

# Sphere and Dot Product Representations of Graphs

Ross J. Kang · Tobias Müller

Received: 6 April 2011 / Revised: 18 November 2011 / Accepted: 4 January 2012 /

Published online: 25 January 2012

© The Author(s) 2012. This article is published with open access at Springerlink.com

**Abstract** A graph  $G$  is a  $k$ -sphere graph if there are  $k$ -dimensional real vectors  $v_1, \dots, v_n$  such that  $ij \in E(G)$  if and only if the distance between  $v_i$  and  $v_j$  is at most 1. A graph  $G$  is a  $k$ -dot product graph if there are  $k$ -dimensional real vectors  $v_1, \dots, v_n$  such that  $ij \in E(G)$  if and only if the dot product of  $v_i$  and  $v_j$  is at least 1.

By relating these two geometric graph constructions to oriented  $k$ -hyperplane arrangements, we prove that the problems of deciding, given a graph  $G$ , whether  $G$  is a  $k$ -sphere or a  $k$ -dot product graph are NP-hard for all  $k > 1$ . In the former case, this proves a conjecture of Breu and Kirkpatrick (Comput. Geom. 9:3–24, 1998). In the latter, this answers a question of Fiduccia et al. (Discrete Math. 181:113–138, 1998).

Furthermore, motivated by the question of whether these two recognition problems are in NP, as well as by the implicit graph conjecture, we demonstrate that, for all  $k > 1$ , there exist  $k$ -sphere graphs and  $k$ -dot product graphs such that each representation in  $k$ -dimensional real vectors needs at least an exponential number of bits to be stored in the memory of a computer. On the other hand, we show that exponentially many bits are always enough. This resolves a question of Spinrad (Efficient Graph Representations, 2003).

**Keywords** Geometric representations of graphs · Sphericity · Dot product dimension

---

R.J. Kang  
Durham University, Durham, UK  
e-mail: [ross.kang@gmail.com](mailto:ross.kang@gmail.com)

T. Müller (✉)  
Centrum Wiskunde & Informatica, Amsterdam, Netherlands  
e-mail: [tobias@cwi.nl](mailto:tobias@cwi.nl)

## 1 Introduction and Statement of Results

A graph  $G$  is a  $k$ -sphere graph if there are vectors  $v_1, \dots, v_n \in \mathbb{R}^k$  such that  $\|v_i - v_j\| \leq 1$  if and only if  $ij \in E(G)$ . Alternatively, a  $k$ -sphere graph can be seen as the intersection graph of equal radius spheres in  $k$ -dimensional space; that is, we can represent each vertex  $i$  by a sphere  $S_i \subseteq \mathbb{R}^k$  of radius 1 in such a way that  $ij \in E(G)$  if and only if  $S_i \cap S_j \neq \emptyset$ . The *sphericity*  $\text{sph}(G)$  of a graph  $G$  is the least  $k$  such that  $G$  is a  $k$ -sphere graph. (Every graph has finite sphericity [10].) Sphericity was first introduced by Havel [7] in his Ph.D. thesis to study the geometry of molecules. Graphs of sphericity one are also called *unit interval graphs*, and graphs of sphericity at most two are called *unit disk graphs*. Unit disk graphs have received considerable attention over the last two decades, in part because of their relevance to applications in wireless networks.

A *k-dot product graph* is a graph  $G$  that can be represented by vectors  $v_1, \dots, v_n \in \mathbb{R}^k$  such that  $ij \in E(G)$  if and only if  $v_i^T v_j \geq 1$ . The *dot product dimension*  $\text{dpd}(G)$  of a graph  $G$  is the least  $k$  such that  $G$  is a  $k$ -dot product graph. (Every graph has finite dot product dimension [6].) Dot product representations of graphs were first introduced by Reiterman et al. [16–18] and (independently, it seems) by Fiduccia et al. [6]. Here, it should be mentioned that Reiterman et al. used an arbitrary threshold  $t \in \mathbb{R}$  in their definition (i.e.  $ij \in E(G)$  if and only if  $v_i^T v_j \geq t$ ). It is however straightforward to adapt our proofs to this definition and we fix  $t = 1$  for convenience. Graphs of dot product one are precisely the class of graphs that have at most two nontrivial connected components each of which is a threshold graph, cf. [6]. The class of threshold graphs has been studied quite extensively, cf. [11].

In this paper, we will consider two related types of questions. First, we study the algorithmic decision problems of recognizing  $k$ -sphere graphs and  $k$ -dot product graphs. Second, we analyse the number of bits needed to store the coordinates of the vectors  $v_1, \dots, v_n \in \mathbb{R}^k$  in a  $k$ -sphere or a  $k$ -dot product representation.

By  $k$ -SPHERICITY, we mean the problem of deciding, given a graph  $G$  as input, whether  $G$  has sphericity at most  $k$ . Put in different words,  $k$ -SPHERICITY is the recognition problem for  $k$ -sphere graphs. It is known that the  $k = 1$  version of this problem can be solved in linear time [9]. On the other hand, Breu and Kirkpatrick [4] proved that 2-SPHERICITY is NP-hard. Breu and Kirkpatrick used a rather complicated reduction from SATISFIABILITY, which crucially relies on a special embedding of instances of SATISFIABILITY on a 2-dimensional grid. They briefly sketched how a similar embedding on a 3-dimensional grid can also be obtained, showing NP-hardness of 3-sphericity, but were unable to extend this further and they conjectured that  $k$ -SPHERICITY is NP-hard for all remaining  $k$ . We prove their conjecture, by providing a simpler NP-hardness proof for 2-SPHERICITY that also extends to all higher dimensions.

**Theorem 1** *k-SPHERICITY is NP-hard for all  $k > 1$ .*

The proof of Theorem 1 builds upon techniques that appeared earlier in [13].

The decision problem  $k$ -DOT PRODUCT asks whether a given input graph has dot product dimension at most  $k$ . For  $k = 1$  it follows from a result of Fiduccia et al. (Theorem 20 in [6]) and a result of Chvátal and Hammer that this is solvable in linear time.

Fiduccia et al. [6] asked whether this recognition problem is NP-hard for other values of  $k$ , and we affirm this.

**Theorem 2**  *$k$ -DOT PRODUCT is NP-hard for all  $k > 1$ .*

Rather than a reduction from (some variant of) SATISFIABILITY, the proofs of Theorems 1 and 2 follow a reduction from the so-called simple stretchability problem, to be defined in the next section.

Having determined that the decision problems of recognizing  $k$ -sphere graphs and  $k$ -dot product graphs are NP-hard, one might wonder whether these decision problems are members of the decision class NP. For this, we need a “polynomial certificate”. That is, for each graph that is a  $k$ -sphere (resp.  $k$ -dot product) graph, there is a proof of this fact that can be checked by an algorithm in polynomial time. An obvious candidate for such a certificate is the set of vectors  $v_1, \dots, v_n$  of a  $k$ -sphere (resp.  $k$ -dot product) representation. For this to be a good certificate, we would however need to guarantee that  $v_1, \dots, v_n$  can be stored using polynomially many bits.

Spinrad [22] also asked how many bits are needed to store the coordinates of a  $k$ -dot product graph, motivated by the so-called *implicit graph conjecture* of Kannan et al. [8]. This conjecture asserts that, for every hereditary graph class  $\mathcal{C}$  (i.e.  $\mathcal{C}$  is closed under taking induced subgraphs) such that the number of graphs on  $n$  vertices in  $\mathcal{C}$  is  $2^{O(n \ln n)}$ , there exists a “labelling scheme” that encodes graphs in the class by assigning bit strings of length  $O(\ln n)$  to the vertices in such a way that the adjacency of two vertices  $u, v$  can be tested by examining only the labels of  $u$  and  $v$ . It is easily shown that the number of  $k$ -sphere graphs (resp.  $k$ -dot product graphs) on  $n$  vertices is  $2^{O(n \ln n)}$  by Warren’s Theorem, cf. [22]. A natural candidate for the labelling scheme would be to use the vectors  $v_1, \dots, v_n$  of some  $k$ -sphere (resp.  $k$ -dot product) representation as the labels. For this to work, we would however need to guarantee that each of  $v_1, \dots, v_n$  can be stored using only  $O(\ln n)$  bits.

The bit size of a natural number  $n \in \mathbb{N}$  is the number of symbols needed to write it in binary notation. It satisfies

$$\text{size}(n) := \lceil \log_2(n + 1) \rceil. \quad (1)$$

The bit size of an integer  $z \in \mathbb{Z}$  is  $\text{size}(z) = 1 + \text{size}(|z|)$ , as we need an extra bit for the sign. In this paper, we shall use the convention that a rational number is stored as a pair of integers (the denominator and numerator) that are relatively prime and these integers are stored in the binary number format. Following Basu et al. [2], we define the size of a rational number  $q = a/b \in \mathbb{Q}$  with  $a, b \in \mathbb{Z}$  relatively prime as

$$\text{size}(q) := \text{size}(a) + \text{size}(b). \quad (2)$$

For a vector  $v \in \mathbb{Q}^k$ , we set  $\text{size}(v) := \sum_{i=1}^k \text{size}(v_i)$ .

We should remark that several other notions of the size of a rational number exist in the literature.

