Simulated Annealing for the MinLA Problem: Sequential and Parallel Heuristics. Report de recerca LSI-97-46-R, Departament de Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, http://www.lsi.upc.es/~jpetit/MinLA/SS+SA, November 1997.

[21] 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.

[22] J.S. Turner. On the probable performance of heuristics for bandwidth minimization. *SIAM J. Comput.*, 15(2):561–580, 1986.

[23] L.G. Valiant. Universality considerations in VLSI circuits. *IEEE Trans. Comput.*, 30(2):135–140, 1981.

A Appendix: Proofs of complexity results

In order to prove the theorems, we need to present a technical definition together with a lemma, and then quote some complexity results.

Definition 4 (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. Let H be a subdivision of a graph G. Then $\textsc{mincut}(G) = \textsc{mincut}(H)$.

Proof. We prove it for the insertion of a new node w between an edge uv of G. Let φ be an optimal layout of G where, without lost of generality, $\varphi(u) < \varphi(v)$. Let ϕ be a layout of H that corresponds to insert w just after u. Then, $\textsc{cut}(G, \varphi) = \textsc{cut}(H, \phi)$. To show that ϕ is optimal for H, let us suppose the contrary: there exists a layout ϕ' of H with $\textsc{cut}(H, \phi') < \textsc{cut}(H, \phi)$. In this case, let us build a layout φ' of G by removing w from φ'. In this case, $\textsc{cut}(G, \varphi') \leq \textsc{cut}(H, \phi') < \textsc{cut}(H, \phi) = \textsc{cut}(G, \varphi') < \textsc{cut}(G, \varphi)$ which contradicts the optimality of φ with respect to G. □

Theorem 6 ([16]). Bandwidth remains **NP**-complete even when restricted to caterpillars with at most one hair attached to each vertex of the body.

Theorem 7 ([17]). MinCut and VertSep remain **NP**-complete even when restricted to planar graphs with maximum vertex degree 3.

Proof of Theorem 1. 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 for the obvious illustration.

We present a reduction from the MinCut problem restricted to planar graphs with maximum vertex degree 3 to the MinCut problem restricted to grid graphs. Let $\langle G, K\rangle$ be an instance of MinCut restricted to planar graphs with maximum vertex degree 3. Using the algorithm of Valiant [23], one can draw G in such a way that its nodes are located at positions $(6x, 6y)$ for some $x, y \in \mathbb{N}$ and that edges only follow horizontal and vertical paths without crossing one another. This embedding only uses an area polynomial in the size of G. Then, replace each edge by a string of unit disks to produce a grid graph H. As, by construction, G' is a subdivision of G, we have $\textsc{mincut}(G) \leq K \Leftrightarrow \textsc{mincut}(G') \leq K$, which proves the claimed result. Figure 3 illustrates this reduction.

Observe that the previous reduction creates graphs with maximum degree 3 and recall that for graphs with maximum degree 3, the SearchNb problem is identical to the MinCut problem [14]. 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 [8]. 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 MinCut. As the resulting graph H is a grid graph homeomorphic to the input graph G, we have $\textsc{minvs}(H) \leq K \Leftrightarrow \textsc{minsn}(G) \leq K$. □

Proof of Theorem 2. We assume that the BISECTION problem is only valid for graphs of even order. Let $\langle G, K \rangle$ be an instance of BISECTION where G is a planar graph with n nodes and maximum vertex degree 4. We will reduce it to an instance $\langle H, K \rangle$ of BISECTION where H is a unit disk graph such that $\text{MINBIS}(G) = \text{MINBIS}(H)$. The reduction for unit square graphs is analogous.

As we did in the proof of Theorem 1, we start by embedding G on the plane using Valiant's algorithm [23] in such a way that its nodes are located at positions $(6x, 6y)$ for some $x, y \in \mathbb{N}$ and that edges only follow horizontal and vertical paths without crossing one another. Now, we identify each "original" node of the embedding with a unit disk and we 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 $V(G)$, call its "gadget" the set of disks that represent its adjacent half edges (a gadget includes the extreme disks, where it ends). Now, give to each non extreme disk multiplicity n^2 (extreme disks retain multiplicity 1) and add multiplicity to the original nodes in such a way that every gadget receives the same amount of disks. H is the resulting graph of this transformation (which can be computed in polynomial time), where disks with multiplicity m are, in fact, m different disks on the same position forming a clique.

