d-Dimensional Arrangement Revisited

Daniel Rotter Jens Vygen

Research Institute for Discrete Mathematics
University of Bonn

Revised version: April 5, 2013

Abstract

We revisit the d-dimensional arrangement problem and analyze the performance ratios of previously proposed algorithms based on the linear arrangement problem with d-dimensional cost. The two problems are related via space-filling curves and recursive balanced bipartitioning. We prove that the worst-case ratio of the optimum solutions of these problems is $\Theta(\log n)$, where n is the number of vertices of the graph. This invalidates two previously published proofs of approximation ratios for d-dimensional arrangement. Furthermore, we conclude that the currently best known approximation ratio for this problem is $O(\log n)$.

1 Introduction

We revisit the d-dimensional arrangement problem (d-DIMAP) for $d \in \mathbb{N}$: given an undirected graph $G = (V(G), E(G))$ and an integer $k \geq \sqrt[d]{|V(G)|}$, find an injection $p : V(G) \to \{1, \ldots, k\}^d$ minimizing $\sum_{\{v,w\} \in E(G)} ||p(v) - p(w)||_1$. Throughout this paper, we write $n = |V(G)|$ and $m = |E(G)|$, and d is a fixed constant. The case $d = 2$ is an interesting (though simplified) model of VLSI placement.

Already the case $d = 1$, known as the OPTIMAL LINEAR ARRANGEMENT PROBLEM, is *NP*-hard (Garey, Johnson and Stockmeyer [1976]). The currently best known approximation guarantee is $O(\sqrt{\log n} \log \log n)$, due to Charikar et al. [2010] and Feige and Lee [2007] (improving an earlier result of Rao and Richa [2004]).

For the general case, Hansen [1989] sketched an algorithm that recursively bipartitions the vertex set using an algorithm proposed by Leighton and Rao [1999]. The Leighton-Rao algorithm computes a c-balanced cut (i.e., the set of edges with exactly one endpoint in U for a set $U \subset V(G)$ with $cn \leq |U| \leq (1 - c)n$) that is at most $O(\log n)$ times larger than a minimum c'-balanced cut, for some constants $0 < c < c' < \frac{1}{2}$. This can lead to an $O(\log^2 n)$-approximation algorithm for d-DIMAP, although Hansen did not give a full proof.

The Leighton-Rao result was improved by Arora, Rao and Vazirani [2009], who obtained $O(\sqrt{\log n})$ instead of $O(\log n)$. Arora, Hazan and Kale [2010] obtained the same ratio by a faster algorithm. Using this algorithm for the recursive bipartitioning improves Hansen's result by a factor of $O(\sqrt{\log n})$.

Even et al. [2000] presented an $O(\log n \log\log n)$-approximation algorithm for the linear arrangement problem with d-dimensional cost (d-LAP): given a graph G, find a bijection $p : V(G) \to \{1,\dots,n\}$ such that $\sum_{\{v,w\}\in E(G)} \sqrt[d]{|p(v)-p(w)|}$ is minimized. Charikar, Makarychev and Makarychev [2007] used the result of Arora, Rao and Vazirani [2009] to obtain an $O(\sqrt{\log n})$-approximation algorithm for d-LAP for any $d \geq 2$.

Both, Even et al. [2000] and Charikar, Makarychev and Makarychev [2007], claimed that their approximation algorithm for d-LAP implies an approximation algorithm for d-DIMAP with the same performance ratio for every fixed d. The idea, proposed by Even et al. [2000], is to transform the linear arrangement into a d-dimensional arrangement according to a discrete space-filling curve; this is essentially [Even et al. [2000], Lemma 12] (except that they did not address the case $n < k^d$ explicitly):

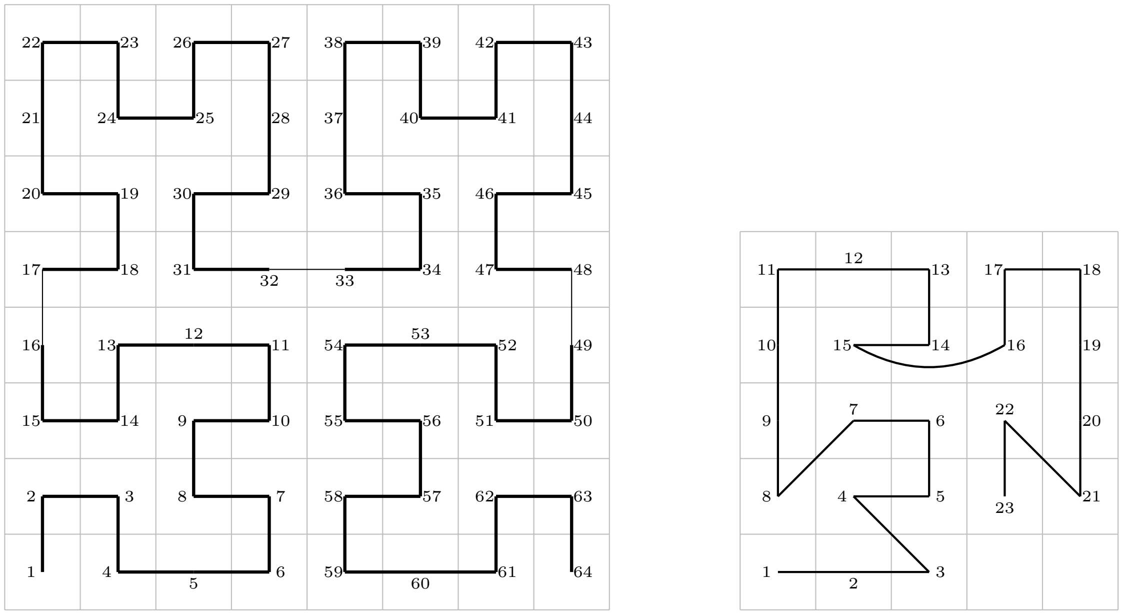
Lemma 1 (essentially Even et al. [2000], Lemma 12) *For any $n, d, k \in \mathbb{N}$ with $n \leq k^d$, there exists an injection $p : \{1,\dots,n\} \to \{1,\dots,k\}^d$ such that $||p(i)-p(j)||_1 \leq 4(d+1)\sqrt[d]{|i-j|}$ for all $i,j \in \{1,\dots,n\}$, and such a mapping can be computed in $O(n(d+\log n)+\log k)$ time.*

Our proof follows Even et al. [2000], but contains an explicit construction of a suitable space-filling curve through the d-dimensional grid, also in the case $n < k^d$.

