

Recognising Multidimensional Euclidean Preferences

Dominik Peters

Abstract

Euclidean preferences are a widely studied preference model, in which decision makers and alternatives are embedded in d -dimensional Euclidean space. Decision makers prefer those alternatives closer to them. This model, also known as multidimensional unfolding, has applications in economics, psychometrics, marketing, and many other fields. We study the problem of deciding whether a given preference profile is d -Euclidean. For the one-dimensional case, polynomial-time algorithms are known. We show that, in contrast, for every other fixed dimension $d > 1$, the recognition problem is equivalent to the existential theory of the reals (ETR), and so in particular NP-hard. We further show that some Euclidean preference profiles require exponentially many bits in order to specify any Euclidean embedding, and prove that the domain of d -Euclidean preferences does not admit a finite forbidden minor characterisation for any $d > 1$. We also study dichotomous preferences and the behaviour of other metrics, and survey a variety of related work.

1 Introduction

The study of preferences spans a multitude of fields: economics (and particularly game theory and social choice), political science, psychology, multi-agent systems, marketing, and others. An important element of working with preferences is *understanding* them by constructing models for them and identifying underlying structure. For example, in a psychological model of preferences, we aim to discover which underlying psychological process has generated the preferences we now observe (Coombs 1964). In political science, one might wonder about the underlying structure of ‘political space’ by analysing voter preferences (Merrill and Grofman 1999). In economics, working with well-structured preferences often allows a model to become analytically tractable.

This paper will mainly use the lens and language of *computational social choice*, but our results also apply to a formal study of any of the concerns outlined above. In recent years, much work in computational social choice has focussed on identifying structure in a given preference profile. The reason for this is simple: many of the problems social choice aims to solve (such as preference aggregation, committee selection, or fair and efficient allocation) are computationally hard. However, if we manage to identify underlying, hopefully low-dimensional structure in the input profile, we can exploit this structure to guide algorithms. This approach has been successful, particularly for the domain of *single-peaked preferences*: once we have identified an *axis* on which an input profile is single-peaked, we may efficiently find a Kemeny ranking (Brandt et al. 2010), identify an optimal committee according to the Chamberlin-Courant rule (Betzler et al. 2013), calculate possible winners for many voting rules (Faliszewski et al. 2009), and solve the stable roommates problem (Bartholdi and Trick 1986).

This leaves the question of whether we can, in fact, efficiently find a certificate for single-peakedness (or another desired domain restriction) that such algorithms can use. Fortunately, the answer is often positive: there are efficient and certificate-producing recognition algorithms for preferences that are *single-peaked* (Escoffier et al. 2008), *single-crossing* (Elkind et al. 2012), *1-Euclidean* (Doignon and Falmagne 1994), or *single-peaked on a tree* (Trick 1989). Indeed, each of these preference domains is quite well-understood, for example

in terms of forbidden-substructure characterisations (Ballester and Haeringer 2011), concise representations of all certificates (Peters and Elkind 2016; Bartholdi and Trick 1986), containment relations between the different domains (Elkind et al. 2014), and the probability with which a random preference profile falls within a given domain (Bruner and Lackner 2015).

Intriguingly, there is another preference domain for which we do not have a comparable amount of understanding, yet it is an extremely popular modelling choice across disciplines. Known as *spatial preferences*, or as the *d-Euclidean domain*, or as *multidimensional unfolding*, this preference domain contains profiles that can be ‘embedded’ into d -dimensional Euclidean space. Precisely, a preference profile is d -Euclidean if we can assign every voter and every alternative a point in \mathbb{R}^d such that voters prefer those alternatives that are closer to them (according to the usual Euclidean metric) to those that are further away. This characterisation of preferences has intuitive appeal: considering \mathbb{R}^d as a continuous ‘policy space’, within which alternatives can vary along multiple dimensions, each voter is identified with an *ideal point* (Bennett and Hays 1960). The best alternative for the voter is the one that minimises the distance to the ideal policy. We could also think of a facility location problem, where a single facility needs to be placed somewhere in the plane, with each decision-maker preferring the facility to be placed as close to them as possible (Hotelling 1929).

In practice, the embedding of voters and alternatives into \mathbb{R}^d is hidden, and we only have access to the ordinal ranking data in form of a preference profile (Anshelevich et al. 2015). Given this data, can we recover an appropriate embedding into \mathbb{R}^d that explains the preferences? This problem is known as multidimensional unfolding (Bennett and Hays 1960), and a large variety of methods that attempt to estimate embeddings have been proposed in the statistics and psychometrics literature [for a modern exposition see, e.g., Borg and Groenen 2005]. None of these methods is guaranteed to return a suitable embedding whenever it exists and terminate in polynomial time.

We analyse this problem from a formal computational perspective. In particular, we prove that the decision problem of identifying d -Euclidean preference profiles is NP-hard for each fixed $d \geq 2$. More precisely, we prove that the problem is equivalent to the existential theory of the reals (ETR), and thus is in fact $\exists\mathbb{R}$ -complete. This means that the recognition problem is unlikely to be contained in NP (though it is decidable in PSPACE). Using recent results about hyperplane arrangements, we deduce that, for each fixed $d \geq 2$, there exist d -Euclidean preference profiles such that the coordinates of any d -Euclidean embedding require exponentially many bits to specify. Thus, there is provably no polynomial-time algorithm that, on input a preference profile, outputs a d -Euclidean embedding if one exists.

While the domains of single-peaked and of single-crossing preferences admit characterisations by a finite list of forbidden configurations (Ballester and Haeringer 2011; Bredereck et al. 2013b), the 1-Euclidean domain does not admit such a characterisation (Chen et al. 2015). Chen et al. (2015) conjectured that the same is true for the domain of d -Euclidean preferences, for each fixed $d \geq 2$. We use a connection with the theory of oriented matroids to prove their conjecture.

The results in this paper cast doubt on the fruitfulness of exploiting structure in d -Euclidean preferences in computational social choice, and they also limit the extent to which the spatial preference model can be used to reliably explain observed preference data. Future work could explore ways of mitigating this situation. At the end of this paper, we briefly sketch one way that this could be done, namely by replacing the Euclidean ℓ_2 -metric by the ℓ_1 - or ℓ_∞ -metric. We show that the recognition problems corresponding to these metrics are contained in NP.

2 Preliminaries

2.1 Euclidean preferences

Let A be a finite set of *alternatives* or *candidates*. A *preference relation* \succsim over A (usually referred to as a *vote*) is a complete and transitive binary relation over A . We denote by \succ the strict part of \succsim , that is, $a \succ b$ if and only if $a \succsim b$ but $b \not\succsim a$; and we denote by \sim the indifference part of \succsim , with $a \sim b$ if and only if $a \succsim b$ and $b \succsim a$. A *profile* V over A is a set of votes over A . For notational convenience, we give voters names like v or i , and refer to their preference relation by \succsim_v and \succsim_i .

Let V be a profile over A . We say that V is *d-Euclidean* (where $d \geq 1$) if there exists a map $x : V \cup A \rightarrow \mathbb{R}^d$ satisfying

$$a \succ_v b \implies \|x(v) - x(a)\| < \|x(v) - x(b)\| \quad \text{for all } v \in V \text{ and } a, b \in A. \quad (1)$$

Thus, voter v prefers those alternatives which are closer to v according to the embedding x . Here, $\|\cdot\|$ refers to the usual Euclidean ℓ_2 -norm on \mathbb{R}^d , that is

$$\|(x_1, \dots, x_d)\| = \|(x_1, \dots, x_d)\|_2 = \sqrt{x_1^2 + \dots + x_d^2}.$$

The (*open*) *ball* of radius r centred at c is $B(c, r) = \{x \in \mathbb{R}^d : \|x - c\| < r\}$.

Typically, we will consider profiles of *strict orders* in which every vote \succsim_v is antisymmetric (that is, we do not allow ties). Notice that in any d -Euclidean embedding of a profile of strict orders, no two alternatives can reside at the same point of \mathbb{R}^d .