We are interested in the following quantities for  $G$  a  $k$ -sphere graph (resp. a  $k$ -dot product graph):

$$f_k(G) := \min_{\substack{v:V(G)\rightarrow\mathbb{Q}^k, \\ v \text{ is a } k\text{-sphere} \\ \text{representation of } G}} \sum_{t\in V(G)} \text{size}(v(t)),$$

$$g_k(G) := \min_{\substack{v:V(G)\rightarrow\mathbb{Q}^k, \\ v \text{ is a } k\text{-dot product} \\ \text{representation of } G}} \sum_{t\in V(G)} \text{size}(v(t)).$$

It may not be immediately clear to the reader that  $f_k(G), g_k(G)$  are well-defined, since this requires that every  $k$ -sphere or  $k$ -dot product graph has a representation with all coordinates of all vectors rational. That such a rational representation always exists follows from (the proofs of) the upper bounds in Theorems 3 and 4 below. Let us set

$$f_k(n) := \max_{\substack{|V(G)|=n, \\ G \text{ is a } k\text{-sphere graph}}} f_k(G),$$

and

$$g_k(n) := \max_{\substack{|V(G)|=n, \\ G \text{ is a } k\text{-dot product graph}}} g_k(G).$$

We will show that there are  $k$ -sphere graphs that require exponentially many bits, but exponentially many bits are always enough.

**Theorem 3** *For every  $k > 1$ , there exist  $c_1, c_2 > 0$  such that*

$$2^{c_1 n} \leq f_k(n) \leq 2^{c_2 n}.$$

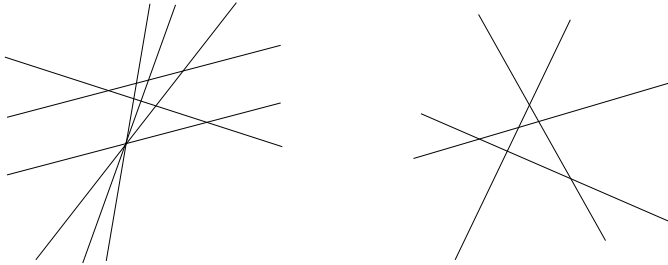
Here is should be mentioned that the special case when  $k = 2$  of Theorem 3 was previously proved by McDiarmid and the second author [12], and a slightly weaker lower bound for the case  $k = 2$  of  $f_2(n) \geq 2^{\Omega(\sqrt{n})}$  was previously proved in [13], the conference version of [12].

We also show that the analogous result holds for  $k$ -dot product graphs.

**Theorem 4** *For every  $k > 1$ , there exist  $c_1, c_2 > 0$  such that*

$$2^{c_1 n} \leq g_k(n) \leq 2^{c_2 n}.$$

Theorems 3 and 4 show that the vectors of a  $k$ -sphere representation or of a  $k$ -dot product representation cannot be used as a labelling scheme to prove a special case of the implicit graph conjecture. They also cannot serve as certificates to prove membership in NP of the recognition problems.



**Fig. 1** A non-simple and a simple line arrangement

## 2 Preliminaries

A *line arrangement* is a system  $\mathcal{L} := (\ell_1, \dots, \ell_n)$  of lines in the plane. If every two lines intersect and no point lies on more than two lines, then we call  $\mathcal{L}$  a *simple line arrangement* (or, also, a *uniform line arrangement*). See Fig. 1 for a depiction of a non-simple and a simple line arrangement.

A line  $\ell$  divides  $\mathbb{R}^2 \setminus \ell$  into two pieces. In an *orientation* of  $\ell$ , we distinguish between these two pieces by (arbitrarily) calling one of them,  $\ell^-$ , the “negative side” and the other,  $\ell^+$ , the “positive side”. An *oriented line arrangement* is a line arrangement for which all the lines have been given an orientation.

For convenience, from now on we shall only work with oriented line arrangements. The *sign vector* of a point  $p \in \mathbb{R}^2$  with respect to an oriented line arrangement  $\mathcal{L} = (\ell_1, \dots, \ell_n)$  is a vector  $\sigma(p) \in \{-, 0, +\}^n$  defined as follows:

$$(\sigma(p))_i = \begin{cases} - & \text{if } p \in \ell_i^-, \\ 0 & \text{if } p \in \ell_i, \\ + & \text{if } p \in \ell_i^+. \end{cases}$$

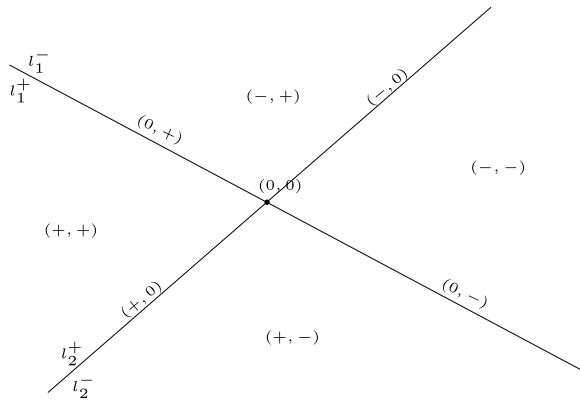
The *combinatorial description*  $\mathcal{D}(\mathcal{L})$  of  $\mathcal{L}$  is the set of all sign vectors

$$\mathcal{D}(\mathcal{L}) := \{\sigma(p) : p \in \mathbb{R}^2\}.$$

The set  $\mathcal{D}(\mathcal{L})$  is almost the same thing as the set of covectors of an oriented matroid and in fact it determines the oriented matroid associated with  $\mathcal{L}$ . (Consult the monograph on oriented matroids by Björner et al. [3] for more details.) It should be mentioned that various alternative combinatorial descriptions of a line arrangement are available, such as local sequences, allowable sequences and wiring diagrams (cf. [5]).

Each connected component of  $\mathbb{R}^2 \setminus (\ell_1 \cup \dots \cup \ell_n)$  is called a *cell* or a *region*. All points in the same cell have the same sign vector, which does not have 0 as a coordinate. A sign vector with two or more zeroes corresponds to an intersection point between two or more lines, and a sign vector with exactly one zero corresponds to a line segment.

If  $\mathcal{D}(\mathcal{L}) = \mathcal{D}(\mathcal{L}')$ , then we say that  $\mathcal{L}$  and  $\mathcal{L}'$  are *isomorphic*. We will use a number of relatively straightforward observations about (oriented) line arrangements in our proofs below. For completeness, we include their proofs in the [Appendix](#).



**Lemma 5** Let  $\mathcal{L}$  be a line arrangement. Then there exists  $\sigma \in \mathcal{D}(\mathcal{L}) \cap \{-, +\}^n$  such that  $-\sigma \in \mathcal{D}(\mathcal{L}) \cap \{-, +\}^n$ .

**Lemma 6** A line arrangement on  $n$  lines has at most  $1 + \binom{n+1}{2}$  cells, and it is simple if and only if it has exactly  $1 + \binom{n+1}{2}$  cells.

**Lemma 7** Let  $\mathcal{L}$  be a simple oriented line arrangement, and let  $\mathcal{L}'$  be an oriented line arrangement with the same number of lines. Then  $\mathcal{L}$  and  $\mathcal{L}'$  are isomorphic if and only if  $\mathcal{D}(\mathcal{L}) \cap \{-, +\}^n = \mathcal{D}(\mathcal{L}') \cap \{-, +\}^n$ .

Similarly to a line arrangement in the plane, a  $k$ -hyperplane arrangement  $\mathcal{H} = (h_1, \dots, h_n)$  is a system of (affine) hyperplanes in  $\mathbb{R}^k$ . We define a cell, an orientation, a sign vector and the combinatorial description in precisely the same way as for lines.

Recall that an affine transformation  $T : \mathbb{R}^k \rightarrow \mathbb{R}^k$  is a linear transformation followed by a translation. Put differently, it is a map of the form  $x \mapsto Ax + b$ , where  $A$  is a nonsingular  $k \times k$  matrix, and  $b \in \mathbb{R}^k$  is a (constant) vector. We will need the following observation in the sequel. The straightforward proof can again be found in the [Appendix](#).

**Lemma 8** Let  $\mathcal{H}$  be an oriented  $k$ -hyperplane arrangement, and let  $T : \mathbb{R}^k \rightarrow \mathbb{R}^k$  be an affine transformation. Define another oriented  $k$ -hyperplane arrangement by setting

$$h'_i := T[h_i], \quad (h'_i)^- := T[h_i^-], \quad (h'_i)^+ := T[h_i^+].$$

Then  $\mathcal{H}$  and  $\mathcal{H}'$  are isomorphic.

STRETCHABILITY is the problem of deciding, given a set  $S \subseteq \{-, 0, +\}^n$ , whether  $S$  is the combinatorial description of some oriented line arrangement of  $n$  lines. We can also define an orientation and a combinatorial description for an arrangement  $\mathcal{C} = (c_1, \dots, c_n)$  of continuous curves (that satisfy some regularity conditions) in the same way. The name “stretchability” now comes from the idea that

we may ask, for an oriented system of curves in the plane, whether there is a line arrangement with the same combinatorial description. In this case, we may imagine the curves “stretching” into lines.

Similarly, SIMPLE STRETCHABILITY is the problem of deciding, given a set  $S \subseteq \{-, 0, +\}^n$ , whether  $S$  is the combinatorial description of some oriented simple arrangement of  $n$  lines.