Proof: Let $s := \lceil \log_2 \sqrt[d]{n} \rceil$ and $l := 2^s$. Consider the s-th step of the construction of the d-dimensional version of Hilbert's [1891] space-filling curve (see Sagan [1994]), say $q : \{1,\dots,l^d\} \to \{1,\dots,l\}^d$. For any $i,j \in \{1,\dots,l^d\}$ with $i \neq j$ let $t = \lceil \log_2 \sqrt[d]{|i-j|} \rceil$; then $2^{d(t-1)} < |i-j| \leq 2^{dt}$ and hence $||q(i)-q(j)||_1 < (d+1)2^t < 2(d+1)\sqrt[d]{|i-j|}$.

Let $k' = \min\{k, l\}$ and $S = \{\lfloor li/k' \rfloor : i = 1,\dots,k'\}^d$. Writing $q(j) = (q_1(j),\dots,q_d(j))$, we finally set $p(j) := (\lceil k'q_1(j')/l \rceil, \dots, \lceil k'q_d(j')/l \rceil)$ for $j = 1,\dots,n$, where $j' = \min\{i : |\{q(1),\dots,q(i)\} \cap S| = j\}$. Note that p is injective and $||p(i)-p(j)||_1 \leq ||q(i')-q(j')||_1 \leq 2(d+1)\sqrt[d]{|i'-j'|} \leq 2(d+1)\sqrt[d]{2^d|i-j|} = 4(d+1)\sqrt[d]{|i-j|}$ for any i and j. □

See Figure 1 for an example. Hence, for any graph, any solution of d-LAP can be transformed to a solution of d-DIMAP such that the cost increases at most by a factor $4(d+1)$. However, this transformation does *not* preserve the approximation ratio, as we point out in this note. This is because the optimum value of d-LAP is not bounded by a constant factor times the optimum value of d-DIMAP. A factor $\Theta(\log n)$ is lost because of the following theorem, our main result:



(a) Hilbert's curve q for $d = 2$ and $s = 3$.

(b) The resulting injection p for $d = 2$, $n = 23$, and $k = 5$.

Figure 1: left: Hilbert's curve q for $d = 2$ and $s = 3$; right: the resulting injection p for $d = 2$, $n = 23$, and $k = 5$. The figure shows the graph with edges $\{q(i), q(i+1)\}$, $i = 1, \ldots, 63$ on the left and the graph with edges $\{p(i), p(i+1)\}$, $i = 1, \ldots, 22$ on the right. Note that p results from q by considering only the points in S (which here means erasing the second, fifth, and seventh row and column), and omitting the last $k^d - n$ points.

Theorem 2 *Let $d \in \mathbb{N}$, $d \geq 2$. For any graph G and any injection $p : V(G) \to \{1, \ldots, k\}^d$ (where $k \in \mathbb{N}$), there exists a bijection $q : V(G) \to \{1, \ldots, n\}$ such that*

$$\sum_{\{v,w\} \in E(G)} \sqrt[d]{|q(v) - q(w)|} \;\leq\; O(\log n) \sum_{\{v,w\} \in E(G)} ||p(v) - p(w)||_1.$$

There are pairs (G, p) for which this bound is tight.

Consequently, the analysis of the algorithms of Even et al. [2000] and Charikar, Makarychev and Makarychev [2007] only yields approximation ratios of $O(\log^2 n \log\log n)$ and $O(\log n \sqrt{\log n})$, respectively. However, a different proof (see Section 4) shows that the algorithm of Even et al. [2000] does indeed achieve the claimed performance ratio $O(\log n \log\log n)$. Moreover, from a result of Fakcharoenphol, Rao and Talwar [2004] we can deduce the currently best known approximation ratio of $O(\log n)$; this will be shown in Section 4.

Banerjee et al. [2009] suggested a similar algorithm for $d = 2$. They claimed an approximation ratio of $O(\sqrt[4]{\log n}\sqrt{m \log\log n})$ (and a weaker ratio for hypergraphs). Unfortunately, their proof contains an error, too (the complete graph is a counterexample to [Banerjee et al. [2009], Lemma 2]). However, the claimed approximation ratio is anyway worse than the trivial $O(\sqrt{m})$, which is obtained by an arbitrary injection of the non-isolated vertices to $\{1, \ldots, \lceil\sqrt{2m}\rceil\}^2$.

We note that d-DIMAP is not known to be *MAXSNP*-hard for any $d \in \mathbb{N}$ (but see Ambühl, Mastrolilli and Svensson [2011] and Devanur et al. [2006]).

The next two sections contain a proof of Theorem 2.

2 Upper Bound

We first consider the direction needed for proving approximation ratios for d-DIMAP via d-LAP and space-filling curves.