In some applications, it makes sense to consider preferences that include indifferences (ties). An extreme case, which nevertheless finds many applications, is that of *dichotomous preferences*. A vote \succsim is *dichotomous* if there are no three alternatives $a, b, c \in A$ with $a \succ b \succ c$. Equivalently, \succsim is dichotomous if A splits into approved and non-approved alternatives. That is, we can write $A = A_1 \cup A_2$ with $A_1 \cap A_2 = \emptyset$ satisfying $a \succ b$ iff $a \in A_1$ and $b \in A_2$. We then say that the voter *approves* of the alternatives in A_1 , while the voter does not approve the alternatives in A_2 . A dichotomous vote can (and will) be specified by just giving the set A_1 of approved alternatives.

Our definition of Euclidean preferences applies to dichotomous preferences as well. Following the terminology of Elkind and Lackner (2015), we call a profile of dichotomous preferences *d-DE* (Dichotomous Euclidean) if it is d -Euclidean. In this context, the definition requires that there is an embedding $x : V \cup A \rightarrow \mathbb{R}^d$ so that for each voter $v \in V$, the set of approved alternatives of v coincides with the set of alternatives contained in some ball $B(x(v), r_v)$ centred at $x(v)$. We call a profile of dichotomous preferences *d-DUE* (Dichotomous *Uniform* Euclidean) if there is an embedding $x : V \cup A \rightarrow \mathbb{R}^d$ so that for each voter $v \in V$, the set of approved alternatives of v coincides with the alternatives contained in the *unit* ball $B(x(v), 1)$ centred at $x(v)$.

Some authors define Euclidean preferences in a subtly different way from us. One popular definition involves reversing the direction of the implication arrow in (1). Notice that under this definition, whenever a voter is equidistant between two alternatives, the voter is free to break the tie in either way. In particular, when $d \geq 2$, then *every* preference profile is “ d -Euclidean” under this definition – place all voters at the origin, and position alternatives on the unit sphere around the origin. In the area of multidimensional unfolding, embeddings of this type are said to include *degeneracies* [see, e.g., Busing et al. 2005]. Our definition circumvents this issue by just outright disallowing degeneracies.

Bogomolnaia and Laslier (2007) use another different definition: they replace the implication in (1) by an if-and-only-if. This definition is equivalent to ours for strict preferences,

but is much more restrictive for preferences including ties: Bogomolnaia and Laslier’s definition requires that whenever a voter v is indifferent between a and b , then a and b are equidistant to v . Our definition does not impose any relation on the relative distances in cases of ties.

2.2 Forbidden Substructures

Let V be a profile over the alternative set A . If we delete some alternatives, and are left with the set $A' \subseteq A$, we can obtain the restricted profile $V|_{A'}$ where every vote is restricted in the obvious way: $\succ|_{A'} := \succ \cap (A' \times A')$. Now, if V is a profile over A , and W is a profile over B , we say that V *contains* W if we can obtain W by first deleting some alternatives and voters from V , and then relabelling the remaining alternatives and reordering the remaining voters. A preference domain \mathcal{R} (that is, a set of profiles) may then be *characterised by forbidden configurations* by giving a set \mathcal{S} of profiles such that $\mathcal{R} = \{\text{profile } V : V \text{ does not contain any } W \in \mathcal{S}\}$.

We call a preference domain \mathcal{R} *hereditary* if it is closed under containment. That is, if $V \in \mathcal{R}$ and V contains W , then $W \in \mathcal{R}$. The d -Euclidean domain is hereditary for any $d \geq 1$. Note that any hereditary domain \mathcal{R} is characterised by its complement $\overline{\mathcal{R}}$ (that is, its set of counterexamples) in this way. However, a satisfying characterisation will either use a finite set of obstructions, or be otherwise highly structured.

2.3 Existential theory of the reals (ETR)

The *language* of the first-order *theory of the reals* consists of formulas using as symbols (i) a countable collection of variable symbols x_i , (ii) constant symbols 0 and 1, (iii) addition, subtraction, multiplication symbols, (iv) the equality (=) and inequality (<) symbols, (v) Boolean connectives (\vee, \wedge, \neg), (vi) universal and existential quantifiers (\forall, \exists). The *theory* of the reals consists of all true sentences in this language, interpreted using the obvious semantics (where quantifiers quantify over the real numbers \mathbb{R}).

The *existential theory of the reals* (ETR) consists of the true sentences of the form

$$\exists x_1 \in \mathbb{R} \exists x_2 \in \mathbb{R} \dots \exists x_n \in \mathbb{R} \quad F(x_1, x_2, \dots, x_n)$$

with $F(x_1, x_2, \dots, x_n)$ a quantifier-free formula in the language just defined. In other words, F is a Boolean combination of equalities and inequalities of real polynomials.

The *decision problem* of ETR is the problem of deciding whether a given sentence of the above form is true, that is whether it is a member of ETR. Schaefer (2010) introduced the complexity class $\exists\mathbb{R}$ as the class of decision problems that admit a polynomial-time many-one reduction to the decision problem of ETR. Thus, $\exists\mathbb{R}$ captures the computational complexity of the existential theory of the reals. We say that a problem A is $\exists\mathbb{R}$ -*hard* if all problems in $\exists\mathbb{R}$ reduce to A in polynomial time. We say that A is $\exists\mathbb{R}$ -*complete* if it is contained in $\exists\mathbb{R}$ and is $\exists\mathbb{R}$ -hard.

From the definition of ETR it is not even clear that the decision problem of ETR is decidable. By introducing a quantifier-elimination procedure, Tarski (1948) showed that ETR is in fact decidable. Since then, a variety of algorithmic improvement have been made over Tarski’s procedure (which does not admit an elementary time bound), and there exist algorithms with a singly-exponential time dependence in the number of variables (Grigor’ev 1988; Renegar 1992). In addition, Canny (1988) obtained the astonishing result that ETR can be solved in polynomial space. Thus $\exists\mathbb{R} \subseteq \text{PSPACE}$.

From the other direction, it is easy to see that ETR can be used to solve the propositional satisfiability problem (3SAT): we can encode that a variable x is either true or false ($x =$

$0 \vee x = 1$), and we can encode a clause like $(x_1 \vee \neg x_2 \vee x_3)$ through $(x_1 + (1 - x_2) + x_3 \geq 1)$.¹ Thus, ETR is NP-hard, and every $\exists\mathbb{R}$ -hard problem is also NP-hard. Together, we have the containments $\text{NP} \subseteq \exists\mathbb{R} \subseteq \text{PSPACE}$.

Multiple $\exists\mathbb{R}$ -complete problems are known, and many of them are questions of the form “can a given combinatorial object be geometrically represented?”. Particular examples include recognising intersection graphs of line segments in the plane (Schaefer 2010), of unit disk graphs (Kang and Müller 2012), or of unit distance graphs (Schaefer 2013). The problem of recognising d -Euclidean preference profiles falls exactly into this category: trying to find a geometric embedding that ‘explains’ a given combinatorial structure. Another $\exists\mathbb{R}$ -complete problem is a decision version of the problem of finding a Nash equilibrium of a non-cooperative game (Schaefer and Štefankovič 2015). In particular, it is $\exists\mathbb{R}$ -complete to decide whether a given 3-player game has a Nash equilibrium within a given ball (in the simplex of mixed strategies). $\exists\mathbb{R}$ -hardness also holds for various other decision problems related to Nash equilibria, and even restricted to symmetric games (Garg et al. 2015).

2.4 Arrangements of hyperplanes

Our exposition and terminology follows Kang and Müller (2012).

An (*affine*) d -*hyperplane* is a set of form $h = \{x \in \mathbb{R}^d : c^T x = b\} \subseteq \mathbb{R}^d$ for some $c \in \mathbb{R}^d$ and $b \in \mathbb{R}$. A particular example of a hyperplane, for given $p, q \in \mathbb{R}^d$, is the set $\{x \in \mathbb{R}^d : \|x - p\| = \|x - q\|\}$ of points that are equidistant to p and q ; in two dimensions, this is the *perpendicular bisector*. Any hyperplane h divides $\mathbb{R}^d \setminus h$ into two connected components, namely the half-planes $c^T x > b$ and $c^T x < b$. We can give h an *orientation* by (arbitrarily) designating one of these components as h ’s *positive side* h^+ , and the other as h ’s *negative side* h^- . We call a hyperplane with a chosen orientation an *oriented hyperplane*. An *oriented hyperplane arrangement* (h_1, \dots, h_n) is a finite ordered collection of oriented hyperplanes in \mathbb{R}^d .