**Theorem 9** (Shor [21]) *SIMPLE STRETCHABILITY is NP-hard.*

This theorem is also a straightforward corollary of a deep topological result by Mněv [14]. Shor’s proof is more direct. It reduces a SAT-variant to stretchability using Pappus’ and Desargues’ theorems, cf. Chap. 8 of [3].

### 3 The Proofs

For the NP-hardness proofs, it will be useful to define an intermediate decision problem. For  $k > 1$  and  $S \subseteq \{-, +\}^n$  with  $(-, \dots, -), (+, \dots, +) \in S$ , we say that  $S$  is  $k$ -realizable if there exists an oriented  $k$ -hyperplane arrangement  $\mathcal{H}$  with  $S \subseteq \mathcal{D}(\mathcal{H})$ . The condition that  $(-, \dots, -), (+, \dots, +) \in S$  is useful for technical reasons later on in the proofs of Theorems 1 and 2. For  $k > 1$ , we let  $k$ -REALIZABILITY denote the algorithmic problem of deciding, given a set  $S \subseteq \{-, +\}^n$  as input, whether  $S$  is  $k$ -realizable.

**Theorem 10**  $k$ -REALIZABILITY is NP-hard for all  $k > 1$ .

The proof of Theorem 10 relies on the following observation:

**Lemma 11** *Let  $S \subseteq \{-, +\}^n$  with  $(-, \dots, -), (+, \dots, +) \in S$  and  $k > 1$  be arbitrary and set  $S' := S \times \{-, +\}$ . Then  $S$  is  $k$ -realizable if and only if  $S'$  is  $(k + 1)$ -realizable.*

*Proof* Let  $k, S, S'$  be as in the statement of the lemma. Let us first suppose that there exists an oriented hyperplane arrangement  $\mathcal{H} = (h_1, \dots, h_n)$  in  $\mathbb{R}^k$  such that  $S \subseteq \mathcal{D}(\mathcal{H})$ . Define a new hyperplane arrangement  $\mathcal{H}'$  by

$$h'_i = h_i \times \mathbb{R}, \quad (h'_i)^- = h_i^- \times \mathbb{R}, \quad (h'_i)^+ = h_i^+ \times \mathbb{R},$$

for  $i = 1, \dots, n$ , and

$$\begin{aligned} h'_{n+1} &= \mathbb{R}^k \times \{0\}, & (h'_{n+1})^- &= \mathbb{R}^k \times \{x : x < 0\}, \\ (h'_{n+1})^+ &= \mathbb{R}^k \times \{x : x > 0\}. \end{aligned} \tag{3}$$

For each  $s \in S$ , there exists  $p \in \mathbb{R}^k$  such that  $\sigma_{\mathcal{H}}(p) = s$ . If we set  $p_1 = (p, -1)$   $p_2 = (p, 1)$ , then clearly  $\sigma_{\mathcal{H}'}(p_1) = (s, -)$  and  $\sigma_{\mathcal{H}'}(p_2) = (s, +)$ . Thus,  $S' \subseteq \mathcal{D}(\mathcal{H}')$ , as required.

Now, suppose that  $S' \subseteq \mathcal{D}(\mathcal{H}')$  for some  $(k + 1)$ -hyperplane arrangement  $\mathcal{H}'$ . By Lemma 8, we can also assume without loss of generality that (3) holds. Let us define a  $k$ -hyperplane arrangement  $\mathcal{H} = (h_1, \dots, h_n)$  by identifying  $\mathbb{R}^k$  with  $\mathbb{R}^k \times \{0\}$  in the obvious way and setting

$$h_i := h'_i \cap \mathbb{R}^k, \quad h_i^- := (h'_i)^- \cap \mathbb{R}^k, \quad h_i^+ := (h'_i)^+ \cap \mathbb{R}^k. \tag{4}$$

It is not immediately clear that each  $h_i$  is a  $k$ -hyperplane, since we might have  $h_i = \emptyset$ , in case  $h'_i$  and  $h'_{n+1}$  happen to be parallel, or  $h_i = \mathbb{R}^k$ , in case  $h'_i$  and  $h'_{n+1}$  coincide. However, if none of these two cases holds, then  $\dim_{\mathbb{R}^{k+1}}(h'_i \cap h'_{n+1}) = k - 1$  so that  $h_i$  is a well-defined hyperplane in  $\mathbb{R}^k$ . We shall see soon that this is in fact the case for all  $i = 1, \dots, n$ .

Let  $s \in S$  be arbitrary, and let us write  $s_1 = (s, -), s_2 = (s, +)$ . Pick  $p_1, p_2 \in \mathbb{R}^{k+1}$  such that  $\sigma_{\mathcal{H}'}(p_i) = s_i$ . Then  $p_1 \in (h'_{n+1})^-$  and  $p_2 \in (h'_{n+1})^+$ , so that the segment  $[p_1, p_2]$  must intersect  $h'_{n+1} = \mathbb{R}^k \times \{0\}$  at some point  $p' = (p, 0)$ . It is clear that  $\sigma_{\mathcal{H}'}(p') = (s, 0)$  and hence  $\sigma_{\mathcal{H}}(p) = s$ .

It remains for us to show that the  $h_i$  are well-defined hyperplanes in  $\mathbb{R}^k$ . To this end, recall that  $(-, \dots, -), (+, \dots, +) \in S$ . Let  $p, q \in \mathbb{R}^k$  be points with  $\sigma_{\mathcal{H}}(p) = (-, \dots, -), \sigma_{\mathcal{H}}(q) = (+, \dots, +)$  and let  $\ell$  be the line between  $p, q$ . Then  $\ell \subseteq \mathbb{R}^k = h'_{n+1}$  and  $\ell$  hits all of  $h_i^-, h_i, h_i^+$  for each  $i = 1, \dots, n$ . Hence, we cannot have  $h_i = \emptyset$  or  $h_i = \mathbb{R}^k$  so that  $h_i$  is a well-defined  $k$ -hyperplane. □

*Proof of Theorem 10* We shall start by considering the case when  $k = 2$ . For  $k = 2$ , we shall use a reduction from SIMPLE STRETCHABILITY. Let  $D \subseteq \{-, 0, +\}^n$  be some set of sign vectors. We need to construct, in polynomial time, a set  $S \subseteq \{-, +\}^n$  such that  $S$  is 2-realizable if and only if the answer to SIMPLE STRETCHABILITY with input  $D$  is “yes”.

We can assume that  $|D \cap \{-, +\}^n| = 1 + \binom{n+1}{2}$ . This is because, if  $|D \cap \{-, +\}^n| \neq 1 + \binom{n+1}{2}$ , then the answer to the SIMPLE STRETCHABILITY-instance is “no” by Lemma 6. We can certainly determine if  $|D \cap \{-, +\}^n| \neq 1 + \binom{n+1}{2}$  in polynomial time, and if so we can simply output a set  $S \subseteq \{-, +\}^n$  which we know is not 2-realizable. Similarly, we can assume that there is some  $s \in D \cap \{-, +\}^n$  such that  $-s \in D$ . (We can check for the existence of such a pair in polynomial time, and non-existence of such a pair would imply that the answer to the SIMPLE STRETCHABILITY-instance is “no” by Lemma 5.)

Now, let us pick such a pair  $s, -s \in D \cap \{-, +\}^n$  and let  $f : \{-, 0, +\}^n \rightarrow \{-, 0, +\}^n$  denote the map that changes  $+$  to  $-$  and changes  $-$  to  $+$  in all coordinates  $i$  where  $s_i = +$  and leaves all other coordinates unchanged. Let us set  $D' := \{f(t) : t \in D\}$ . Note that  $f(s) = (-, \dots, -)$  and  $f(-s) = (+, \dots, +)$ . If  $D = \mathcal{D}(\mathcal{L})$  for some line arrangement, then  $D'$  corresponds to a “reorientation” (i.e. switching the roles of  $h_i^-$  and  $h_i^+$  for some indices  $i$ ). Thus,  $D = \mathcal{D}(\mathcal{L})$  for some simple oriented line arrangement  $\mathcal{L}$  if and only if  $D' = \mathcal{D}(\mathcal{L}')$  for some simple oriented line arrangement  $\mathcal{L}'$ . ( $D'$  is constructed from  $D$  to ensure the technical promise that  $(-, \dots, -), (+, \dots, +) \in S$  given earlier.)

We now set  $S = D' \cap \{-, +\}^n$ . We claim that  $S$  is 2-realizable if and only if  $D = \mathcal{D}(\mathcal{L})$  for some simple oriented line arrangement  $\mathcal{L}$ .



To see this, first note that if  $D = \mathcal{D}(\mathcal{L})$  for some simple  $\mathcal{L}$ , then  $S$  is 2-realizable by construction. Let us thus assume that  $S$  is 2-realizable. By Lemma 6, there is a simple  $\mathcal{L}$  such that  $S \subseteq \mathcal{D}(\mathcal{L})$ . Lemma 7 now implies that  $\mathcal{D}(\mathcal{L}) = D'$ . But then we also have that  $D$  is the combinatorial description of some simple oriented line arrangement. Thus,  $S$  is 2-realizable if and only if  $D$  is the combinatorial description of a simple oriented line arrangement, as claimed. Since all the steps above to produce  $S$  from  $D$  can clearly be implemented in polynomial time, this proves the theorem for  $k = 2$ .