We have to prove that $\langle G, K \rangle$ is a positive instance of BISECTION if and only if $\langle H, K \rangle$ is also a positive instance of BISECTION. 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 BISECTION then there exists a bisection B of G such that $\text{BIS}(G, B) \leq K$. 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 BISECTION.

On the other hand, if $\langle H, K \rangle$ is a positive instance of BISECTION, we have two cases: When $K > 2n$, the bisection width of G cannot exceed $2n$ (as G has maximum degree 4), thus $\langle G, K \rangle$ surely is a positive instance of BISECTION. When $K \leq 2n$, 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, one obtains that $\langle G, K \rangle$ is a positive instance of BISECTION. $\square$

Notice that this last proof works both in l_2 and l_∞. All pictures use the Euclidean norm (l_2) for readability purposes.

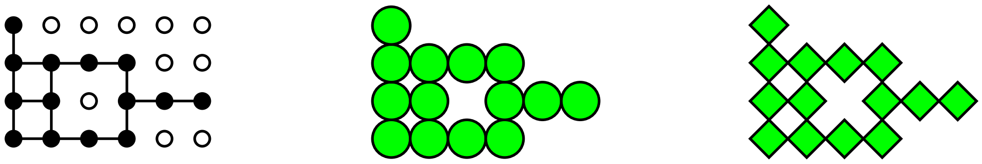

Figure 1: Any grid graph is a unit disk graph both in l_2 and l_∞.

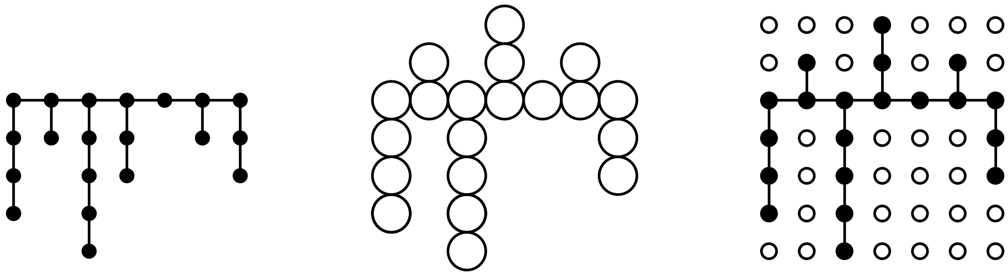

Figure 2: Any caterpillar with at most one hair attached to each vertex of the body is a unit disk graph and a grid graph.

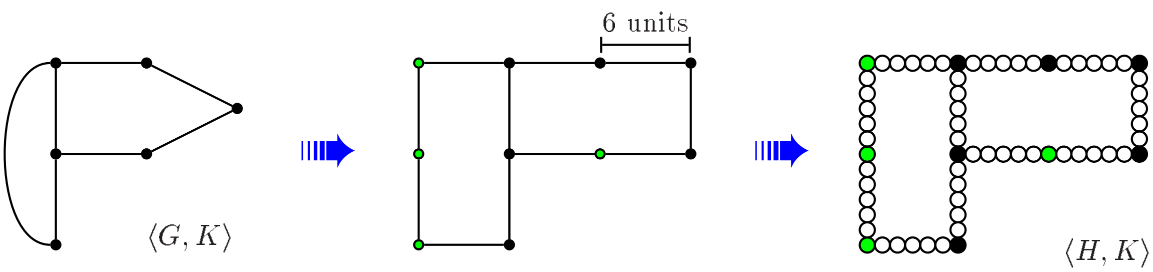

Figure 3: Reduction from MinCut restricted to planar graphs with maximum vertex degree 3 to MinCut restricted to grid graphs. At left, the input graph; at the center, the input graph embedded with Valiant's algorithm; at right, substitution of the edges with paths of disks.