Given an oriented hyperplane arrangement $\mathcal{H} = (h_1, \dots, h_n)$, we can assign to each point $x \in \mathbb{R}^d$ its *sign vector* $\sigma(x) \in \{-, 0, +\}^n$ by setting

$$\sigma(x)_i = \begin{cases} + & \text{if } x \in h_i^+ \\ 0 & \text{if } x \in h_i \\ - & \text{if } x \in h_i^- \end{cases}$$

Thus, the sign vector of x records, for each oriented hyperplane in the arrangement, on which side of the hyperplane x lies. The *combinatorial description* $\mathcal{D}(\mathcal{H})$ of \mathcal{H} is the collection of all sign vectors induced by the arrangement \mathcal{H} , that is

$$\mathcal{D}(\mathcal{H}) = \{\sigma(x) : x \in \mathbb{R}^d\}.$$

If $\mathcal{D}(\mathcal{H}) = \mathcal{D}(\mathcal{H}')$, then we say that \mathcal{H} and \mathcal{H}' are *isomorphic*.

Every connected component of $\mathbb{R}^d \setminus \mathcal{H} = \mathbb{R}^d \setminus (h_1 \cup \dots \cup h_n)$ is called a *cell* (or *chamber* or *region*). All points in the same cell have the same sign vector.

3 The Recognition Problem for Euclidean Preferences

In this section we will show that the problem of recognising d -Euclidean preferences is $\exists\mathbb{R}$ -complete for each fixed $d \geq 2$. We will do this by reducing from a problem concerning arrangements of hyperplanes. But first let us formally define the relevant decision problem, and verify that the problem is in fact contained in $\exists\mathbb{R}$.

¹Some definitions of ETR and $\exists\mathbb{R}$ do not allow use of the equality symbol, but this makes no difference up to polynomial-time transformations (Schaefer and Štefankovič 2015).

d-EUCLIDEAN

Instance: set A of alternatives, profile V of strict orders over A
Question: is V d -Euclidean?

Proposition 1. *d*-EUCLIDEAN is contained in $\exists\mathbb{R}$ for every $d \geq 1$. In particular it is contained in PSPACE.

Proof. This is almost immediate from the definition of d -Euclidean preferences. Namely, a profile is d -Euclidean if and only if there exist reals $x_{r,i} \in \mathbb{R}$ for each $r \in A \cup V$ and $i = 1, \dots, d$ such that whenever $a \succ_r b$, we have

$$\|x_v - x_a\| < \|x_v - x_b\| \iff \sum_{i=1}^d (x_{v,i} - x_{a,i})^2 < \sum_{i=1}^d (x_{v,i} - x_{b,i})^2.$$

Thus, the problem is equivalent to asking whether a system of polynomial inequalities has a solution. This system can be constructed in polynomial time, given the profile. \square

This proposition in particular shows that d -EUCLIDEAN is *decidable*, a fact that is not *a priori* obvious, and to the best of our knowledge has not been previously noted.

Our starting point in the reductions is the following problem about combinatorial descriptions of hyperplane arrangements.

d-REALISABILITY

Instance: a set $S \subseteq \{-, +\}^n$ of sign vectors with $(-, \dots, -), (+, \dots, +) \in S$
Question: is there an oriented d -hyperplane arrangement \mathcal{H} with $S \subseteq \mathcal{D}(\mathcal{H})$?

For example, $S = \{(+, +, +, +), (-, +, +, -), (-, +, -, +), (-, +, -, -), (-, -, -, +), (-, -, -, -)\}$ is 2-realised by the four red lines in Figure 1, where the red label of the line is placed on the positive side of the line.

Theorem 1 (Kang and Müller 2012). *d*-REALISABILITY is $\exists\mathbb{R}$ -complete for $d \geq 2$.

Kang and Müller establish this by a reduction from SIMPLE STRETCHABILITY, the problem of deciding whether an arrangement of pseudolines can be stretched into an isomorphic arrangement of lines. That problem is $\exists\mathbb{R}$ -complete by Mnëv's 1985 universality theorem, a deep topological result about representing semialgebraic varieties. Shor (1991) gives a direct proof of NP-hardness by a reduction from SAT.

We are now ready to prove our main result.

Theorem 2. *d*-EUCLIDEAN is $\exists\mathbb{R}$ -complete for each $d \geq 2$.

Proof. We have already seen that d -EUCLIDEAN is contained in $\exists\mathbb{R}$ (Proposition 1). We now show $\exists\mathbb{R}$ -hardness by a reduction from *d*-REALISABILITY.

Let $S \subseteq \{-, +\}^n$ be a given set of sign vectors with $(-, \dots, -), (+, \dots, +) \in S$. We construct a profile of $|S|$ votes over a total of $2n$ alternatives. Precisely, we take as alternatives the set $A = \{a_1, b_1, \dots, a_n, b_n\}$. For each $\sigma \in S$, we introduce a voter v_σ with strict order \succ_σ specified by

$$\{a_1, b_1\} \succ_\sigma \{a_2, b_2\} \succ_\sigma \dots \succ_\sigma \{a_n, b_n\} \quad \text{and} \quad \begin{cases} a_i \succ_\sigma b_i \iff \sigma_i = +, \\ b_i \succ_\sigma a_i \iff \sigma_i = -. \end{cases}$$

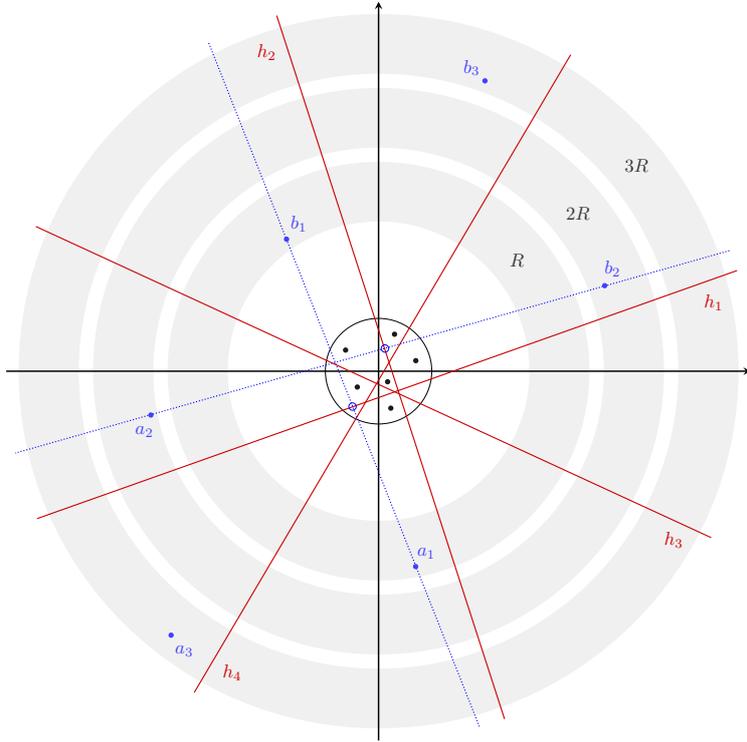


Figure 1: 2-Euclidean embedding of a profile obtained from a 2-realizable sign vector set S through the reduction of Theorem 2. Black dots represent voters v_σ , positioned within the unit ball and within the cell with sign vector σ induced by the red hyperplanes (lines). The red labels are on the positive side of each line. Blue circles denote the points p_i . Blue dots correspond to alternatives; note that a_i and b_i are at radius $Ri \pm 2$ from the origin.

This completes the description of the reduction. We now show its correctness.