Let us now consider the case  $k > 2$ , and suppose that we have already shown that  $(k - 1)$ -REALIZABILITY is NP-hard. Since it is clear that we can construct  $S' = S \times \{-, +\}$  in polynomial time from  $S$ , it is a direct consequence of Lemma 11 that  $k$ -REALIZABILITY is also NP-hard. □

To prove Theorems 1 and 2, we will also need the following technical lemma.

**Lemma 12** *Let  $\varepsilon > 0$ ,  $K > 0$  be arbitrary, and let  $\mathcal{H}$  be an oriented hyperplane arrangement with  $(-, \dots, -), (+, \dots, +) \in \mathcal{D}(\mathcal{H})$ . Then there exists  $\mathcal{H}'$  with the following properties.*

- (i)  $\mathcal{D}(\mathcal{H}') = \mathcal{D}(\mathcal{H})$ .
- (ii) The origin  $\underline{0}$  belongs to the  $(-, \dots, -)$ -cell of  $\mathcal{H}'$ .
- (iii) For each oriented hyperplane  $h'_i$  of  $\mathcal{H}'$ , we can write

$$h'_i = \{z : v_i^T z = c_i\}, \quad h_i^+ = \{z : v_i^T z < c_i\}, \quad h_i^- = \{z : v_i^T z > c_i\},$$

with  $\|v_i - e_1\| < \varepsilon$  (where  $e_1 = (1, 0, \dots, 0)$ ).

- (iv) Every cell of  $\mathcal{H}'$  intersects the ball  $B((K, 0, \dots, 0)^T, \varepsilon)$ .

*Proof* The proof is by means of repeated applications of Lemma 8. By using a suitable scaling  $z \mapsto \lambda z$  if needed, we can ensure that every cell of  $\mathcal{H}$  intersects the ball  $B(\underline{0}, \varepsilon/2)$  of radius  $\varepsilon/2$  around the origin. In particular, there are points  $p, q \in B(\underline{0}, \varepsilon/2)$  such that  $\sigma(q) = (-, \dots, -)$ ,  $\sigma(p) = (+, \dots, +)$ . Let  $\ell$  denote the line through  $p$  and  $q$ . Notice that all the hyperplanes  $h_1, \dots, h_n$  must intersect  $\ell$  in the open segment  $(p, q)$ .

Let  $r$  be the connected component of  $\ell \setminus (p, q)$  that contains  $q$ . Thus,  $r$  is a “ray” emanating from  $q$ . All points of  $r$  must have sign vector  $(-, \dots, -)$  (since none of the hyperplanes  $h_1, \dots, h_n$  hits  $r$ ). By applying a suitable translation followed by a rotation, we can map  $q$  to the origin  $\underline{0}$  and  $r$  to the nonnegative  $e_1$ -axis  $\{(x, 0, \dots, 0) : x \geq 0\}$ . Notice that now, since we have translated over a distance of  $\|q\| < \varepsilon/2$ , every cell intersects  $B(\underline{0}, \varepsilon)$ . Also, notice that the positive  $e_1$ -axis intersects each of the hyperplanes  $h_1, \dots, h_n$ .

Next, we will rescale only the  $e_1$ -axis to make all hyperplanes have normals that are close to the  $e_1$ -vector. To see that this can be done, let  $v_1, \dots, v_n$  be vectors and  $c_1, \dots, c_n$  be constants such that

$$h_i = \{z : v_i^T z = c_i\}, \quad h_i^- = \{z : v_i^T z < c_i\}, \quad h_i^+ = \{z : v_i^T z > c_i\}.$$

Note that, since  $\underline{0} \in h_i^-$ , we have  $c_i > 0$  for all  $i$ . Similarly, since  $(x, 0, \dots, 0) \in h_i^-$  for all  $x \leq 0$ , we must then also have  $(v_i)_1 \geq 0$ . Also, note that, since  $e_1$  is not parallel

to  $h_i$  (recall that  $h_i$  intersects the positive  $e_1$ -axis), we must have  $(v_i)_1 \neq 0$  for all  $i$ . Now, let  $T$  denote the linear transformation  $T : (z_1, \dots, z_n) \mapsto (\lambda z_1, z_2, \dots, z_n)$ , for some  $\lambda > 0$ . Then we obtain

$$h'_i = \{T(z) : v_i^T z = c_i\}$$

$$= \{z : \lambda^{-1}(v_i)_1 z_1 + (v_i)_2 z_2 + \dots + (v_i)_n z_n = c_i\}.$$

Let us write

$$w_i := (\lambda^{-1}(v_i)_1, (v_i)_2, \dots, (v_i)_n)^T, \quad u_i := w_i / \|w_i\|, \quad d_i := c_i / \|w_i\|.$$

Clearly,  $h'_i = \{z : u_i^T z = d_i\}$ . Now, note that  $u_i \rightarrow e_1$  as  $\lambda \searrow 0$ . Thus, if we take  $\lambda$  sufficiently small, then we obtain  $\|u_i - e_1\| < \varepsilon$  for all  $i$ . Notice that, since we have merely “scaled down” the  $e_1$ -coordinate of all points in  $\mathbb{R}^k$ , it still holds that every cell intersects  $B(\underline{0}, \varepsilon)$  and the nonpositive  $e_1$ -axis belongs to the all-minus cell.

Finally, if we apply the translation  $z \mapsto z + K e_1$ , then it is clear that the normals of the hyperplanes do not change and the ball  $B(\underline{0}, \varepsilon)$  is mapped into  $B((K, 0, \dots, 0)^T, \varepsilon)$ . Hence, every cell now hits  $B((K, 0, \dots, 0)^T, \varepsilon)$ . Moreover, the origin still lies in the cell corresponding to  $(-, \dots, -)$ , since the nonpositive  $e_1$ -axis was mapped onto  $\{(x, 0, \dots, 0) : x \leq K\}$ . □

*Proof of Theorem 1* We will use a reduction from  $k$ -REALIZABILITY. Thus, given a set  $S \subseteq \{-, +\}^n$  with  $(-, \dots, -), (+, \dots, +) \in S$ , we will construct (in polynomial time) a graph  $G$  such that  $G$  is a  $k$ -sphere graph if and only if  $S \subseteq \mathcal{D}(\mathcal{H})$  for some oriented  $k$ -hyperplane arrangement  $\mathcal{H}$ .

Given  $S \subseteq \{-, +\}^n$ , let  $G = G(S)$  be a graph on  $2n + |S|$  vertices which is defined as follows. Let  $V(G) = A \cup B \cup C$ , where  $A = \{a_1, \dots, a_n\}$ ,  $B = \{b_1, \dots, b_n\}$  and  $C = \{c_\sigma : \sigma \in S\}$ . The edge set of  $G$  is as follows:

- $a_i a_j, b_i b_j \in E(G)$  for all  $1 \leq i < j \leq n$ ;
- $a_i b_j \notin E(G)$  for all  $1 \leq i, j \leq n$ ;
- $c_\sigma c_\tau \in E(G)$  for all  $\sigma \neq \tau \in S$ ;
- $a_i c_\sigma \in E(G)$  if and only if  $\sigma_i = -$ ;
- $b_i c_\sigma \in E(G)$  if and only if  $\sigma_i = +$ .

It should be clear that  $G(S)$  can be constructed in polynomial time. It remains to show that  $G$  is a  $k$ -sphere graph if and only if there exists a  $k$ -hyperplane arrangement  $\mathcal{H}$  such that  $S \subseteq \mathcal{D}(\mathcal{H})$ .

First, suppose that  $G$  can be realized as a  $k$ -sphere graph. That is, there is a function  $v : V(G) \rightarrow \mathbb{R}^k$  such that  $\|v(s) - v(t)\| \leq 1$  if and only if  $st \in E(G)$ . Let us define a hyperplane arrangement  $\mathcal{H}$  as follows:

$$h_i := \{z : \|z - v(a_i)\| = \|z - v(b_i)\|\},$$

$$h_i^- := \{z : \|z - v(a_i)\| < \|z - v(b_i)\|\},$$

$$h_i^+ := \{z : \|z - v(a_i)\| > \|z - v(b_i)\|\}.$$
(5)

Notice that, for each  $\sigma \in S$ , if  $\sigma_i = -$ , then  $v(c_\sigma) \in B(v(a_i), 1) \setminus B(v(b_i), 1)$ . So  $v(c_\sigma)$  is closer to  $v(a_i)$  than to  $v(b_i)$  and hence it lies in  $h_i^-$ . Completely analogously,  $v(c_\sigma) \in h_i^+$  if  $\sigma_i = +$ . This shows that  $S \subseteq \mathcal{D}(\mathcal{H})$ .