Lemma 3 *For any graph G and any injection $p : V(G) \to \{1, \ldots, k\}^d$ (where $k, d \in \mathbb{N}$), there exists a bijection $q : V(G) \to \{1, \ldots, n\}$ such that*

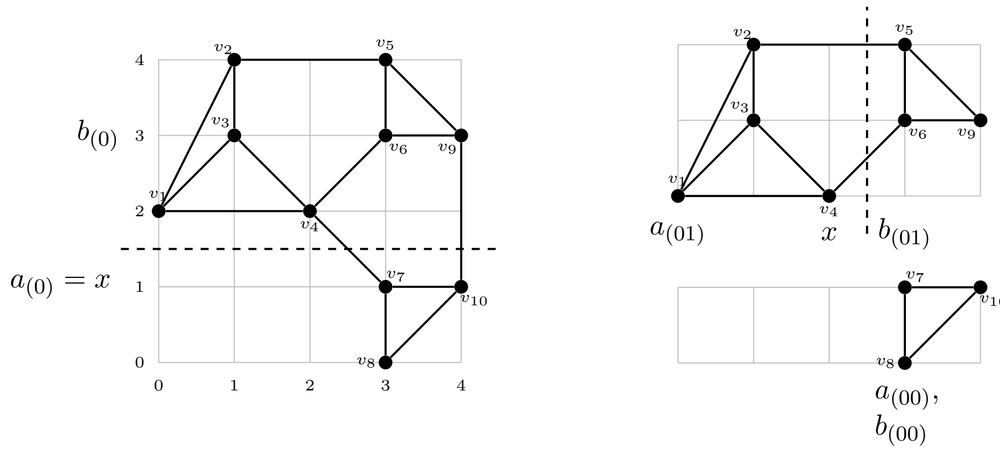
$$\sum_{\{v,w\}\in E(G)} \sqrt[d]{|q(v) - q(w)|} \;\le\; 32d \ln n \sum_{\{v,w\}\in E(G)} ||p(v) - p(w)||_1.$$

This essentially generalizes [Charikar, Makarychev and Makarychev [2007], Theorem 2.1, part I] (they consider the case where p is a one-dimensional bijection). The following proof is inspired by theirs, but the analysis is more involved. The basic idea is to partition the vertex set recursively. In each iteration, large vertex sets are partitioned into two sets of approximately equal size (up to a constant factor) according to their j-th coordinates in p, where j changes in each iteration. Then all vertices in one set will precede all vertices in the other set in q.

Proof: We write $p(v) = (p_1(v), \ldots, p_d(v))$ for $v \in V(G)$. Let $\gamma = \frac{4d^2-1}{4d^2}$. Note that $1 - \gamma^d = (1-\gamma)\sum_{i=0}^{d-1} \gamma^i \le d(1-\gamma) = \frac{1}{4d}$ and $\frac{1}{2} < \gamma^d < 1$.

We construct q as follows. Let $i := 1$ and $r_1(v) := 0$ for all $v \in V(G)$. Repeat the following until $(r_1(v), \ldots, r_i(v)) \ne (r_1(w), \ldots, r_i(w))$ for all $v, w \in V(G)$ with $v \ne w$. At termination, the lexicographical order of these vectors determines q.

In iteration i we will consider coordinates $p_{j_i}(v)$ for $v \in V(G)$, where $j_i = 1+(i \bmod d)$.



(a) Iteration 1

(b) Iteration 2

Figure 2: Example of the first two iterations of the algorithm defined in the proof of Lemma 3. Here, $n = 10$ and $d = 2$, hence $\gamma = \frac{15}{16}$ and $\gamma^d = \frac{225}{256}$. In iteration 2, we have $|X_{(0,0)}| = 1 < \frac{1}{2}\sqrt[d]{\gamma^i n/2} \approx 1.048$. Hence, we will not partition $V_{(0,0)} = \{v_7, v_8, v_{10}\}$ in iteration 2.

We write $S_i = \{(r_1(v), \ldots, r_i(v)) : v \in V(G)\}$. For each $s \in S_i$ let

$$\begin{aligned} V_s &= \{v \in V(G) : (r_1(v), \ldots, r_i(v)) = s\}, \\ a_s &= \min\{x : |\{v \in V_s : p_{j_i}(v) \geq x + 1\}| \leq \gamma^d |V_s|\}, \\ b_s &= \max\{x : |\{v \in V_s : p_{j_i}(v) \leq x\}| \leq \gamma^d |V_s|\}, \\ X_s &= \{a_s, \ldots, b_s\}. \end{aligned}$$

Note that $b_s + 1 \geq a_s$ because $|\{v \in V_s : p_{j_i}(v) \geq a_s\}| + |\{v \in V_s : p_{j_i}(v) \leq b_s + 1\}| > \gamma^d |V_s| + \gamma^d |V_s| > |V_s|$.

Let us sketch the idea behind these definitions. Partitioning V_s into the set of vertices for which the j_i-th coordinate is at most x and the rest yields sufficiently small parts if $x \in X_s$. However, we will only perform such a partitioning step if $|X_s|$ is sufficiently large, and then we will pick a coordinate $x \in X_s$ that yields the smallest cut. If $|V_s|$ is large, then $|X_s|$ will be large for at least one coordinate, and so we will make progress after at most d iterations. We will now give the details.

There are two cases. If $|X_s| \leq \frac{1}{2}\sqrt[d]{\gamma^i n/2}$, then we set $r_{i+1}(v) := 0$ for all $v \in V_s$.

Otherwise, we "split" V_s: for $x \in X_s$ and $v \in V_s$ let $r^x(v) := 0$ if $p_{j_i}(v) \leq x$ and $r^x(v) := 1$ otherwise. Choose $x \in X_s$ such that $\sum_{\{v,w\} \in E(G[V_s])} |r^x(v) - r^x(w)|$ is minimized, and set $r_{i+1}(v) := r^x(v)$ for all $v \in V_s$.

After doing this for each $s \in S_i$, we increment i. This ends the description of the procedure that ultimately defines q. See Figure 2 for an illustration.

To see that this procedure terminates, we prove that

$$|V_s| \;\leq\; \max\left\{1,\, \gamma^{i-d} n\right\} \tag{1}$$

for any $s \in S_i$ and any iteration i.

This is trivial for $i \leq d$. We proceed by induction. Let $i > d$ and $s \in S_i$. For $1 \leq h < i$ let s^h denote the prefix of s of length h, i.e., the vector resulting from s by omitting the last $i - h$ components.