Suppose the profile constructed is d -Euclidean, and let $x : V \cup A \rightarrow \mathbb{R}^d$ be a Euclidean embedding. Take the oriented hyperplane arrangement $\mathcal{H} = (h_1, \dots, h_n)$ defined by

$$\begin{aligned} h_i &= \{x \in \mathbb{R}^d : \|x - x(a_i)\| = \|x - x(b_i)\|\}, \\ h_i^+ &= \{x \in \mathbb{R}^d : \|x - x(a_i)\| < \|x - x(b_i)\|\}, \\ h_i^- &= \{x \in \mathbb{R}^d : \|x - x(a_i)\| > \|x - x(b_i)\|\}. \end{aligned}$$

Then, clearly, $S \subseteq \mathcal{D}(\mathcal{H})$: Let $\sigma \in S$ and let $i \in \{1, \dots, d\}$. If $\sigma_i = +$, we have $a_i \succ_\sigma b_i$, and thus by definition of Euclidean preferences, we must have $\|x(v_\sigma) - x(a_i)\| < \|x(v_\sigma) - x(b_i)\|$ and hence $x(v_\sigma) \in h_i^+$ so that $\sigma(x(v_\sigma))_i = + = \sigma_i$. Similarly if $\sigma_i = -$. It follows that $\sigma \in \mathcal{D}(\mathcal{H})$. Hence $S \subseteq \mathcal{D}(\mathcal{H})$.

Conversely, suppose that $S \subseteq \mathcal{D}(\mathcal{H})$ for some oriented d -hyperplane arrangement \mathcal{H} . By applying an appropriate scaling map $x \mapsto \lambda x$ if needed, we may assume that every cell of \mathcal{H} intersects the unit ball $B(0, 1) \subseteq \mathbb{R}^d$. Write $\mathcal{H} = (h_1, \dots, h_n)$ with $h_i = \{x \in \mathbb{R}^d : u_i^T x = b_i\}$, where without loss of generality $\|u_i\| = 1$, so that u_i is a unit vector. Further, we will say that $h_i^+ = \{x : u_i^T x > b_i\}$ and $h_i^- = \{x : u_i^T x < b_i\}$.

We now construct a Euclidean embedding $x : A \cup V \rightarrow \mathbb{R}^d$. We start by placing the

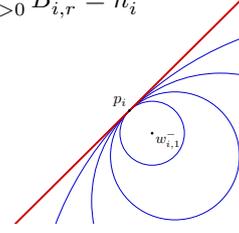
voter v_σ corresponding to $\sigma \in S$ at an arbitrary point $x(v_\sigma) \in B(0, 1)$ of the cell of \mathcal{H} with sign vector σ . This exists by our assumption that $S \subseteq \mathcal{D}(\mathcal{H})$.

Next, for each $i = 1, \dots, d$, pick some point $p_i \in h_i \cap B(0, 1)$ (this is possible because $B(0, 1)$ meets both h_i^- and h_i^+ since $(-, \dots, -), (+, \dots, +) \in S$). Following an argument by Kang and Müller (2012), we set for $r > 0$

$$\begin{aligned} w_{i,r}^+ &:= p_i + ru_i & \text{and} & & w_{i,r}^- &:= p_i - ru_i, \\ B_{i,r}^+ &:= B(w_{i,r}^+, r) & \text{and} & & B_{i,r}^- &:= B(w_{i,r}^-, r). \end{aligned}$$

Note that $B_{i,r}^+ \subseteq h_i^+$ and $B_{i,r}^- \subseteq h_i^-$. In fact, $\bigcup_{r>0} B_{i,r}^+ = h_i^+$ and $\bigcup_{r>0} B_{i,r}^- = h_i^-$ (see figure on the right). Hence, for all r sufficiently large, we have

$$\begin{aligned} x(v_\sigma) &\in B_{i,r}^+ & \text{for all } \sigma & \text{ with } \sigma_i = +, \text{ and} \\ x(v_\sigma) &\in B_{i,r}^- & \text{for all } \sigma & \text{ with } \sigma_i = -. \end{aligned}$$



Fix a value $R > 4$ of r for which this holds. We now pick the positions of the alternatives in the Euclidean embedding: Set

$$x(a_i) = w_{i,Ri}^+ \quad \text{and} \quad x(b_i) = w_{i,Ri}^-.$$

We are left to verify that the map $x : A \cup V \rightarrow \mathbb{R}^d$ thus constructed actually corresponds to voters' preferences. First let us show that, according to the embedding x , every voter's preference has the form

$$\{a_1, b_1\} \succ_\sigma \{a_2, b_2\} \succ_\sigma \dots \succ_\sigma \{a_n, b_n\}.$$

So let v_σ be a voter, let $1 \leq i < j \leq n$, and let $c_i \in \{a_i, b_i\}$ and $c_j \in \{a_j, b_j\}$. Then

$$\begin{aligned} \|x(v_\sigma) - x(c_i)\| &\leq \|x(v_\sigma) - p_i\| + \|p_i - x(c_i)\| && \text{(triangle inequality)} \\ &\leq 2 + Ri && (u_i \text{ is a unit vector)} \\ &< Rj - 2 && (j > i \text{ and } R > 4) \\ &\leq \|x(c_j) - p_j\| - \|p_j - x(v_\sigma)\| && \text{(as before)} \\ &\leq \|x(v_\sigma) - x(c_j)\|. && \text{(reverse triangle inequality)} \end{aligned}$$

Thus, it follows that $c_i \succ_\sigma c_j$, as desired. Finally, we need to confirm that

$$\|x(v_\sigma) - x(a_i)\| < \|x(v_\sigma) - x(b_i)\| \iff \sigma_i = +.$$

So suppose $\sigma_i = +$. By choice of R , we have $x(v_\sigma) \in B_{i,Ri}^+$, so that $\|x(v_\sigma) - x(a_i)\| < Ri$. On the other hand, we have $\|x(v_\sigma) - x(b_i)\| \geq Ri$: for suppose not. Then $x(v_\sigma) \in B_{i,Ri}^-$, and thus $x(v_\sigma) \in B_{i,Ri}^+ \cap B_{i,Ri}^- = \emptyset$, a contradiction. \square

Certainly, this hardness result implies that it is also hard to recognise d -Euclidean profiles of *weak orders* (since strict orders form a special case). For *dichotomous orders*, hardness does not follow immediately, but a similar reduction can be used. The decision problems for dichotomous preferences are defined as follows:

d -DICHOTOMOUS-EUCLIDEAN

Instance: set A of alternatives, profile V of dichotomous votes over A
Question: is V d -DE?

d-DICHOTOMOUS-UNIFORM-EUCLIDEAN

Instance: set A of alternatives, profile V of dichotomous votes over A
Question: is V d -DUE?

Perhaps unsurprisingly, the argument employed is almost identical to the hardness result for recognising unit disk graphs (Kang and Müller 2012).

Theorem 3. *Both d -DICHOTOMOUS-EUCLIDEAN and d -DICHOTOMOUS-UNIFORM-EUCLIDEAN are $\exists\mathbb{R}$ -complete for each $d \geq 2$.*

Proof Sketch. The proof is similar to the previous reduction. We again reduce from d -REALISABILITY. The same reduction works for both DE and DUE. Let $S \subseteq \{-, +\}^n$ be a given set of sign vectors with $(-, \dots, -), (+, \dots, +) \in S$. We construct a profile of $|S|$ dichotomous votes over the same set of alternatives $A = \{a_1, b_1, \dots, a_n, b_n\}$ as before. For each $\sigma \in S$, the voter v_σ has dichotomous preferences approving the set

$$v_\sigma = \{a_i : i = 1, \dots, n \text{ with } \sigma_i = +\} \cup \{b_i : i = 1, \dots, n \text{ with } \sigma_i = -\}.$$

This completes the description of the reduction.