Next, suppose that there exists  $\mathcal{H}$  with  $S \subseteq \mathcal{D}(\mathcal{H})$ . We can assume that  $\mathcal{H}$  satisfies the conclusions of Lemma 12 with  $K := 1000$  and  $\varepsilon := 1/1000$ . Let  $v_1, \dots, v_n$  and  $c_1, \dots, c_n$  be as provided by part (iii) of Lemma 12. For notational convenience, let us write  $B := B((K, 0, \dots, 0)^T, \varepsilon)$  and  $u_i := v_i / \|v_i\|$ . It follows from  $\|v_i - e_1\| \leq \varepsilon$  that

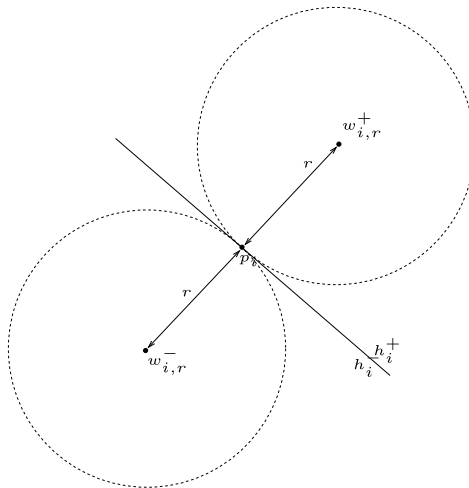
$$\|u_i - e_1\| \leq 3\varepsilon \quad \text{for all } i. \tag{6}$$

For each  $\sigma \in S$ , let  $w(c_\sigma)$  be a point of  $B$  with sign vector  $\sigma$  with respect to  $\mathcal{H}$ . For each  $i = 1, \dots, n$ , let  $p_i$  be a point of  $h_i \cap B$ . (To see that such a point exists, note that  $B$  contains a point of  $h_i^-$  and a point of  $h_i^+$ . The segment between these two points must hit  $h_i$ .) For  $r > 0$ , let us set

$$w_{i,r}^- := p_i - ru_i, \quad w_{i,r}^+ := p_i + ru_i,$$

and

$$B_{i,r}^- := B(w_{i,r}^-, r), \quad B_{i,r}^+ := B(w_{i,r}^+, r).$$

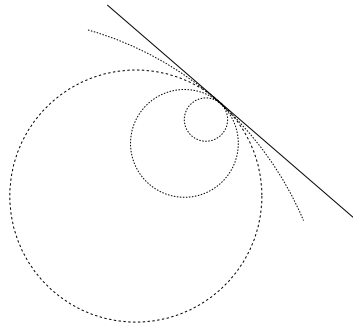


Notice that  $B_{i,r}^- \subseteq h_i^-$  and  $B_{i,r}^+ \subseteq h_i^+$ . Moreover, if  $r' > r$ , then

$$B_{i,r}^- \subseteq B_{i,r'}^- \quad \text{and} \quad B_{i,r}^+ \subseteq B_{i,r'}^+.$$

Also, note that

$$\bigcup_{r>0} B_{i,r}^- = h_i^-, \quad \bigcup_{r>0} B_{i,r}^+ = h_i^+.$$



Hence, for all large enough  $r$ , it must hold that  $w(c_\sigma) \in B_{i,r}^-$  for all  $i$  and all  $\sigma$  with  $\sigma_i = -$ , and  $w(c_\sigma) \in B_{i,r}^+$  for all  $i$  and all  $\sigma$  with  $\sigma_i = +$ . Let us choose such an  $r$ , and set  $w(a_i) := w_{i,r}^-, w(b_i) := w_{i,r}^+$  for all  $i = 1, \dots, n$ .

If we now set  $v(t) = w(t)/r$  for all  $t \in V(G)$ , then we see that  $\|v(a_i) - v(c_\sigma)\| \leq 1$  if and only if  $\sigma_i = -$ , and  $\|v(b_i) - v(c_\sigma)\| \leq 1$  if and only if  $\sigma_i = +$ . Moreover, since  $w(c_\sigma), w(c_\tau) \in B$ ,

$$\|v(c_\sigma) - v(c_\tau)\| = \|w(c_\sigma) - w(c_\tau)\|/r \leq 2\varepsilon/r \leq 1,$$

for all  $\sigma, \tau \in S$ . Similarly, we find

$$\begin{aligned} \|v(a_i) - v(b_j)\|, \|v(b_i) - v(b_j)\| &\leq \|u_i - u_j\| + \|p_i - p_j\|/r \\ &\leq 6\varepsilon + 2\varepsilon/r \\ &\leq 1, \end{aligned}$$

where we have used (6), and

$$\begin{aligned} \|v(a_i) - v(b_j)\| &= \|r(u_i + u_j) + p_i - p_j\|/r \\ &\geq 2 - 6\varepsilon - 2\varepsilon/r \\ &> 1. \end{aligned}$$

Thus,  $v(\cdot)$  defines a good representation of  $G$  as a  $k$ -sphere graph and we are done.  $\square$

*Proof of Theorem 2* We again use a reduction from  $k$ -REALIZABILITY. Given  $S \subseteq \{-, +\}^n$  with  $(-, \dots, -), (+, \dots, +) \in S$ , let  $G = G(S)$  be a graph on  $n + |S|$  vertices which is defined as follows. Let us write  $V(G) = A \cup B$ , where  $A = \{a_1, \dots, a_n\}$  and  $B = \{b_\sigma : \sigma \in S\}$ . The edge set of  $G$  is as follows:

- $a_i a_j \in E(G)$  for all  $1 \leq i < j \leq n$ ;
- $b_\sigma b_\tau \notin E(G)$  for all  $\sigma \neq \tau \in S$ ;
- $b_\sigma a_i \in E(G)$  if and only if  $\sigma_i = +$ .

It should be clear that  $G$  can be constructed from  $S$  in polynomial time. It remains to show that  $G$  is a  $k$ -dot product graph if and only if there exists a  $k$ -hyperplane arrangement  $\mathcal{H}$  with  $S \subseteq \mathcal{D}(\mathcal{H})$ .

First, let us assume that  $G$  is a  $k$ -dot product graph. Thus, there is a function  $v : V(G) \rightarrow \mathbb{R}^k$  such that  $v(s)^T v(t) \geq 1$  if and only if  $st \in E(G)$ . For each  $i = 1, \dots, n$ , let  $c_i$  be a number with

$$\max\{v(a_i)^T v(b_\sigma) : \sigma \in S, \sigma_i = -\} < c_i < 1. \quad (7)$$

This choice of  $c_i$  ensures that  $v(a_i)^T v(b_\sigma) < c_i$  if  $\sigma_i = -$ , and  $v(a_i)^T v(b_\sigma) > c_i$  if  $\sigma_i = +$ . Now, let the oriented hyperplane arrangement  $\mathcal{H}$  be defined by

$$\begin{aligned} h_i &:= \{z : v(a_i)^T z = c_i\}, & h_i^- &:= \{z : v(a_i)^T z < c_i\}, \\ h_i^+ &:= \{z : v(a_i)^T z > c_i\}. \end{aligned}$$

Under this definition, it is clear that, for each  $\sigma \in S$ , the sign vector of  $v(b_\sigma)$  with respect to  $\mathcal{H}$  is exactly  $\sigma$ .

Now, suppose that there exists some  $\mathcal{H}$  such that  $S \subseteq \mathcal{D}(\mathcal{H})$ . We may assume that  $\mathcal{H}$  satisfies the conditions of Lemma 12 with  $K := 1000$ ,  $\varepsilon := 1/k$ . Let  $v_1, \dots, v_n$  and  $c_1, \dots, c_n$  be as provided by part (iii) of Lemma 12. For notational convenience, let us write  $B := B((K, 0, \dots, 0), \varepsilon)$ .

For each  $\sigma \in S$ , let us pick a vector  $v(b_\sigma) \in B$  such that  $v(b_\sigma)$  has sign vector  $\sigma$  with respect to  $\mathcal{H}$ . For each  $i = 1, \dots, n$ , let us set  $v(a_i) := v_i/c_i$ . It is clear that

$$v(b_\sigma)^T v(b_\tau) \geq (K - \varepsilon)^2 - (k - 1)\varepsilon^2 \geq 1,$$

for all  $\sigma, \tau$  and that  $v(a_i)^T v(b_\sigma) \geq 1$  if and only if  $\sigma_i = +$ .

We now claim that  $\|v(a_i)\| < 1$  for all  $i$ . To see this, recall that by part (iv) of Lemma 12 there is some  $p \in B$  such that  $(\sigma(p))_i = -$  and hence  $c_i > v_i^T p$ . Since  $\|v_i - e_1\| < \varepsilon$ , we have

$$c_i > v_i^T p > (1 - \varepsilon)(K - \varepsilon) - (k - 1)\varepsilon^2.$$

As  $\|v_i\| < 1 + \varepsilon$ , we indeed obtain

$$\begin{aligned} \|v(a_i)\| &= \|v_i\|/c_i \\ &\leq (1 + \varepsilon)/((1 - \varepsilon)K - (k - 1)\varepsilon^2) \\ &< 1. \end{aligned}$$

(Here we used that, since the origin lies in the all-minus cell, we have  $c_i > 0$  for all  $i = 1, \dots, n$ .) This proves the claim that  $\|v(a_i)\| < 1$  for all  $i$ , which in turn gives that  $v(a_i)^T v(a_j) < 1$  for all  $i, j$ .