Case 1: $V_s \neq V_{s^{i-d}}$. Then the set V_s resulted from splitting during at least one of the iterations $i-d, \ldots, i-1$. Then $|V_s| \leq \gamma^d |V_{s^{i-d}}|$. Since $s^{i-d} \in S_{i-d}$, we are done by induction.

Case 2: $V_s = V_{s^{i-d}}$. Then the set V_s was not split during any of the iterations $h \in \{i-d, \ldots, i-1\}$. This implies $b_{s^h} + 1 - a_{s^h} = |X_{s^h}| \leq \frac{1}{2}\sqrt[d]{\gamma^h n/2} \leq \frac{1}{2}\sqrt[d]{\gamma^{i-d} n/2}$ for $h = i-d, \ldots, i-1$.

Moreover, by the choice of a_{s^h} and b_{s^h}, we have $|\{v \in V_s : p_{j_h}(v) < a_{s^h}\}| < (1-\gamma^d)|V_s|$ and $|\{v \in V_s : p_{j_h}(v) > b_{s^h} + 1\}| < (1-\gamma^d)|V_s|$.

Combining this for $h = i-d, \ldots, i-1$ yields

$$\begin{aligned}
|V_s| \;&\leq\; |\{v \in V_s : a_{s^h} \leq p_{j_h}(v) \leq b_{s^h} + 1 \text{ for } h = i-d, \ldots, i-1\}| + \\
&\quad\; \sum_{h=i-d}^{i-1} \left(|\{v \in V_s : p_{j_h}(v) > b_{s^h} + 1\}| + |\{v \in V_s : p_{j_h}(v) < a_{s^h}\}|\right) \\
&<\; 2d(1-\gamma^d)|V_s| + |\{v \in V_s : a_{s^h} \leq p_{j_h}(v) \leq b_{s^h} + 1 \text{ for } h = i-d, \ldots, i-1\}| \\
&\leq\; \frac{1}{2}|V_s| + \prod_{h=i-d}^{i-1} (b_{s^h} + 2 - a_{s^h}),
\end{aligned}$$

and hence

$$|V_s| \;<\; 2 \prod_{h=i-d}^{i-1} (b_{s^h} + 2 - a_{s^h}).$$

If $\frac{1}{2}\sqrt[d]{\gamma^{i-d} n/2} < 1$, we have $b_{s^h} + 1 = a_{s^h}$ for $h = i-d, \ldots, i-1$, and conclude $|V_s| < 2$. If $\frac{1}{2}\sqrt[d]{\gamma^{i-d} n/2} \geq 1$, we have

$$\prod_{h=i-d}^{i-1} (b_{s^h} + 2 - a_{s^h}) \;\leq\; \left(1 + \frac{1}{2}\sqrt[d]{\gamma^{i-d} n/2}\right)^d \;\leq\; \left(\sqrt[d]{\gamma^{i-d} n/2}\right)^d \;=\; \gamma^{i-d} n/2,$$

and conclude $|V_s| < \gamma^{i-d} n$.

In both cases (1) is proved. Let t denote the index i of the last iteration; then $|V_s| = 1$ for all $s \in S_{t+1}$. From (1) we immediately get $t \leq \log_{1/\gamma}(\gamma^{-d}n) = d + \log_{1/\gamma} n \leq d + 4d^2 \ln n$.

Next, we compute an upper bound on the number of edges separated in one partitioning step, in iteration i for $s \in S_i$ with $|X_s| > \frac{1}{2}\sqrt[d]{\gamma^i n/2}$:

$$\begin{aligned}
&\sum_{\{v,w\}\in E(G[V_s])} |r_{i+1}(v) - r_{i+1}(w)| \\
&\qquad \leq \frac{1}{|X_s|} \sum_{x\in X_s} \sum_{\{v,w\}\in E(G[V_s])} |r^x(v) - r^x(w)| \\
&\qquad = \frac{1}{|X_s|} \sum_{\{v,w\}\in E(G[V_s])} |\{x \in X_s : p_{j_i}(v) \leq x < p_{j_i}(w) \text{ or } p_{j_i}(w) \leq x < p_{j_i}(v)\}| \\
&\qquad \leq \frac{1}{|X_s|} \sum_{\{v,w\}\in E(G[V_s])} |p_{j_i}(v) - p_{j_i}(w)| \\
&\qquad < \frac{2}{\sqrt[d]{\gamma^i n/2}} \sum_{\{v,w\}\in E(G[V_s])} |p_{j_i}(v) - p_{j_i}(w)|. \qquad (2)
\end{aligned}$$

For an edge $e = \{v, w\}$ let i_e be the smallest i such that $r_{i+1}(v) \neq r_{i+1}(w)$ differ (i.e., i_e is the index of the iteration in which e is separated). Then, both endpoints of e are in $V_{(r_1(v),\ldots,r_{i_e}(v))}$, and these vertices are placed consecutively in q (see Figure 3 for an illustration). Hence, using (1),

$$\begin{aligned}
\sum_{\{v,w\}\in E(G)} \sqrt[d]{|q(v) - q(w)|} &\leq \sum_{e=\{v,w\}\in E(G)} \sqrt[d]{|V_{(r_1(v),\ldots,r_{i_e}(v))}| - 1} \\
&< \sum_{e=\{v,w\}\in E(G)} \sqrt[d]{\gamma^{i_e - d} n} \\
&= \frac{1}{\gamma} \sum_{i=1}^{t} \sqrt[d]{\gamma^i n} \ |\{e \in E(G) : i_e = i\}| \\
&= \frac{1}{\gamma} \sum_{i=1}^{t} \sqrt[d]{\gamma^i n} \sum_{s\in S_i} \sum_{\{v,w\}\in E(G[V_s])} |r_{i+1}(v) - r_{i+1}(w)| \\
&< \frac{2\sqrt[d]{2}}{\gamma} \sum_{i=1}^{t} \sum_{s\in S_i} \sum_{\{v,w\}\in E(G[V_s])} |p_{j_i}(v) - p_{j_i}(w)| \\
&\leq \frac{2\sqrt[d]{2}}{\gamma} \sum_{i=1}^{t} \sum_{\{v,w\}\in E(G)} |p_{j_i}(v) - p_{j_i}(w)|
\end{aligned}$$