Correctness can be established using almost the same argument as in (Kang and Müller 2012, Theorem 1); we omit the details here for lack of space. Compared to the argument for Theorem 2, we need to do slightly more scaling and shifting. \square

4 Precision

In this section, we consider the question of how many bits are needed to specify a Euclidean embedding $x : A \cup V \rightarrow \mathbb{R}^d$. We only consider the ‘natural encoding’ where the coordinates of each point are given as rational numbers. Note that if every d -Euclidean profile were to admit an embedding that can be specified in polynomially many bits, then this would put the problem d -EUCLIDEAN in NP. Yet in this section we show that there is a family of profiles which need exponentially many bits in order to specify any Euclidean embedding. This result, by itself, does not rule out that the *decision* problem d -EUCLIDEAN is in NP: there could be a ‘clever’ way to certify that an embedding exists, without explicitly giving the embedding (finding such a ‘clever’ certificate would prove $\text{NP} = \exists\mathbb{R}$, which in the words of Kang and Müller (2012) would constitute a “minor breakthrough in complexity theory”). On the other hand, our result shows that the *function* problem associated with the problem d -EUCLIDEAN is *provably* not in P.

Let us now make precise the notion of the *size* of an embedding $x : A \cup V \rightarrow \mathbb{R}^d$. Here, we follow the definitions of McDiarmid and Müller (2013). The number of bits needed to store a natural number $n \in \mathbb{N}$ is the number of digits in its binary representation: $\text{size}(n) := \lceil \log_2(n+1) \rceil$. To represent an integer $k \in \mathbb{Z}$, we need an extra bit to store its sign: $\text{size}(k) := 1 + \text{size}(|k|)$. Finally, we represent a rational number $q \in \mathbb{Q}$ as a pair of integers representing a fraction: if $q = m/n$, where $m, n \in \mathbb{Z}$ are relatively prime, we set $\text{size}(q) := \text{size}(m) + \text{size}(n)$. The size of a rational vector $x \in \mathbb{Q}^d$ is $\text{size}(x) := \sum_{i=1}^d \text{size}(x_i)$. Then, the size of a *rational* Euclidean embedding $x : A \cup V \rightarrow \mathbb{Q}^d$ is defined as

$$\text{size}(x) := \sum_{r \in A \cup V} \text{size}(x(r)).$$

Before we establish the promised lower bound, let us first give a corresponding upper bound. Namely, while some d -Euclidean profiles require exponentially many bits to specify,

(single-)exponentially many bits are always enough. To see this, we will first need a guarantee that every d -Euclidean profile admits a rational embedding, because we have only assigned sizes to rational embeddings.

Theorem 4. *Every d -Euclidean profile admits a rational embedding. Further, for each $d \geq 1$, there is a constant $c = c(d)$ such that any d -Euclidean profile with n voters and m alternatives admits a rational embedding $x : A \cup V \rightarrow \mathbb{Q}^d$ with $\text{size}(x) \leq 2^{c(n+m)}$.*

This theorem is an essentially immediate corollary of the following general result about the bit sizes of solutions to polynomial inequalities.

Theorem 5 (Basu et al. 1996). *Fix $d, \tau \in \mathbb{N}$. There is a constant $C = C(d, \tau)$ such that for all sets \mathcal{P} of polynomials in n variables of degree at most d and with integer coefficients of bit size at most τ , we have that whenever the system $\{p(x) > 0 : p \in \mathcal{P}\}$ has a real solution, then it has a rational solution of bit size at most $|\mathcal{P}| \cdot \tau \cdot d^{Cn}$.*

Proof of Theorem 4. From Proposition 1, we know that d -Euclidean embeddings are precisely the solutions to a certain system of strict polynomial inequalities, in which all polynomials have degree 2, and all coefficients are at most 2 in absolute value. So we can apply Theorem 5 to obtain our result. \square

The upper bound of Theorem 4 is not tight for the case $d = 1$. Recall from Appendix A that there are polynomial-time algorithms for recognising the 1-Euclidean domain. These work through a (non-trivial) reduction to linear programming. The linear programs produced in this reduction have polynomially bounded coefficients (in fact, bounded by 2), and are thus *combinatorial* linear programs, and so admit a *strongly polynomial* algorithm (Tardos 1986). Thus, we can say the following:

Proposition 2. *Every 1-Euclidean profile with n voters and m alternatives admits a rational embedding $x : A \cup V \rightarrow \mathbb{Q}$ with $\text{size}(x)$ bounded by $\text{poly}(n, m)$.*

For the lower bound, we use techniques developed by McDiarmid and Müller (2013) and Kang and Müller (2012) and apply them to the reduction of Theorem 2. For a profile V over alternative set A , we define $\|V\| := |V| + |A|$.

Theorem 6. *Fix $d \geq 2$. For a d -Euclidean preference profile V , let $e(V)$ denote the minimum size of a rational Euclidean embedding of V . For each $m \geq 1$, let $e(n + m)$ be the maximum $e(V)$ among d -Euclidean preference profiles V with $\|V\| = n + m$. Then $e(n + m) \geq 2^{\Omega(n+m)}$.*

Proof Sketch. McDiarmid and Müller (2013) construct a family of combinatorial descriptions of line arrangements that have doubly-exponential ‘*span*’, which (roughly) corresponds to the size of the numbers needed to represent any realisation of the combinatorial description. Kang and Müller (2012) generalise this construction to d -dimensional hyperplane arrangements. They then transform these line arrangements into unit disk graphs in the proof of their Theorem 3. That proof goes through almost verbatim for our case (using the reduction of our Theorem 2). At one step in the proof, Kang and Müller need to introduce isolated vertices to pad the graph in question. In our setting, we may introduce new alternatives and add them to the bottom of each voter’s preference list; in a Euclidean embedding we just place these alternatives far away from all the voters. \square

5 Forbidden Minor Characterisations

Consider a possible characterisation of the d -Euclidean domain by a set \mathcal{S} of forbidden configurations. We will call this characterisation *good* if the set \mathcal{S} is polynomial-time recognisable:

that is, given a profile, there should be a polynomial-time algorithm deciding whether the given profile is one of the configurations contained in \mathcal{S} . Certainly, if \mathcal{S} were *finite*, then \mathcal{S} provides a good characterisation. However, there exist infinite characterisations that are still good in this sense, for example for interval graphs (Lekkerkerker and Boland 1962) and matrices with the consecutive ones property (Tucker 1972). (It could be argued that in order for \mathcal{S} to be good it needs to be recognisable in LOGSPACE or another complexity class below P.)

However, given the complexity result of Section 3, it is a straightforward observation that for each $d \geq 2$, no good characterisation by forbidden substructures will exist for the d -Euclidean domain, subject to a reasonable complexity-theoretic assumption.

Proposition 3. *For each $d \geq 2$, the set of d -Euclidean preference profiles does not admit a good characterisation by forbidden substructures unless $\exists\mathbb{R} \subseteq \text{coNP}$.*

Note that $\exists\mathbb{R} \subseteq \text{coNP}$ would imply $\text{NP} \subseteq \text{coNP}$, itself a rather unlikely event.

Proof. Suppose a good characterisation by \mathcal{S} exists. We give a coNP-algorithm that recognises d -Euclidean preferences: Given an input profile, guess some subprofile, guess a relabeling of voters and alternatives in this subprofile, and check whether the result is contained in \mathcal{S} . \square

By a similar argument, no *finite* characterisation can exist unless $\text{P} = \exists\mathbb{R} = \text{NP}$. In the remainder of this section, we prove this weaker result *without* appealing to any complexity-theoretic assumptions. To do this, we use a connection between the theory of arrangements of hyperplanes and the theory of *ordered matroids*.

Theorem 7 (Bokowski and Sturmfels 1989). *There exist infinitely many nonrealisable uniform oriented matroids of rank 3 such that every proper minor of them is realisable. In particular, by the Topological Realisation Theorem, for every $n_0 \in \mathbb{N}$, there exists a non-stretchable simple pseudoline arrangement with $n > n_0$ lines such that removing any line results in a stretchable arrangement.*

To prove our result about d -Euclidean preferences, we will use the examples from Theorem 7 and apply to them the chain of many-one reductions that yielded the hardness result of Theorem 2. This argument is presented in several lemmas in the appendix.

Theorem 8. *The domain of d -Euclidean preferences does not admit a finite characterisation by forbidden configurations, for any fixed $d \geq 1$.*

6 Other metrics

In Section 2.1, we defined Euclidean preferences using the usual Euclidean ℓ_2 -metric, measuring distances by shortest paths in the plane. Other choices of metric may be preferred in certain contexts, and in this section we will briefly consider the effect of using other metrics on the complexity of the recognition problem.