Thus, we have constructed a valid  $k$ -dot product representation of  $G$ , which concludes the proof.  $\square$

We now turn our attention towards the proofs of Theorems 3 and 4. For the upper bound on the bit sizes, we will use a result by Basu, Pollack and Roy. The following is a reformulation of Theorem 1.3.5 of [1]:

**Theorem 13** [1] *For each  $d, \tau \in \mathbb{N}$ , there exists a constant  $C = C(d, K)$  such that the following holds. Suppose that  $\mathcal{F}$  is a set of polynomials in  $n$  variables, each with degree at most  $d$  and with integer coefficients of bit size at most  $\tau$ . If there exists a solution  $(x_1, \dots, x_n) \in \mathbb{R}^n$  of the system  $\{f(x) > 0 : f \in \mathcal{F}\}$ , then there also exists one with each  $x_i$  a rational  $x_i = a_i/b_i$  with  $\text{size}(a_i), \text{size}(b_i) \leq \tau \cdot d^{Cn}$ .*

*Proof of the upper bound in Theorem 3* The vectors  $v_1, \dots, v_n \in \mathbb{R}^k$  form a  $k$ -sphere representation of  $G$  if and only if they satisfy the following set of inequalities:

$$\sum_{t=1}^k ((v_i)_t - (v_j)_t)^2 \leq 1 \quad \text{for all } ij \in E(G),$$

$$\sum_{t=1}^k ((v_i)_t - (v_j)_t)^2 > 1 \quad \text{for all } ij \notin E(G).$$
(8)

Suppose they do. Now, notice that, if we “shrink” the vectors very slightly (i.e. if we put  $v'_i := \lambda v_i$  for some  $\lambda < 1$  but  $\lambda$  very close to 1), then we obtain a representation in which all inequalities of (8) are strict. Observe that all degrees of the polynomials in (8) are two, and all coefficients are at most two in absolute value.

Since the inequalities can be taken strict, we can apply Theorem 13 and find a solution  $v_1, \dots, v_n$  of (8) such that each coordinate of each  $v_i$  is a rational number  $a/b$  with  $\text{size}(a), \text{size}(b) \leq 2^{O(n)}$ . Hence, the sum of the bit sizes of the coordinates is no more than  $2kn2^{O(n)} = 2^{O(n)}$ . This concludes the proof. □

*Proof of the upper bound in Theorem 4* The vectors  $v_1, \dots, v_n \in \mathbb{R}^k$  form a  $k$ -dot product representation of  $G$  if and only if

$$v_i^T v_j \geq 1 \quad \text{for all } ij \in E(G),$$

$$v_i^T v_j < 1 \quad \text{for all } ij \notin E(G).$$
(9)

Suppose they do. By “expanding” very slightly (i.e. by multiplying all vectors by the same scalar  $\lambda > 1$  with  $\lambda$  very close to 1), we get a solution of (9) for which all inequalities are strict. We now proceed completely analogously to the proof of Theorem 3. □

An *intersection point* of a  $k$ -hyperplane arrangement  $\mathcal{H}$  is a point that is the unique point common to a subset of the hyperplanes of  $\mathcal{H}$ . That is, there are hyperplanes  $h_{i_1}, \dots, h_{i_m} \in \mathcal{H}$  such that

$$\{p\} = h_{i_1} \cap \dots \cap h_{i_m}.$$

Let us observe that we can always take such a subset to consist of precisely  $m = k$  hyperplanes. Let  $i(\mathcal{H})$  denote the set of intersection points of  $\mathcal{H}$ .

The *span*  $\text{span}(\mathcal{H})$  of a hyperplane arrangement  $\mathcal{H}$  denotes the ratio of the largest distance between two distinct intersection points to the smallest non-zero distance

between two intersection points. That is,

$$\text{span}(\mathcal{H}) := \max_{p,q \in i(\mathcal{H})} \|p - q\| / \min_{\substack{p,q \in i(\mathcal{H}) \\ p \neq q}} \|p - q\|.$$

The proof of the lower bound makes use of the following recent result of McDiarmid and the second author [12] (the result stated is an immediate corollary of Theorem 3.2 in [12]):

**Theorem 14** [12] *There exist universal constants  $\alpha, \beta$  such that the following holds. For every  $r \in \mathbb{N}$ , there exists  $n \leq \alpha r$  and a set  $S \subseteq \{-, +\}^n$  with  $|S| \leq \beta r$  such that*

- (i) *there exists a line arrangement  $\mathcal{L}$  with  $S \subseteq \mathcal{D}(\mathcal{L})$ , and*
- (ii)  *$\text{span}(\mathcal{L}) \geq 2^{2^r}$  for every line arrangement with  $S \subseteq \mathcal{D}(\mathcal{L})$ .*

This result has the following corollary.

**Corollary 15** *For every  $k > 1$ , there exist  $\alpha = \alpha(k), \beta = \beta(k) > 0$  such that the following holds. For every  $r \in \mathbb{N}$ , there exists  $n \leq \alpha r$  and a set  $S \subseteq \{-, +\}^n$  such that  $|S| \leq \beta r$  and*

- (i)  $(-, \dots, -), (+, \dots, +) \in S$ ,
- (ii) *there exists a  $k$ -hyperplane arrangement  $\mathcal{L}$  with  $S \subseteq \mathcal{D}(\mathcal{L})$ , and*
- (iii)  *$\text{span}(\mathcal{L}') \geq 2^{2^r}$  for every  $k$ -hyperplane arrangement  $\mathcal{L}'$  with  $S \subseteq \mathcal{D}(\mathcal{L}')$ .*

*Proof* We start with the case  $k = 2$ . Let  $\alpha, \beta$  be as in Theorem 14. Let  $r \in \mathbb{N}$  be arbitrary, and let  $n = n(r), S = S(r)$  as provided by Theorem 14. Let  $\mathcal{L}$  be an oriented line arrangement with  $S \subseteq \mathcal{D}(\mathcal{L})$ . By Lemma 5, we can find a pair  $s, -s \in \mathcal{D}(\mathcal{L})$ . It is clear that, if we set  $S' := S \cup \{s, -s\}$ , then  $S'$  still has properties (i) and (ii) of Theorem 14.

Recall that in a reorientation of  $\mathcal{L}$  we switch the roles of  $\ell_i^-$  and  $\ell_i^+$  for some indices  $i$ . Thus, a reorientation certainly preserves  $\text{span}(\mathcal{L})$ . Let  $f : \{-, 0, +\}^n \rightarrow \{-, 0, +\}^n$  be the “reorientation” map from the proof of Theorem 10 with  $s \mapsto (-, \dots, -)$  and  $-s \mapsto (+, \dots, +)$ , and set  $S' := f[S \cup \{s, -s\}]$ . Then  $|S'| \leq \beta r + 2 \leq 2\beta r$ . Moreover, any oriented line arrangement  $\mathcal{L}'$  with  $S' \subseteq \mathcal{D}(\mathcal{L}')$  can be reoriented (via the same map  $f$ , in fact) to an oriented line arrangement  $\mathcal{L}$  with  $S \cup \{s, -s\} \subseteq \mathcal{D}(\mathcal{L})$ . Hence, any such oriented line arrangement has  $\text{span}(\mathcal{L}') \geq 2^{2^r}$ . This proves the case  $k = 2$  of Corollary 15 with  $\alpha(2) := \alpha, \beta(2) := 2\beta$ .

Let us now assume that Corollary 15 has been proved for  $k \geq 2$ . Let  $r \in \mathbb{N}$  be arbitrary, and let  $\alpha(k), \beta(k), n = n(k, r), S = S(k, r)$  be as provided by the case  $k$  of Corollary 15. Let us set

$$\begin{aligned} \alpha(k + 1) &:= 2\alpha(k), & \beta(k + 1) &:= 2\beta(k), & n(k + 1, r) &= n(k, r) + 1, \\ S' = S(k + 1, r) &:= S \times \{-, +\}. \end{aligned}$$

These choices clearly satisfy  $n(k + 1, r) \leq \alpha(k + 1)r, |S(k + 1, r)| \leq \beta(k + 1)r$  and  $(-, \dots, -), (+, \dots, +) \in S(k + 1, r)$ . By Lemma 11, there is at least one oriented

$(k + 1)$ -hyperplane arrangement  $\mathcal{H}'$  such that  $S' \subseteq \mathcal{D}(\mathcal{H}')$ . This shows part (ii) of Corollary 15 also holds. It remains to show part (iii).

Now, let  $\mathcal{H}'$  be an arbitrary oriented  $(k + 1)$ -hyperplane arrangement  $\mathcal{H}'$  with  $S' \subseteq \mathcal{D}(\mathcal{H}')$ . Let us apply an isometry  $T : \mathbb{R}^k \rightarrow \mathbb{R}^k$  that maps  $h'_{n+1}$  to  $\mathbb{R}^{k-1} \times \{0\}$ . Since  $T$  is an isometry, this does not affect  $\text{span}(\mathcal{H}')$ . If the oriented  $k$ -hyperplane arrangement  $\mathcal{H}$  is defined as in (4), then the proof of Lemma 11 shows that  $S \subseteq \mathcal{D}(\mathcal{H})$  and hence  $\text{span}(\mathcal{H}) \geq 2^{2^r}$ . Now, let  $p \in i(\mathcal{H})$  be arbitrary. Then  $p$  is the unique intersection point of some  $k$ -hyperplanes  $h_{i_1}, \dots, h_{i_k}$  of  $\mathcal{H}$ . Thus, we can write

$$\{p\} = h_{i_1} \cap \dots \cap h_{i_k} = h'_{i_1} \cap \dots \cap h'_{i_k} \cap h'_{n+1}.$$

Hence, we also have  $p \in i(\mathcal{H}')$ . Since  $p$  was arbitrary, we have  $i(\mathcal{H}) \subseteq i(\mathcal{H}')$ , which immediately implies

$$\text{span}(\mathcal{H}') \geq \text{span}(\mathcal{H}) \geq 2^{2^r}.$$

Because  $\mathcal{H}'$  was an arbitrary  $(k + 1)$ -hyperplane arrangement with  $S \subseteq \mathcal{D}(\mathcal{H}')$ , this proves part (iii) of Corollary 15. □

Before we continue with the proofs of the lower bounds in Theorems 3 and 4, it is convenient to state one more auxiliary result. We have already defined the bit size of an integer, a rational and a rational vector. For  $A$  a rational matrix, we defined  $\text{size}(A) := \sum_{i,j} \text{size}(A_{ij})$ . The following is a reformulation of Corollary 3.2b in [20]:

**Lemma 16** [20] *There exists a universal constant  $K > 0$  such that the following holds. If the system  $Ax = b$  of rational linear equations has a solution, then there is also a solution with  $\text{size}(x) \leq (\text{size}(A) + \text{size}(b))^K$ .*

Here it should be noted that Schrijver actually uses a notion of the bit size of a vector or matrix that is slightly different from ours. Schrijver’s bit size is however never more than a factor two larger than our bit size and never smaller than half our bit size, so that Lemma 16 is a straightforward corollary of Corollary 3.2b in [20].