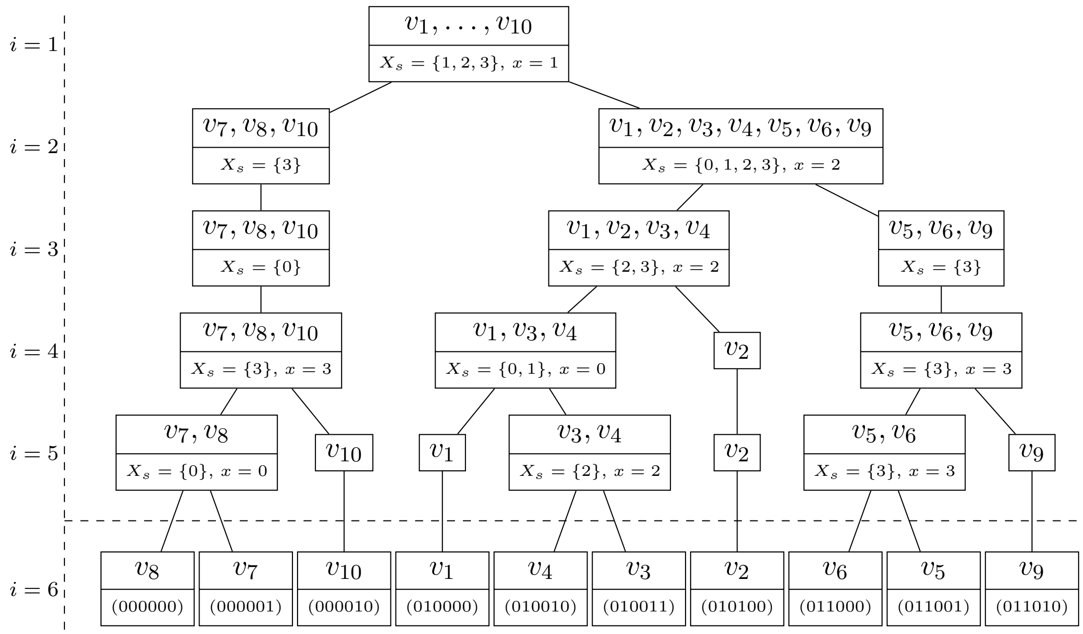


Figure 3: Visualization of the hierarchical decomposition $(\{V_s : s \in S_i\})_{i=1,\ldots,t}$ constructed in the proof of Theorem 3 on the instance of Figure 2. The resulting linear order q is the left-to-right order indicated at the bottom of the figure.

$$
\begin{aligned}
&\leq \frac{2\sqrt[d]{2}}{\gamma} \left\lceil \frac{t}{d} \right\rceil \sum_{\{v,w\} \in E(G)} ||p(v) - p(w)||_1 \\
&< \frac{8d^2 \sqrt[d]{2}}{4d^2 - 1} \lceil 1 + 4d \ln n \rceil \sum_{\{v,w\} \in E(G)} ||p(v) - p(w)||_1 \\
&\leq 32 d \ln n \sum_{\{v,w\} \in E(G)} ||p(v) - p(w)||_1.
\end{aligned}
$$

□

Charikar, Makarychev and Makarychev [2007] called a sequence $P_0, P_1, \ldots, P_t$ of partitions of $V(G)$ a *hierarchical decomposition* if $P_0 = \{V(G)\}$, $P_t = \{\{v\} : v \in V(G)\}$, and P_{i+1} is a refinement of P_i for each $i = 1, \ldots, t-1$. For a constant $0 < b < 1$, a hierarchical decomposition is called *b-balanced* if $|C| \leq b^i n$ for each $C \in P_i$. We remark that the sequence $(\{V_s : s \in S_i\})_{i=1,d+1,d+2,\ldots,t}$ defined in the proof of Lemma 3 is a γ-balanced hierarchical decomposition (see Figure 3).

3 Lower Bound

We will now show that the bound is tight up to a factor that only depends on d. The graphs that we will consider are d-dimensional grids themselves: let G_k^d be given by $V(G_k^d) = \{1,\ldots,k\}^d$ and $E(G_k^d) = \{\{x,y\} : x,y \in V(G_k^d), ||x-y||_1 = 1\}$. Note that the identity function p embeds $V(G_k^d)$ in itself with $\sum_{\{v,w\}\in E(G_k^d)} ||p(v)-p(w)||_1 = |E(G_k^d)| = d(k^d - k^{d-1}) < dn$. Therefore, the following lemma shows the lower bound, and hence, with Lemma 3, implies Theorem 2.