The two metrics we consider here are the ℓ_1 -metric and the ℓ_∞ -metric. The ℓ_1 -metric is also known as the *cityblock* or *taxicab* or *Manhattan* distance, because it measures distances by shortest paths on a grid like the street network of Manhattan. Formally, the ℓ_1 -norm is defined by

$$\|(x_1, \dots, x_d)\|_1 := |x_1| + \dots + |x_d|.$$

Thus, the ℓ_1 -distance $\|x - y\|_1$ of two points x and y is the sum of the absolute distances along each coordinate axis. The ℓ_∞ -metric, on the other hand, measures the *maximum*

distance along a coordinate axis:

$$\|(x_1, \dots, x_d)\|_\infty := \max\{|x_1|, \dots, |x_d|\}.$$

For each of these metrics (or indeed, any metric space), we can obtain a notion of d -Euclidean preferences by just plugging this metric into the definition in line (1). For certain settings, the ℓ_∞ metric has a nice interpretation as corresponding to ‘pessimistic’ voters who judge candidate according to their (subjectively) worst feature. The ℓ_1 metric also has intuitive appeal – see Eguia (2011) and the references therein for arguments in favour of using Euclidean preferences with respect to this metric.

Comparing the ℓ_1 - and ℓ_∞ -metrics to the ℓ_2 -metric we have used so far, one gets the sense that ℓ_1 and ℓ_∞ are more ‘discrete’ or ‘combinatorial’ than the more geometric ℓ_2 . Supporting this intuition, we find that the complexity of the recognition problem changes (unless $\text{NP} = \exists\mathbb{R}$) when we use one of these metrics (proof in the appendix).

Theorem 9. *The problems of recognising preference profiles that are d -Euclidean with respect to the ℓ_1 -metric or the ℓ_∞ -metric are contained in NP for every $d \geq 1$.*

7 Conclusions

The results of this paper are bad news for the d -Euclidean domain: because producing a Euclidean embedding will in general be infeasible, we are stuck with heuristic algorithms that may or may not produce a correct output in their allotted time. In some sense, our hardness results show that the estimation algorithms developed for the multidimensional unfolding problem over the past several decades are best possible, in the sense that we cannot hope for exact efficient algorithms. Still, future developments in ETR-solver technology might allow solving practical instances in reasonable time, and perhaps some of the ideas in the area of multidimensional unfolding can be formalised and yield exact algorithms. We have run some preliminary experiments on the PrefLib dataset (Mattei and Walsh 2013) using the `nlsat` solver (Jovanović and De Moura 2012) which is part of the `z3` theorem prover (De Moura and Bjørner 2008), and appears to be the strongest ETR-solver available. However, `nlsat` was unable to decide whether *any* of the PrefLib profiles that we tried was 2- or 3-Euclidean within a time bound of one hour (except for trivial profiles on 3 alternatives).

The general infeasibility of identifying membership in the d -Euclidean domain ($d \geq 2$) also means that any efficient algorithm developed for actual voting problems that exploits the spatial structure will need to be given a Euclidean embedding as part of the input. Except perhaps for facility-location type problems, not too many examples come to mind in which the underlying spatial structure is known *a priori*. Thus, such algorithms may turn out to be of limited use.

While the multidimensional case seems nasty, there is hope that we will be able to develop a better understanding of the one-dimensional Euclidean domain in the future. We would like to reiterate here two open problems posed elsewhere in the literature: Elkind and Faliszewski (2014) ask whether the 1-Euclidean domain can be recognised by a ‘combinatorial’ algorithm that does not rely on solving a linear program, and Chen et al. (2015) ask whether there is an explicit and good characterisation of the 1-Euclidean domain by (infinitely many) forbidden configurations. It appears likely that the answer to both of these questions is the same, and a positive answer would require a better structural understanding of the 1-Euclidean domain.

Finally, the versions of the problem we discussed in Section 6 for ℓ_1 and ℓ_∞ are intriguing: What is the precise complexity of the recognition problem? How are these notions related to multidimensional single-peakedness?

Acknowledgements

I thank Edith Elkind and Martin Lackner for helpful discussions, and Jürgen Bokowski for correspondence about the theory of oriented matroids.

References

- E. Anshelevich, O. Bhardwaj, and J. Postl. Approximating optimal social choice under metric preferences. In *AAAI '15*, pages 777–783, 2015.
- M. A. Ballester and G. Haeringer. A characterization of the single-peaked domain. *Social Choice and Welfare*, 36(2):305–322, 2011.
- S. Barberà, F. Gul, and E. Stacchetti. Generalized median voter schemes and committees. *Journal of Economic Theory*, 61(2):262–289, 1993.
- J. Bartholdi and M. A. Trick. Stable matching with preferences derived from a psychological model. *Operations Research Letters*, 5(4):165–169, 1986.
- S. Basu, R. Pollack, and M.-F. Roy. On the combinatorial and algebraic complexity of quantifier elimination. *Journal of the ACM*, 43(6):1002–1045, 1996.
- J. F. Bennett and W. L. Hays. Multidimensional unfolding: Determining the dimensionality of ranked preference data. *Psychometrika*, 25(1):27–43, 1960.
- N. Betzler, A. Slinko, and J. Uhlmann. On the computation of fully proportional representation. *Journal of Artificial Intelligence Research*, 47(1):475–519, 2013.
- A. Bogomolnaia and J.-F. Laslier. Euclidean preferences. *Journal of Mathematical Economics*, 43(2):87–98, 2007.
- J. Bokowski and B. Sturmfels. An infinite family of minor-minimal nonrealizable 3-chirotopes. *Mathematische Zeitschrift*, 200(4):583–589, 1989.
- I. Borg and P. J. Groenen. *Modern multidimensional scaling: Theory and applications*. Springer Science & Business Media, 2005.
- F. Brandt, M. Brill, E. Hemaspaandra, and L. A. Hemaspaandra. Bypassing combinatorial protections: Polynomial-time algorithms for single-peaked electorates. In *AAAI '10*, 2010.
- R. Bredereck, J. Chen, and G. J. Woeginger. Are there any nicely structured preference profiles nearby? In *IJCAI '13*, pages 62–68, Aug. 2013a.
- R. Bredereck, J. Chen, and G. J. Woeginger. A characterization of the single-crossing domain. *Social Choice and Welfare*, 41(4):989–998, 2013b.
- M.-L. Bruner and M. Lackner. On the likelihood of single-peaked preferences. Technical Report arXiv:1505.05852 [cs.GT], arXiv.org, May 2015.
- F. M. Busing, P. J. Groenen, and W. J. Heiser. Avoiding degeneracy in multidimensional unfolding by penalizing on the coefficient of variation. *Psychometrika*, 70(1):71–98, 2005.
- J. Canny. Some algebraic and geometric computations in pspace. In *STOC '88*, pages 460–467. ACM, 1988.
- J. Chen, K. Pruhs, and G. J. Woeginger. The one-dimensional euclidean domain: Finitely many obstructions are not enough. Technical Report arXiv:1506.03838 [cs.GT], arXiv.org, 2015.
- C. H. Coombs. Psychological scaling without a unit of measurement. *Psychological review*, 57(3):145, 1950.