With Lemma 16 and Corollary 15 in hand, we are now able to prove the lower bounds in Theorems 3 and 4.

*Proof of the lower bound in Theorem 3* Let us fix an arbitrary  $r \in \mathbb{N}$  and let  $n = n(r)$ ,  $S = S(r)$  be as provided by Corollary 15. Let  $G = G(S)$  be as in the proof of Theorem 1. Now, let  $v : V(G) \rightarrow \mathbb{R}^k$  be an arbitrary  $k$ -sphere representation of  $G$ , and let  $\mathcal{H} = (h_1, \dots, h_n)$  be as defined in (5). As shown in the proof of Theorem 1, we have  $S \subseteq \mathcal{D}(\mathcal{H})$ . Observe that we can also express  $h_i$  as

$$h_i = \{z \in \mathbb{R}^k : (v(a_i) - v(b_i))^T z = (v(a_i) - v(b_i))^T (v(a_i) + v(b_i))/2\}. \quad (10)$$

Now, let  $p$  be an intersection point of  $\mathcal{H}$ . Then there are  $h_{i_1}, \dots, h_{i_k} \in \mathcal{H}$  such that  $\{p\} = h_{i_1} \cap \dots \cap h_{i_k}$ . Using (10), we can write  $p$  as the unique solution of

$$Ap = b,$$



where the rows of  $A$  are  $(v(a_{i_1}) - v(b_{i_1}))^T, \dots, (v(a_{i_k}) - v(b_{i_k}))^T$  and the entries of  $b$  are  $(v(a_{i_1}) - v(b_{i_1}))^T (v(a_{i_1}) + v(b_{i_1}))/2, \dots, (v(a_{i_k}) - v(b_{i_k}))^T (v(a_{i_k}) + v(b_{i_k}))/2$ .

Let  $\sigma \in \mathbb{N}$  be such that  $\sum_{t \in V(G)} \text{size}(v(t)) \leq \sigma$ . It is easily seen that  $\text{size}(A) + \text{size}(b) = O(\sigma^C)$  for some  $C = C(k) > 0$ , so that by Theorem 16 every intersection point  $p$  also satisfies  $\text{size}(p) \leq \sigma^K$  for some constant  $K = K(k) > 0$ .

Now, let  $p, q$  be two distinct intersection points. Observe that by definition of bit size,

$$|p_j - q_j| \leq 2^{\sigma^K},$$

for each  $j = 1, \dots, k$ . Similarly,

$$\text{if } p_j \neq q_j \text{ then } |p_j - q_j| \geq 1/2^{2\sigma^K}.$$

Thus, we have that

$$\|p - q\| \leq \sum_{j=1}^k |p_j - q_j| \leq k2^{2\sigma^K}$$

and

$$\|p - q\| \geq \max_j |p_j - q_j| \geq 1/2^{2\sigma^K}$$

for any pair of distinct intersection points. We see that

$$k2^{4\sigma^K} \geq \text{span}(\mathcal{H}) \geq 2^{2^r}.$$

This amounts to

$$\sigma \geq \left( \frac{2^r - \log_2 k}{4} \right)^{1/K} = 2^{(r-2-\log_2(1-\log_2 k/2^r))/K} \geq 2^{cr} \tag{11}$$

for some  $c > 0$ , provided  $r$  is sufficiently large. Since  $v : V(G) \rightarrow \mathbb{R}^k$  was an arbitrary  $k$ -sphere representation,

$$f_k(G) \geq 2^{cr},$$

for  $r$  sufficiently large.

We now claim that this implies

$$f_k(n) \geq 2^{c'n} \tag{12}$$

for all  $n$ , with  $c' > 0$  a suitably chosen universal constant. To see this, pick an arbitrary  $n \in \mathbb{N}$  and set  $r := \lfloor n/(2\alpha + \beta) \rfloor$ . For  $n$  sufficiently large, our proof gives a graph  $G$  on at most  $n$  vertices with  $f_k(G) \geq 2^{cr} \geq 2^{n \cdot c/(4\alpha + 2\beta)}$ . We can add additional isolated vertices to get a graph with exactly  $n$  vertices. This shows  $f_k(n) \geq 2^{n \cdot c/(4\alpha + 2\beta)}$  for all  $n \geq n_0$ , where  $n_0$  is some constant. Now, note that trivially  $f_k(n) \geq 2kn$  for all  $n$

(since we need at least one bit for each denominator and numerator). Hence, if we set  $c' := \min(\frac{1}{n_0}, \frac{c}{4\alpha+2\beta})$ , then (12) indeed holds for all  $n$ .  $\square$

*Proof of the lower bound in Theorem 4* Let us fix an arbitrary  $r \in \mathbb{N}$  and let  $n = n(r)$ ,  $S = S(r)$  be as provided by Corollary 15. Let  $G = G(S)$  be as in the proof of Theorem 2, and let  $v : V(G) \rightarrow \mathbb{R}^k$  be an arbitrary  $k$ -dot product representation of  $G$ . For  $i = 1, \dots, n$ , let us set

$$c_i := \frac{1}{2} (1 + \max\{v(a_i)^T v(b_\sigma) : \sigma \in S, \sigma_i = -\}).$$

Then the  $c_i$  clearly satisfy (7). (By the construction of  $G(S)$ , we have  $v(a_i)^T v(b_\sigma) < 1$  for all  $i$  and  $\sigma \in S$  with  $\sigma_i = -$ .) Now, let  $\mathcal{H}$  be defined by

$$\begin{aligned} h_i &:= \{z : v(a_i)^T z = c_i\}, & h_i^- &:= \{z : v(a_i)^T z < c_i\}, \\ h_i^+ &:= \{z : v(a_i)^T z > c_i\}, \end{aligned}$$

as in the proof of Theorem 2. Thus, we have  $S \subseteq \mathcal{D}(\mathcal{H})$ .

Let  $p \in i(\mathcal{H})$  be an intersection point. Then there are hyperplanes  $h_{i_1}, \dots, h_{i_k}$  of  $\mathcal{H}$  such that  $\{p\} = h_{i_1} \cap \dots \cap h_{i_k}$ . In other words,  $p$  is the unique solution of the matrix equation

$$Ap = b,$$

where  $A$  is a  $(k \times k)$ -matrix with rows  $v(a_{i_1})^T, \dots, v(a_{i_k})^T$  and  $b \in \mathbb{R}^k$  has coordinates  $c_{i_1}, \dots, c_{i_k}$ . Observe that, writing  $\sigma := \sum_{t \in V(G)} \text{size}(v(t))$ , we have  $\text{size}(A) + \text{size}(b) = O(\sigma^K)$  for some constant  $K = K(k)$ . We can thus proceed as in the proof of the lower bound in Theorem 3 to show that

$$f_k(n) \geq 2^{cn},$$

for some universal constant  $c = c(k)$ , as required.  $\square$

### 4 Discussion and Further Work

Mněv’s work [14, 15] (see also [3, 21]) actually proves that SIMPLE STRETCH-ABILITY is polynomially equivalent to the decision problem for the “existential theory of the reals” (ERT). ERT is the problem of deciding, given as input a set of polynomial equalities, inequalities and strict inequalities with integer coefficients, whether there exists a simultaneous real solution. While membership in NP for  $k$ -SPHERICITY and  $k$ -DOT PRODUCT remains open, our proofs in fact imply that both are polynomially equivalent to ERT. So our results imply in particular that if one were to give a polynomial certificate for either  $k$ -SPHERICITY or  $k$ -DOT PRODUCT then this would prove that ERT is in NP. As ERT is a pretty central notion in complexity theory (many other decision problems can for instance be written as special cases of ERT), proving that either  $k$ -SPHERICITY or  $k$ -DOT PRODUCT is

in NP would thus constitute a minor breakthrough in complexity theory. For more discussion of the complexity class characterized by ERT as well as its connection to other geometric and topological graph representations, see the recent survey by Schaefer [19].

We have seen that the vectors of a  $k$ -sphere (resp.  $k$ -dot product) representation may require exponentially many bits to be stored in the memory of a computer, which means that they cannot serve as a polynomial certificate for the recognition problem and certainly not as a labelling scheme to prove a special case of the implicit graph conjecture. However, this is only if we use the obvious way of encoding the vectors as bit strings using the binary number format. Perhaps there is a more economical way to “encode the geometry” of the vectors.