Lemma 4 *Let $d \geq 2$. If $q : V(G_k^d) \to \{1,\ldots,n\}$ is any bijection, then*

$$\sum_{\{v,w\}\in E(G_k^d)} \sqrt[d]{|q(v)-q(w)|} \;>\; \tfrac{3}{16}\left(1-\left(\tfrac{3}{4}\right)^{\frac{d-1}{d}}\right)\left(1-\left(\tfrac{3}{4}\right)^{1/d}\right) dn\log_2 n - \tfrac{3dn}{64}.$$

Proof: Let $G = G_k^d$, and let $q : V(G) \to \{1,\ldots,n\}$ be any bijection. Apply the procedure in the proof of Lemma 3 to q (in the role of p, for dimension $d=1$) to compute vectors $(r_1(v),\ldots,r_{t+1}(v))$ for $v \in V(G)$ with

$$\sum_{\{v,w\}\in E(G[V_s])} |r_{i+1}(v)-r_{i+1}(w)| \;<\; \frac{4}{(\frac{3}{4})^i n} \sum_{\{v,w\}\in E(G[V_s])} |q(v)-q(w)|$$

for $i = 1,\ldots,t$ and $s \in S_i$ (cf. inequality (2); note that $\gamma = \frac{3}{4}$). Hence,

$$\begin{aligned}
\sum_{e\in E(G)} \sqrt[d]{\left(\tfrac{3}{4}\right)^{i_e} n} &= \sum_{i=1}^{t} \sqrt[d]{\left(\tfrac{3}{4}\right)^{i} n}\;\; |\{e\in E(G) : i_e = i\}| \\
&= \sum_{i=1}^{t} \sqrt[d]{\left(\tfrac{3}{4}\right)^{i} n} \sum_{s\in S_i} \sum_{\{v,w\}\in E(G[V_s])} |r_{i+1}(v)-r_{i+1}(w)| \\
&< \sum_{i=1}^{t} \sqrt[d]{\left(\tfrac{3}{4}\right)^{i} n} \sum_{s\in S_i} \frac{4}{\left(\frac{3}{4}\right)^{i} n} \sum_{\{v,w\}\in E(G[V_s])} |q(v)-q(w)| \\
&= 4 \sum_{\{v,w\}\in E(G)} |q(v)-q(w)| \left(\left(\tfrac{3}{4}\right)^{i_e} n\right)^{\frac{1-d}{d}} \sum_{i=1}^{i_e} \left(\left(\tfrac{3}{4}\right)^{\frac{1-d}{d}}\right)^{i-i_e} \\
&< \frac{4\left(\frac{3}{4}\right)^{\frac{1-d}{d}}}{1-\left(\frac{3}{4}\right)^{\frac{d-1}{d}}} \sum_{\{v,w\}\in E(G)} \sqrt[d]{|q(v)-q(w)|}. \qquad (3)
\end{aligned}$$

The last inequality holds because for $e = \{v, w\} \in E(G)$ there is an $s \in S_{i_e}$ with $v, w \in V_s$ and $|q(v) - q(w)| < |V_s| \leq \left(\frac{3}{4}\right)^{i_e - 1} n$ (cf. inequality (1), implying $\left(\frac{3}{4}\right)^{i_e} n > \frac{3}{4} |q(v) - q(w)|$.

A subgraph of G with c vertices has at most $d(c - c^{1-1/d})$ edges, and this is tight if the subgraph is induced by a product of d intervals of length $c^{1/d}$. There are at most 2^{i-1} subgraphs $G[V_s]$, $s \in S_i$, each with at most $\left(\frac{3}{4}\right)^{i-1} n$ vertices.
Therefore,

$$\begin{aligned}
|\{e \in E(G) : i_e \leq i\}| &= |E(G)| - \sum_{s \in S_i} |E(G[V_s])| \\
&\geq d(n - k^{d-1}) - \sum_{s \in S_i} d\left(|V(G[V_s])| - |V(G[V_s])|^{1-1/d}\right) \\
&= \sum_{s \in S_i} d|V(G[V_s])|^{1-1/d} - dk^{d-1} \\
&\geq \frac{n}{\left(\frac{3}{4}\right)^{i-1} n} d\left(\left(\frac{3}{4}\right)^{i-1} n\right)^{1-1/d} - dk^{d-1} \\
&= \left(\left(\frac{3}{4}\right)^{(1-i)/d} - 1\right) dk^{d-1},
\end{aligned}$$

and hence,

$$\begin{aligned}
\sum_{e \in E(G)} \sqrt[d]{\left(\frac{3}{4}\right)^{i_e} n} &\geq \sum_{i=1}^{t} \left(\sqrt[d]{\left(\frac{3}{4}\right)^{i} n} - \sqrt[d]{\left(\frac{3}{4}\right)^{i+1} n}\right) |\{e \in E(G) : i_e \leq i\}| \\
&\geq \sum_{i=1}^{t} \left(\sqrt[d]{\left(\frac{3}{4}\right)^{i} n} - \sqrt[d]{\left(\frac{3}{4}\right)^{i+1} n}\right) \left(\left(\frac{3}{4}\right)^{(1-i)/d} - 1\right) dk^{d-1} \\
&= \left(1 - \left(\frac{3}{4}\right)^{1/d}\right) \left(\frac{3}{4}\right)^{1/d} \sum_{i=0}^{t-1} \left(1 - \left(\frac{3}{4}\right)^{i/d}\right) dn \\
&\geq \left(1 - \left(\frac{3}{4}\right)^{1/d}\right) \left(\frac{3}{4}\right)^{1/d} \left(t - \frac{1}{1 - \left(\frac{3}{4}\right)^{1/d}}\right) dn \\
&\geq \left(\left(1 - \left(\frac{3}{4}\right)^{1/d}\right) \log_2 n - 1\right) \left(\frac{3}{4}\right)^{1/d} dn. \qquad (4)
\end{aligned}$$

The inequalities (3) and (4) imply

$$\begin{aligned}
&\sum_{\{v,w\}\in E(G)} \sqrt[d]{|q(v)-q(w)|} \\
&\qquad > \frac{\left(\frac{3}{4}\right)^{1/d}}{4\left(\frac{3}{4}\right)^{\frac{1-d}{d}}} dn \left(\left(1-\left(\tfrac{3}{4}\right)^{\frac{d-1}{d}}\right)\left(1-\left(\tfrac{3}{4}\right)^{1/d}\right)\log_2 n - \left(1-\left(\tfrac{3}{4}\right)^{\frac{d-1}{d}}\right)\right) \\
&\qquad > \frac{3}{16} dn \left(\left(1-\left(\tfrac{3}{4}\right)^{\frac{d-1}{d}}\right)\left(1-\left(\tfrac{3}{4}\right)^{1/d}\right)\log_2 n - \tfrac{1}{4}\right).
\end{aligned}$$