- C. H. Coombs. *A Theory of Data*. John Wiley & Sons, 1964.
- D. Cornaz, L. Galand, and O. Spanjaard. Bounded single-peaked width and proportional representation. In *ECAI '12*, pages 270–275, 2012.
- L. De Moura and N. Bjørner. Z3: An efficient smt solver. In *Tools and Algorithms for the Construction and Analysis of Systems*, pages 337–340. Springer, 2008.
- J.-P. Doignon and J.-C. Falmagne. A polynomial time algorithm for unidimensional unfolding representations. *Journal of Algorithms*, 16(2):218–233, 1994.
- J. X. Eguia. Foundations of spatial preferences. *Journal of Mathematical Economics*, 47(2): 200–205, 2011.
- E. Elkind and P. Faliszewski. Recognizing 1-euclidean preferences: An alternative approach. In *Algorithmic Game Theory*, pages 146–157. Springer, 2014.
- E. Elkind and M. Lackner. On detecting nearly structured preference profiles. In *AAAI '14*, 2014.
- E. Elkind and M. Lackner. Structure in dichotomous preferences. In *IJCAI '15*, pages 2019–2025, 2015. ISBN 978-1-57735-738-4.
- E. Elkind, P. Faliszewski, and A. Slinko. Clone structures in voters' preferences. In *EC '12*, pages 496–513. ACM, 2012.
- E. Elkind, P. Faliszewski, and P. Skowron. A characterization of the single-peaked single-crossing domain. In *AAAI '14*, pages 654–660, 2014.
- G. Erdélyi, M. Lackner, and A. Pfandler. The complexity of nearly single-peaked consistency. In *AAAI '13*, pages 283–289, July 2013.
- B. Escoffier, J. Lang, and M. Öztürk. Single-peaked consistency and its complexity. In *ECAI '08*, volume 8, pages 366–370, 2008.
- P. Faliszewski, E. Hemaspaandra, L. A. Hemaspaandra, and J. Rothe. The shield that never was: Societies with single-peaked preferences are more open to manipulation and control. In *TARK '09*, pages 118–127. ACM, 2009.
- J. Garg, R. Mehta, V. V. Vazirani, and S. Yazdanbod. Etr-completeness for decision versions of multi-player (symmetric) nash equilibria. In *Automata, Languages, and Programming*, pages 554–566. Springer, 2015.
- J.-M. Grandmont. Intermediate preferences and the majority rule. *Econometrica*, 46(2): 317–30, 1978.
- D. Y. Grigor'ev. Complexity of deciding tarski algebra. *Journal of Symbolic Computation*, 5(1):65–108, 1988.
- W. L. Hays and J. F. Bennett. Multidimensional unfolding: Determining configuration from complete rank order preference data. *Psychometrika*, 26(2):221–238, 1961.
- H. Hotelling. Stability in competition. *The Economic Journal*, 39(153):41–57, 1929.
- D. Jovanović and L. De Moura. Solving non-linear arithmetic. In *Automated Reasoning*, pages 339–354. Springer, 2012.
- H. Kamiya, P. Orlik, A. Takemura, and H. Terao. Arrangements and ranking patterns. *Annals of Combinatorics*, 10(2):219–235, 2006.
- H. Kamiya, A. Takemura, and H. Terao. Ranking patterns of unfolding models of codimension one. *Advances in Applied Mathematics*, 47(2):379–400, 2011.

- R. J. Kang and T. Müller. Sphere and dot product representations of graphs. *Discrete & Computational Geometry*, 47(3):548–568, 2012.
- V. Knoblauch. Recognizing one-dimensional euclidean preference profiles. *Journal of Mathematical Economics*, 46(1):1–5, 2010.
- J. B. Kruskal and J. D. Carroll. Geometrical models and badness-of-fit functions. *Multivariate Analysis*, 2:639–671, 1969.
- C. G. Lekkerkerker and J. Boland. Representation of a finite graph by a set of intervals on the real line. *Fundamenta Mathematicae*, 51(1):45–64, 1962.
- N. Mattei and T. Walsh. Preflib: A library for preferences <http://www.preflib.org>. In *Algorithmic Decision Theory*, pages 259–270. Springer, 2013.
- C. McDiarmid and T. Müller. Integer realizations of disk and segment graphs. *Journal of Combinatorial Theory, Series B*, 103(1):114–143, 2013.
- S. Merrill and B. Grofman. *A unified theory of voting: Directional and proximity spatial models*. Cambridge University Press, 1999.
- N. E. Mnëv. Realizability of combinatorial types of convex polyhedra over fields. *Journal of Soviet Mathematics*, 28(4):606–609, 1985.
- D. Peters and E. Elkind. Preferences single-peaked on nice trees. In *AAAI '16*, 2016.
- J. Renegar. On the computational complexity and geometry of the first-order theory of the reals. part i: Introduction. preliminaries. the geometry of semi-algebraic sets. the decision problem for the existential theory of the reals. *Journal of Symbolic Computation*, 13(3):255–299, 1992.
- E. E. C. I. Roskam. *Metric analysis of ordinal data in psychology*. PhD thesis, 1968.
- M. Schaefer. Complexity of some geometric and topological problems. In *Graph Drawing: 17th International Symposium*, volume 5849, pages 334–344. Springer, 2010.
- M. Schaefer. Realizability of graphs and linkages. In *Thirty Essays on Geometric Graph Theory*, pages 461–482. Springer, 2013.
- M. Schaefer and D. Štefankovič. Fixed points, nash equilibria, and the existential theory of the reals. *Theory of Computing Systems*, pages 1–22, 2015. ISSN 1433-0490.
- P. Shor. Stretchability of pseudolines is np-hard. *Applied Geometry and Discrete Mathematics, Amer. Math. Soc., Providence, RI*, 4:531–554, 1991.
- X. Sui, A. Francois-Nienaber, and C. Boutilier. Multi-dimensional single-peaked consistency and its approximations. In *IJCAI '13*, pages 375–382, 2013.
- Y. Takane, F. W. Young, and J. De Leeuw. Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features. *Psychometrika*, 42(1):7–67, 1977.
- E. Tardos. A strongly polynomial algorithm to solve combinatorial linear programs. *Operations Research*, 34(2):250–256, 1986.
- A. Tarski. *A decision method for elementary algebra and geometry*. Rand report. Rand Corporation, 1948.
- M. A. Trick. Recognizing single-peaked preferences on a tree. *Mathematical Social Sciences*, 17(3):329–334, 1989.
- A. Tucker. A structure theorem for the consecutive 1's property. *Journal of Combinatorial Theory, Series B*, 12(2):153–162, 1972.

A Related Work

In this section, we survey work connected to the topics of this paper.

In the psychometrics literature, the problem of detecting d -Euclidean preferences is known as *multidimensional unfolding* (Hays and Bennett 1961; Bennett and Hays 1960), a notion based on Coomb’s 1950; 1964 *unidimensional unfolding* (which in our terminology corresponds to 1-Euclidean preferences). Many ways of solving the multidimensional unfolding problem have been proposed [e.g., Roskam 1968; Kruskal and Carroll 1969; Takane et al. 1977] which are based on iteratively fitting the data by optimising some badness-of-fit function. These methods face the problem of producing *degenerate* embeddings in which voters are placed equidistantly to multiple alternatives. Some methods have attempted to partly remedy this problem [see, e.g., Busing et al. 2005 and the references therein]. However, the papers in this area do not take a formal approach and typically do not establish formal guarantees concerning correctness and runtime bounds.

The domain of *one*-dimensional Euclidean preferences exhibits a lot of structure and has many useful theoretical properties. In particular, every 1-Euclidean profile is both single-peaked and single-crossing (Grandmont 1978). In fact, in every 1-Euclidean embedding, the order of the alternatives forms a single-peaked axis, and the order of the voters forms a single-crossing order. Being a (strictly) stronger condition than both single-peakedness and single-crossingness [see, e.g., Elkind et al. 2014], the 1-Euclidean preference domain inherits all of their good properties: 1-Euclidean profiles admit a transitive majority relation (and thus Condorcet winners), have the representative voter property, admit a non-manipulable voting rule, and a large number of important hard computational problems become tractable on this domain.

The combinatorial substructure of the 1-Euclidean domain has allowed researchers to devise polynomial-time algorithms that recognise this domain. The first such algorithm appears to have been found by Doignon and Falmagne (1994) in the context of unidimensional unfolding, and subsequently rediscovered by Elkind and Faliszewski (2014). This algorithm proceeds by finding a single-crossing ordering of the voters (which is unique up to reversal), then discovers a suitable ordering of the alternatives, and finally uses Linear Programming to check for an embedding into the real line making the input profile 1-Euclidean. An alternative approach is described by Knoblauch (2010). Knoblauch’s algorithm exploits single-peakedness rather than the single-crossing condition: the algorithm first identifies a suitable single-peaked axis, and then again uses Linear Programming to produce the real embedding.