**Acknowledgements** We thank the anonymous referee for helpful comments and pointers to the literature that have improved the presentation.

R.J. Kang’s research partially supported by the Engineering and Physical Sciences Research Council (EPSRC), Grant EP/G066604/1.

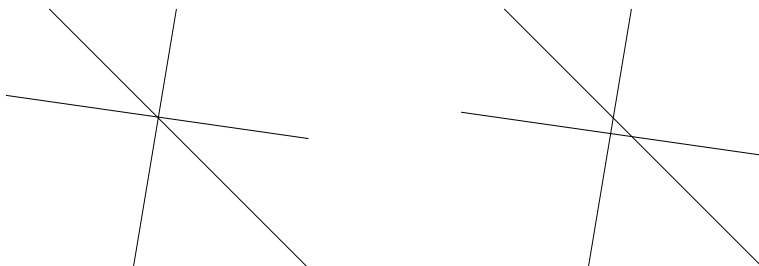
T. Müller’s research partially supported by a VENI Grant from Netherlands Organization for Scientific Research (NWO).

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## Appendix: The Missing Proofs

*Proof of Lemma 5* Let  $\mathcal{L} = (\ell_1, \dots, \ell_n)$  be an arbitrary line arrangement. Let  $\ell$  be a line whose slope is different from the slopes of  $\ell_1, \dots, \ell_n$ . Then each line  $\ell_i$  intersects  $\ell$  in a point  $p_i$ . Let  $p, q \in \ell$  be points such that the open segment  $(p, q) \subseteq \ell$  contains all these intersection points. Hence,  $p$  and  $q$  lie on opposite sides of  $\ell_i$  for every  $i$ . This implies that  $(\sigma(p))_i = -(\sigma(q))_i$  for all  $i$ .  $\square$

*Proof of Lemma 6* Let us first observe that every non-simple line arrangement has strictly fewer cells than a simple line arrangement on the same number of lines. This is because, if three or more lines go through a common point  $p$ , then we can shift one of the lines very slightly so that it no longer goes through  $p$  and at least one extra cell is created.



It remains to show that a simple line arrangement  $\mathcal{L} = (\ell_1, \dots, \ell_n)$  has exactly  $1 + \binom{n+1}{2}$  lines. We shall use induction on  $n$ . When  $n = 1$ , it is clear that there are exactly two regions. Now, suppose that every simple line arrangement on  $n - 1$  lines has exactly  $1 + \binom{n}{2}$  cells. The line arrangement  $\mathcal{L}' = (\ell_1, \dots, \ell_{n-1})$  is certainly simple, and therefore has exactly  $1 + \binom{n}{2}$  cells. Since  $\mathcal{L}$  is simple,  $\ell_n$  intersects every line of  $\mathcal{L}'$  exactly once. Thus, the intersection points with the  $n - 1$  lines of  $\mathcal{L}'$  divide  $\ell_n$  into exactly  $n$  segments. Each of these segments divides a cell of the old line arrangement  $\mathcal{L}'$  into two cells of the new line arrangement  $\mathcal{L}$ . Hence, the number of cells of  $\mathcal{L}$  is precisely  $n + (1 + \binom{n}{2}) = 1 + \binom{n+1}{2}$ .  $\square$

*Proof of Lemma 7* It suffices to show that  $\mathcal{D}(\mathcal{L}) \cap \{-, +\}^n$  determines all the other sign vectors of  $\mathcal{D}(\mathcal{L})$ . To see this, notice that two cells have a common edge if and only if their sign vectors differ in precisely one coordinate. Hence, we can reconstruct the sign vectors with exactly one zero, by finding all pairs of sign vectors in  $\mathcal{D}(\mathcal{L}) \cap \{-, +\}^n$  that differ in exactly one coordinate, and substituting 0 in this coordinate.

Similarly, two edges belonging to the same line (which is true if and only if the unique zeroes of their sign vectors occur in the same position) meet in an intersection point if and only if their sign vectors differ in exactly one coordinate. Hence, we can reconstruct sign vectors with exactly two zeroes by finding all pairs of sign vectors with exactly one zero that agree on the position of the zero and in all other positions but one, and then substituting 0 in the position where they disagree.

Since there are no sign vectors with more than two zeroes in a simple line arrangement, this procedure will indeed produce  $\mathcal{D}(\mathcal{L})$ .  $\square$

*Proof of Lemma 8* Let  $z \in \mathbb{R}^k$  and  $1 \leq i \leq n$  be arbitrary. Observe that  $z \in h_i^-$  if and only if  $T(z) \in (h_i')^-$ , and similarly for  $h_i^+, h_i$ . This shows that  $\sigma_{\mathcal{H}}(z) = \sigma_{\mathcal{H}'}(T(z))$  for all  $z \in \mathbb{R}^k$ . Since  $T$  is a bijection, we must have  $\mathcal{D}(\mathcal{H}) = \mathcal{D}(\mathcal{H}')$ , as required.  $\square$

## References

1. Basu, S., Pollack, R., Roy, M.-F.: On the combinatorial and algebraic complexity of quantifier elimination. *J. ACM* **43**(6), 1002–1045 (1996)
2. Basu, S., Pollack, R., Roy, M.-F.: *Algorithms in Real Algebraic Geometry*. Algorithms and Computation in Mathematics, 2nd edn. Springer, Berlin (2006)
3. Björner, A., Las Vergnas, M., Sturmfels, B., White, N., Ziegler, G.M.: *Oriented Matroids*. Encyclopedia of Mathematics and Its Applications, vol. 46, 2nd edn. Cambridge University Press, Cambridge (1999)
4. Breu, H., Kirkpatrick, D.G.: Unit disk graph recognition is NP-hard. *Comput. Geom.* **9**(1–2), 3–24 (1998)
5. Felsner, S.: In: *Geometric Graphs and Arrangements*. Advanced Lectures in Mathematics. Friedr. Vieweg, Wiesbaden (2004). Some chapters from combinatorial geometry
6. Fiduccia, C.M., Scheinerman, E.R., Trenk, A., Zito, J.S.: Dot product representations of graphs. *Discrete Math.* **181**(1–3), 113–138 (1998)
7. Havel, T.F.: The combinatorial distance geometry approach to the calculation of molecular conformation. Ph.D. thesis, University of California, Berkeley (1982)
8. Kannan, S., Naor, M., Rudich, S.: Implicit representation of graphs. *SIAM J. Discrete Math.* **5**(4), 596–603 (1992)
9. Looges, P.J., Olariu, S.: Optimal greedy algorithms for indifference graphs. *Comput. Math. Appl.* **25**(7), 15–25 (1993)

10. Maehara, H.: Space graphs and sphericity. *Discrete Appl. Math.* **7**(1), 55–64 (1984)
11. Mahadev, N.V.R., Peled, U.N.: *Threshold Graphs and Related Topics*. *Annals of Discrete Mathematics*, vol. 56. North-Holland, Amsterdam (1995)
12. McDiarmid, C.J.H., Müller, T.: Integer representation of disk and segment graphs (2011, submitted). arXiv: [1111.2931](https://arxiv.org/abs/1111.2931)
13. McDiarmid, C.J.H., Müller, T.: The number of bits needed to represent a unit disk graph. In: *Proceedings of 36th International Workshop on Graph Theoretic Concepts in Computer Science (WG 2010)*. *Lecture Notes in Computer Science*, vol. 6410, pp. 315–323 (2010)
14. Mněv, N.E.: Realizability of combinatorial types of convex polyhedra over fields. *J. Sov. Math.* **28**, 606–609 (1985) (in Russian)
15. Mněv, N.E.: The universality theorems on the classification problem of configuration varieties and convex polytopes varieties. In: *Topology and Geometry—Rohlin Seminar*. *Lecture Notes in Math.*, vol. 1346, pp. 527–543. Springer, Berlin (1988)
16. Reiterman, J., Rödl, V., Šiňajová, E.: Embeddings of graphs in Euclidean spaces. *Discrete Comput. Geom.* **4**(4), 349–364 (1989)
17. Reiterman, J., Rödl, V., Šiňajová, E.: Geometrical embeddings of graphs. *Discrete Math.* **74**(3), 291–319 (1989)
18. Reiterman, J., Rödl, V., Šiňajová, E.: On embedding of graphs into Euclidean spaces of small dimension. *J. Comb. Theory, Ser. B* **56**(1), 1–8 (1992)
19. Schaefer, M.: Complexity of some geometric and topological problems. In: *Graph Drawing, 17th International Symposium (GD 2009)*, Chicago, 22–25 September 2009. *Lecture Notes in Computer Science*, vol. 5849, pp. 334–344 (2010)
20. Schrijver, A.: *Theory of Linear and Integer Programming*. *Wiley-Interscience Series in Discrete Mathematics*. Wiley, Chichester (1986)
21. Shor, P.W.: Stretchability of pseudolines is NP-hard. In: *Applied Geometry and Discrete Mathematics*. *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, vol. 4, pp. 531–554. Am. Math. Soc., Providence (1991)
22. Spinrad, J.P.: *Efficient Graph Representations*. *Fields Institute Monographs*, vol. 19. Am. Math. Soc., Providence (2003)