□

4 Approximation Algorithms

After showing him the above proofs, Guy Even [personal communication, 2011] sent us a sketch of a revised proof of the performance ratio of the d-dimensional arrangement algorithm of Even et al. [2000]. This algorithm begins by solving the following linear program (cf. [Even et al. [2000], page 606]):

$$\begin{aligned}
\min \quad & \sum_{\{v,w\}\in E(G)} l(v,w) & & & (5)\\
\text{s.t.} \quad & \sum_{u\in U} l(u,v) \;\geq\; \frac{(|U|-1)^{1+1/d}}{4} & & \forall\, U\subseteq V(G),\ \forall\, v\in U & (6)\\
& l(u,v)+l(v,w) \;\geq\; l(u,w) & & \forall\, u,v,w\in V(G) & (7)\\
& l(v,w) \;\geq\; 0 & & \forall\, v,w\in V(G) & (8)
\end{aligned}$$

An optimum solution $l : V(G)\times V(G) \to \mathbb{R}_{\geq 0}$ of this LP can be found in polynomial time [Even et al. [2000], Section 6.1]. The following lemma strengthens [Even et al. [2000], Lemma 14] by showing that the LP (5)–(8) constitutes a lower bound for the cost of any d-dimensional arrangement, up to a constant factor.

Lemma 5 (Guy Even, personal communication 2011) *Let q^* be an optimum solution to d-DIMAP. Then $l(v,w) := 4^d(d+1)(d-1)! \cdot ||q^*(v)-q^*(w)||_1$ for $v,w\in V(G)$ defines a feasible solution to the LP (5)–(8).*

Proof: Since (7) and (8) hold evidently, we prove (6). Let $\emptyset \neq U \subseteq V(G)$ and $v\in U$, w.l.o.g. $q^*(v)=0$. If $|U|\leq 6^d$, then,

$$\sum_{u\in U} ||q^*(u)||_1 \;\geq\; |U|-1 \;>\; \frac{1}{6}(|U|-1)^{1+1/d} \;\geq\; \frac{1}{4^{d+1}(d+1)(d-1)!}(|U|-1)^{1+1/d}.$$

If $|U| > 6^d$, then let $R := \lfloor \sqrt[d]{|U|}/2 \rfloor - 1 \geq \frac{1}{4}\sqrt[d]{|U| - 1}$ and $S(d,r) := \{x \in \mathbb{Z}^d : ||x||_1 = r\}$ for $r \in \mathbb{N}$.

Observe that $|\{x \in S(d,r) : x \geq 0\}| = \binom{r+d-1}{d-1}$, and thus,

$$\frac{(r+1)^{d-1}}{(d-1)!} \;\leq\; \binom{r+d-1}{d-1} \;\leq\; |S(d,r)| \;\leq\; 2^d\binom{r+d-1}{d-1} \;\leq\; 2^d(r+1)^{d-1}.$$

Since

$$\sum_{r=1}^{R}|S(d,r)| \;\leq\; \sum_{r=1}^{R}2^d(r+1)^{d-1} \;\leq\; 2^dR(R+1)^{d-1} \;\leq\; |U|-1,$$

we have

$$\begin{aligned}
\sum_{u\in U}||q^*(u)||_1 &\geq \sum_{r=1}^{R}\sum_{u\in S(d,r)}||u||_1 \\
&= \sum_{r=1}^{R} r|S(d,r)| \\
&\geq \sum_{r=1}^{R}\frac{r^d}{(d-1)!} \\
&\geq \int_0^R \frac{x^d}{(d-1)!}\,\mathrm{d}x \\
&\geq \frac{R^{d+1}}{(d+1)(d-1)!} \\
&\geq \frac{1}{4^{d+1}(d+1)(d-1)!}(|U|-1)^{1+1/d}
\end{aligned}$$

□

The algorithm of Even et al. [2000] computes a solution within an $O(\log n \log\log n)$ factor of the cost of an optimum solution to LP (5)–(8) and hence, of the cost of q^* (Lemma 5). Therefore, the result in that paper (though not its original proof) is correct.

Now we show that we can even get an $O(\log n)$-approximation algorithm. To this end, we use a result of Fakcharoenphol, Rao and Talwar [2004], who showed how to approximate an arbitrary metric by a special kind of tree metric:

We call a tree T together with a vertex $r \in V(T)$ and a weight function $c : E(T) \to \mathbb{R}_{\geq 0}$ *2-hierarchically well separated* if there exists a constant $\gamma > 0$ such that $c(e) = \gamma \cdot 2^{-h}$, where h is the number of edges in the unique path starting in r and ending with e. This induces a metric $l' : V(T) \to \mathbb{R}_{\geq 0}$, where $l'(v,w)$ is the weight of the v-w-path in (T,c). See Figure 4.

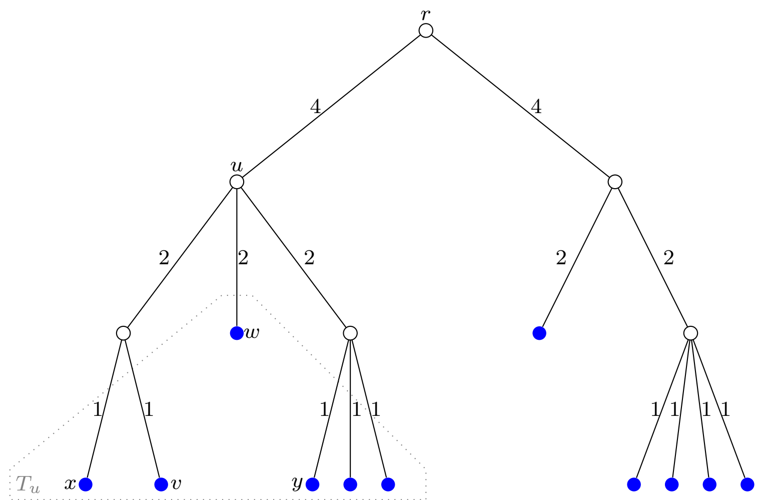


Figure 4: Illustration of a tree metric defined by a 2-hierarchically well separated tree (T, r, c). Here, $\gamma = 8$, $l'(v, w) = 5$ and $l'(x, y) = 6$. The function q defined in the proof of Theorem 7 orders the leaves from left to right.