The *Euclidean dimension* of a preference profile is the smallest integer d such that the profile is d -Euclidean. Any preference profile over m alternatives is $(m-1)$ -Euclidean: to see this, place alternatives on the vertices of the $(m-1)$ -simplex in \mathbb{R}^{m-1} ; then any voter can be placed appropriately within the simplex to reflect the voter’s preferences. Bogomolnaia and Laslier (2007) study the question of which dimension is sufficient to represent all preference profiles of a certain size. They show that the dimension that suffices to embed all profiles with n voters and m alternatives is between $\min\{n-1, m-1\}$ and $\min\{n, m-1\}$. They also show that the Condorcet cycle on k alternatives has Euclidean dimension $k-1$.

Kamiya et al. (2006, 2011) study the question of how many different voter-maximal d -Euclidean preference profiles there are for each fixed number m of alternatives. They introduce and use ‘mid-hyperplane arrangements’ to study this question, and answer it for the cases $d=1$ and $d=m-2$. Kamiya et al. found the same problem for other d to be “quite difficult at this stage”. Our complexity result gives an indication that d -Euclidean preferences are much more chaotic when $d \geq 2$, giving an indication why their counting problem may be difficult to solve for $d \geq 2$.

Every 1-Euclidean profile is also single-peaked. Similarly, there seems to be a connection

between profiles that are multidimensional single-peaked (Barberà et al. 1993; Sui et al. 2013) and ones that are multidimensional Euclidean, though we are not aware of a formal result to that effect. In recent years, there has been a lot of interest in profiles that are “almost structured” (Erdélyi et al. 2013; Cornaz et al. 2012; Bredereck et al. 2013a; Elkind and Lackner 2014). To formalise the notion of closeness, a wide variety of metrics have been proposed (such as finding the minimum number of voters that need to be removed to obtain a structured profile, or finding a way of partitioning the profile into structured subprofiles). Let us note here that the Euclidean dimension of a profile could also be viewed as such a metric – the smaller the dimension of the profile, the more structured it is. Of course, it follows from the results of this paper that it is hard to evaluate this metric. This hardness phenomenon has also been found for most other metrics [see the papers cited previously].

B Missing Proofs

B.1 Forbidden Minor Characterisations

Lemma 1. *Fix $d \geq 2$. For every $n_0 \in \mathbb{N}$, there exists $n > n_0$ and a set $S \subseteq \{-, +\}^n$ with $(+, \dots, +), (-, \dots, -) \in S$ such that S is not d -realisable, but for each $i = 1, \dots, n$, the set $S_{-i} = \{s_{-i} : s \in S\}$ obtained by deleting coordinate i is d -realisable.*

Proof. For $d = 2$, this is a direct consequence of Theorem 7 after applying the relabelling procedure described in the proof of Theorem 10 of (Kang and Müller 2012) – relabelling is necessary to ensure that $(+, \dots, +), (-, \dots, -) \in S$.

We are left to show the result for $d > 2$, and we proceed by induction. For this, we will need Lemma 11 of (Kang and Müller 2012) which states that a set $S \subseteq \{-, +\}^n$ with $(+, \dots, +), (-, \dots, -) \in S$ is d -realisable if and only if $S \times \{+, -\}$ is $(d + 1)$ -realisable. Now, let $n_0 \in \mathbb{N}$ be given, and use the inductive hypothesis to find $n \geq n_0 + 2$ and $S \subseteq \{-, +\}^n$ such that S is not d -realisable, but deleting any coordinate yields a d -realisable set. Now consider $S' := S \times \{+, -\}$. By the result quoted, S' is not $(d + 1)$ -realisable, but for $i = 1, \dots, n$, the sets S'_{-i} are $(d + 1)$ -realisable. If $S'_{-(n+1)}$ ($= S$) also happens to be $(d + 1)$ -realisable, then we are done, since in this case S' is minimally non- $(d + 1)$ -realisable. So suppose that $S'_{-(n+1)}$ is *not* $(d + 1)$ -realisable. In this case, $S'_{-(n+1)}$ is minimally non- $(d + 1)$ -realisable, since deleting any further coordinate leaves a minor of one of the S'_{-i} which we know to be $(d + 1)$ -realisable. In either case, we have found a minimal counterexample of size greater than n_0 , as required. \square

Lemma 2. *Fix $d \geq 2$. For every $m_0 \in \mathbb{N}$, there is $m > m_0$ such that there exists a preference profile over m alternatives which is not d -Euclidean, yet removing any alternative yields a d -Euclidean profile.*

Proof. Let $m_0 \in \mathbb{N}$ be given, and find $m' > m_0$ and $S \subseteq \{+, -\}^{m'}$ satisfying the conditions of Lemma 1. From S , construct the preference profile V over alternative set $\{a_1, b_1, \dots, a_{m'}, b_{m'}\}$ with $2m' > m_0$ alternatives as in the proof of Theorem 2.

According to the proof of Theorem 2, V cannot be d -Euclidean since S is not d -realisable. The profile V' obtained from the d -realisable set S_{-i} , however, *is* d -Euclidean: From the definition of the reduction it is clear that V' is just V with the alternatives a_i and b_i removed. Suppose only one of these is removed: we need to argue that the resulting profile is also d -Euclidean. Without loss of generality, we remove a_i from the profile V to obtain profile V'' . Take the d -Euclidean embedding of V' that is produced in the proof of Theorem 2, and place alternative b_i at any point $x(b_i)$ that is distance Ri away from the origin. Just like in the proof of Theorem 2, we can see that this embedding makes V'' d -Euclidean. \square

It is worth noting that all these infinitely many minimal counterexamples have the shape $\{a_1, b_1\} \succ \{a_2, b_2\} \succ \dots \succ \{a_n, b_n\}$, and in particular they are single-peaked.

Theorem 10. *The domain of d -Euclidean preferences does not admit a finite characterisation by forbidden configurations, for any fixed $d \geq 1$.*

Proof. For $d = 1$, this is the main result of Chen et al. (2015). For fixed $d \geq 2$, suppose for a contradiction that such a characterisation exists, and let M be the maximum number of alternatives in any of the forbidden configurations. By Lemma 2, there exists a profile V over at least $M + 1$ alternatives which is not d -Euclidean. Since the forbidden configurations characterise the d -Euclidean domain, one of the configurations must be contained in V . In fact, considering the size of the configuration, it must be contained in V even with one alternative deleted. However this profile *is* d -Euclidean, contradicting the fact that the d -Euclidean domain is hereditary. \square

B.2 Other metrics

Proof of Theorem 9. We start with the ℓ_1 -metric and show containment in NP by giving a nondeterministic reduction to linear programming. For each of the d coordinate axes of \mathbb{R}^d , nondeterministically guess in which order the points corresponding to voters and alternatives appear along that axis. Once we have decided these orderings, we can rewrite the definition of ℓ_1 -Euclidean preferences without using absolute values. Then, we can replace strict inequalities with weak inequalities by introducing additive ‘slack’ constants (Elkind and Faliszewski 2014, Prop. 3). The result is a linear (feasibility) program, which produces a suitable ℓ_1 -Euclidean embedding if one exists.

$$d_{v,c,i} = \pm(x_{v,i} - x_{c,i}) \quad (\text{distance between } v \text{ and } c \text{ along } i) \quad (2)$$

$$d_{v,c} = \sum_{i=1}^d d_{v,c,i} \quad (\text{distance between } v \text{ and } c) \quad (3)$$

$$d_{v,a} \leq d_{v,b} - 1 \quad \text{when } a \succ_v b \quad (4)$$

$$x_{y,i} \leq x_{y',i} - 1 \quad \text{when } y \text{ occurs to the left of } y' \text{ on axis } i \quad (5)$$

In constraints of form (2), the \pm can be replaced by plus or minus at ‘compile’-time so that the quantity reflects $|x_{v,i} - x_{c,i}|$.

The argument for the ℓ_∞ -metric is similar: here we additionally guess for each pair (v, c) in which direction the maximum distance is achieved. \square

Dominik Peters
 Department of Computer Science
 University of Oxford
 Oxford, UK
 Email: dominik.peters@cs.ox.ac.uk