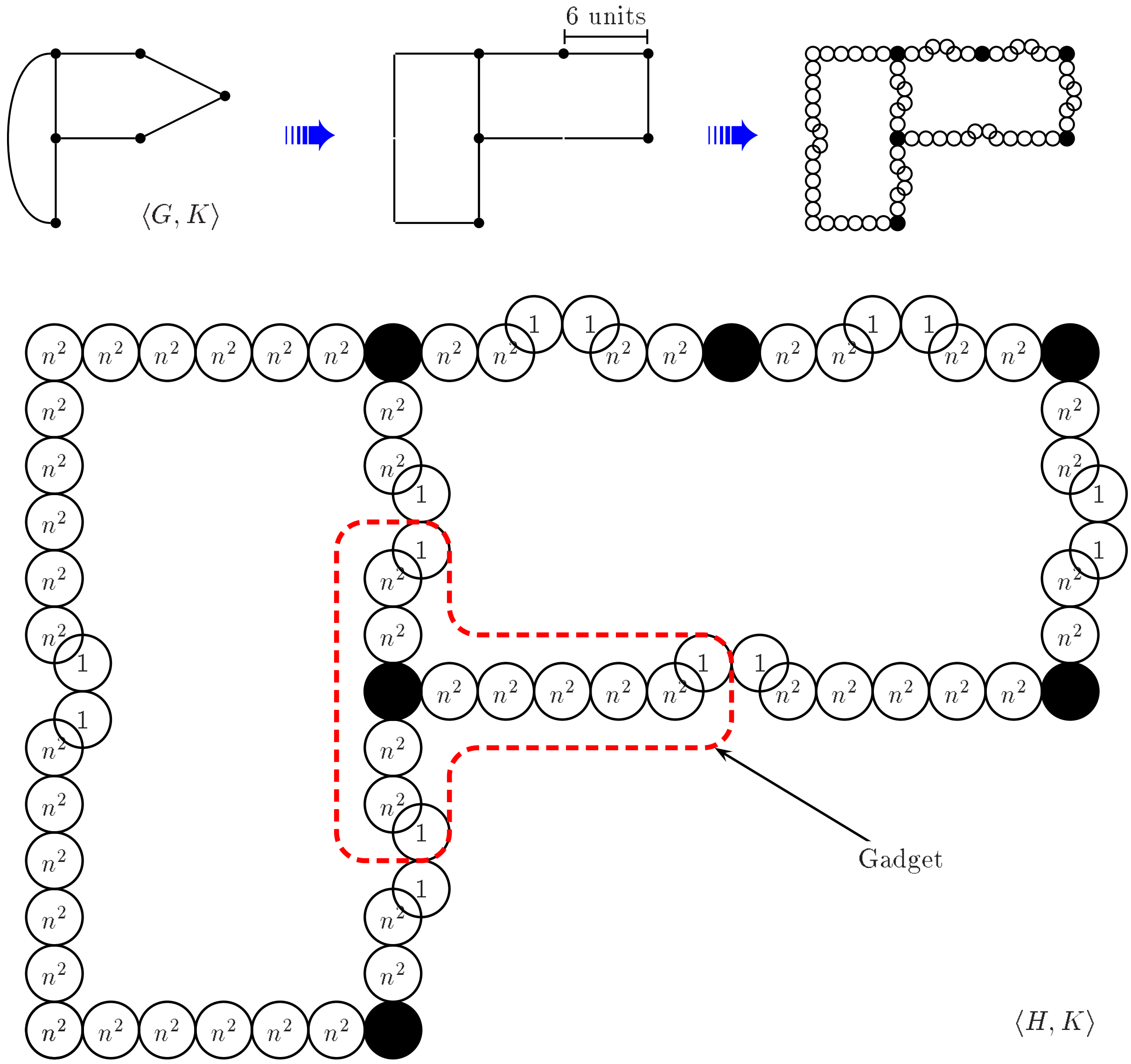

Figure 4: Reduction from BISECTION restricted to planar graphs with maximum vertex degree 4 to BISECTION restricted to unit disk graphs. At top left, the input graph with $n = 6$ nodes; at the top center, the input graph embedded with Valiant's algorithm; at top right, substitution of the edges with paths of disks with even length. At bottom, we show how non extreme disks receive multiplicy 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.