Lemma 6 (Fakcharoenphol, Rao and Talwar [2004]) *Let G be a graph with $n \geq 2$ vertices and $l : V(G) \times V(G) \to \mathbb{R}_{\geq 0}$ a metric. Then one can compute in polynomial time a 2-hierarchically well separated tree (T, r, c) such that $V(G)$ is the set of leaves of T and the induced tree metric l' satisfies the following properties:*

(a) *$l'(v, w) \geq l(v, w)$ for all $v, w \in V(G)$; and*

(b) *$\sum_{\{v,w\} \in E(G)} l'(v, w) \leq O(\log n) \sum_{\{v,w\} \in E(G)} l(v, w)$.*

We conclude:

Theorem 7 *There is an $O(\log n)$-approximation algorithm for d-DIMAP.*

Proof: Let l be an optimum solution to the LP (5)–(8). Let (T, r, c) and l' be as defined in Lemma 6. For $u \in V(T)$ let T_u denote the set of leaves $v \in V(G)$ such that the r-v-path in T contains u. Define a bijection $q : V(G) \to \{1, \ldots, n\}$ such that for all $u \in V(T)$ the elements of T_u are numbered consecutively. Let $\{v, w\} \in E(G)$ and u be the unique vertex with $v, w \in T_u$ and $|T_u|$ maximal (see Figure 4 for an illustration). Due to the spreading constraints (6), there are $x, y \in T_u$ such that $l(x, y) \geq \frac{1}{4} \sqrt[d]{|T_u| - 1}$. Note that $l'(x, y) \leq 2 \cdot l'(v, w)$ since (T, r, c) is 2-hierarchically well separated. Therefore,

$$\sqrt[d]{|q(v) - q(w)|} \;\leq\; \sqrt[d]{|T_u| - 1} \;\leq\; 4\, l(x, y) \;\leq\; 4\, l'(x, y) \;\leq\; 8\, l'(v, w),$$

and hence,

$$\sum_{\{v,w\}\in E(G)} \sqrt[d]{|q(v)-q(w)|} \;\leq\; 8 \sum_{\{v,w\}\in E(G)} l'(v,w) \;\leq\; O(\log n) \sum_{\{v,w\}\in E(G)} l(v,w).$$

The result now follows from Lemma 1 and 5. □

We considered the unweighted version of d-DIMAP in this paper, but only to simplify the exposition. It is straightforward that all results also hold for the weighted version (where nonnegative edge weights are given and the weighted sum is minimized).

Acknowledgement

We thank Guy Even for Lemma 5 and the anonymous reviewers for careful reading and their suggestions.

References

Ambühl, C., Mastrolilli, M., and Svensson, O. [2011]: Inapproximability results for maximum edge biclique, minimum linear arrangement, and sparsest cut. SIAM Journal on Computing 40 (2011), 567–596

Arora, S., Hazan, E., and Kale, S. [2010]: $O(\sqrt{\log n})$ approximation to SPARSEST CUT in $\tilde{O}(n^2)$ time. SIAM Journal on Computing 39 (2010), 1748–1771

Arora, S., Rao, S., and Vazirani, U. [2009]: Expander flows, geometric embeddings and graph partitioning. Journal of the ACM 56 (2009), Article 5

Banerjee, P., Sur-Kolay, S., Bishnu, A., Das, S., Nandy, S.C., and Bhattacharjee, S. [2009]: FPGA placement using space-filling curves: theory meets practice. ACM Transactions on Embedded Computing Systems 9 (2009), Article 12

Charikar, M., Hajiaghayi, M.T., Karloff, H. and Rao, S. [2010]: ℓ_2^2 spreading metrics for vertex ordering problems. Algorithmica 56 (2010), 577–604

Charikar, M., Makarychev, K., and Makarychev, Y. [2007]: A divide and conquer algorithm for d-dimenisonal arrangement. Proceedings of the 18th ACM-SIAM Symposium on Discrete Algorithms (2007), 541–546

Devanur, N, Khot, S., Saket, R., and Vishnoi, N. [2006]: On the hardness of minimum linear arrangement. Manuscript, 2006

Even, G., Naor, J., Rao, S., and Schieber, B. [2000]: Divide-and-conquer approximation algorithms via spreading metrics. Journal of the ACM 47 (2000), 585–616

Fakcharoenphol, J., Rao, S., and Talwar, K. [2004]: A tight bound on approximating arbitrary metrics by tree metrics. Journal of Computer and System Sciences 69 (2004), 485–497

Feige, U., and Lee, J.R. [2007]: An improved approximation ratio for the minimum linear arrangement problem. Information Processing Letters 101 (2007), 26–29

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

Hansen, M.D. [1989]: Approximation algorithms for geometric embeddings in the plane with applications to parallel processing problems. Proceedings of the 30th Annual IEEE Symposium on Foundations of Computer Science (1989), 604–609

Hilbert, D. [1891]: Über die stetige Abbildung einer Linie auf ein Flächenstück. Mathematische Annalen 38 (1891), 459–460

Leighton, T., and Rao, S. [1999]: Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. Journal of the ACM 46 (1999), 787–832

Rao, S., and Richa, A.W. [2004]: New approximation techniques for some linear ordering problems. SIAM Journal on Computing 34 (2004), 388–404

Sagan, H. [1994]: Space-Filling Curves. Springer, New York 1994