On the number of realizations of certain Henneberg graphs arising in protein conformation[☆]

Leo Liberti[a], Benoît Masson[b], Jon Lee[c], Carlile Lavor[d], Antonio Mucherino[b]

[a] *LIX, École Polytechnique, 91128 Palaiseau, France*
[b] *IRISA, University of Rennes 1, Rennes, France*
[c] *IOE, University of Michigan, Ann Arbor (MI), USA*
[d] *Department of Applied Mathematics (IMECC-UNICAMP), University of Campinas, 13081-970, Campinas - SP, Brazil*

Abstract

Several application fields require finding Euclidean coordinates consistent with a set of distances. More precisely, given a simple undirected edge-weighted graph, we wish to find a realization in a Euclidean space so that adjacent vertices are placed at a distance which is equal to the corresponding edge weight. Realizations of a graph can be either flexible or rigid. In certain cases, rigidity can be seen as a property of the graph rather than the realization. In the last decade, several advances have been made in graph rigidity, but most of these, for applicative reasons, focus on graphs having a unique realization. In this paper we consider a particular type of weighted Henneberg graphs that model protein backbones and show that almost all of them give rise to sets of incongruent realizations whose cardinality is a power of two.

Keywords: Distance geometry, graph rigidity, Branch-and-Prune, partial reflection, protein conformation.

1. Introduction

The fundamental problem of Distance Geometry (DG) is that of determining the Euclidean coordinates corresponding to a given set of distances [9]. By "distances" we mean here a set E of unordered pairs $\{u, v\}$ of vertices of a set V together with a function $d : E \to \mathbb{R}_+$ mapping each pair $\{u, v\}$ to a distance value d_{uv} between u and v. We thus formalize this problem as the

[☆] This paper (significantly) extends the conference paper [30]. The first author was partially supported by the Microsoft-CNRS Chair "OSD" and by the ANR "Bip:Bip" project; the fourth author was partially supported by FAPESP and CNPq.

Email addresses: liberti@lix.polytechnique.fr (Leo Liberti), benoit.masson@inria.fr (Benoît Masson), jonxlee@umich.edu (Jon Lee), clavor@ime.unicamp.br (Carlile Lavor), antonio.mucherino@irisa.fr (Antonio Mucherino)

DISTANCE GEOMETRY PROBLEM (DGP). Given a positive integer K and a weighted simple undirected graph $G = (V, E, d)$, where $d : E \to \mathbb{R}_+$, find a function $x : V \to \mathbb{R}^K$ such that

$$\forall \{u, v\} \in E \quad \|x_u - x_v\|_2 = d_{uv}. \tag{1}$$

A vertex function $x : V \to \mathbb{R}^K$ is a *realization*. A realization is *valid* if it satisfies Eq. (1). The DGP, also known as the EMBEDDABILITY problem, is strongly **NP**-hard even when $K = 1$ and the edge weights are restricted to take values in $\{1, 2\}$ [42]. It was shown in [1] that all graphs having a realization in some Euclidean space are realizable when $K \geq \frac{1}{2}(\sqrt{8|E| + 1} - 1)$, so, in a sense, the smaller the dimension, the harder the problem.

Historically, DG was first studied by Menger in the 1930s [31] and then fully discussed and extended by Blumenthal [4]. At that time, the main notion was that of a *distance space* (S, d), where S is a set and $d : S \times S \to \mathbb{R}$ is a distance function. The "fundamental problem" (also called *subset problem* by Blumenthal) was that of determining necessary and sufficient conditions for given weighted sets (S, d) to be distance spaces. In other words, given a square $n \times n$ matrix $D = (d_{ij})$, determine whether it is a *distance matrix*, i.e. a matrix for which there exists a realization x in K dimensions such that $d_{ij} = \|x_i - x_j\|$ for all $i, j \in S$. Schoenberg [44] proved in 1935 that this is the case if and only if the matrix $-V^\top DV$ is Positive Semi-Definite (PSD), where V is the $n \times (n-1)$ matrix:

$$\frac{1}{\sqrt{2}} \begin{pmatrix} -1 & -1 & \dots & -1 \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

We remark that Schoenberg used another notation; our notation is taken from [10, Sect. 5.4.1]. Since a matrix can be ascertained to be PSD in polynomial time, Blumenthal's subset problem is in **P**.

A finite simple weighted undirected graph may have uncountably many or finitely many realizations. In this paper we are interested in rigid graphs, which have finitely many realizations. Graph rigidity was discussed time and again in connection with several different application fields, the main ones being statics, molecular biology, robotics and localization of wireless sensor networks. Historically, this generated several slightly different definitions of the concept of graph rigidity. An advanced but didactical account can be found in [14].

We consider the case where the graph G has a vertex order such that each $(K+1)$-tuplet of consecutive vertices of V induces a $(K+1)$-clique as a subgraph of G (see Fig. 1). The class of DGP instances with such an order, collectively known as the generalized DISCRETIZABLE MOLECULAR DISTANCE GEOMETRY PROBLEM ($^{\mathsf{K}}$DMDGP), is also **NP**-hard when $K = 3$ [22]. The $^{\mathsf{K}}$DMDGP order is a particular type of Henneberg order [46], and therefore characterizes a rigid graph. However, in most cases the graph is not uniquely rigid. In this paper

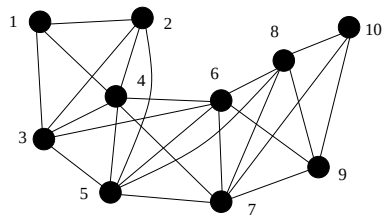


Figure 1: A $^{\mathsf{K}}$DMDGP graph with $K = 3$.

we prove that, for almost all edge weight functions, $^{\mathsf{K}}$DMDGP graphs have a number of incongruent realizations which is a power of two. As Hendrickson pointed out in [16, Sect. 3], nonuniquely rigid graphs are not well studied. This might be because one of the most important applications of DG (the localization of sensor networks [12]) requires the input data to be dense enough so that the graph is uniquely rigid [3]. More importantly, Hendrickson also pointed out that the key to non-unique rigidity is partial reflection of the realization: we shall make use of this concept to derive our result.

Our motivation for studying rigid graphs which are not uniquely rigid arises from the conformation of proteins. In general, a molecule can be seen as a unit sphere graph [19] defined by a distance threshold given by the resolution scope of Nuclear Magnetic Resonance (NMR) machinery (this is between 5Å and 6Å [43]). In other words, all atom pairs closer than this threshold are adjacent. The class of DGP instances defined by these unit sphere graphs, with $K = 3$, is known as the Molecular Distance Geometry Problem (MDGP) (see [29] and citations therein). We argued in [28, 22] that whenever the instance represents a protein backbone, the natural backbone order is a $^{\mathsf{K}}$DMDGP order. Since, however, the DGP and its variants require exact distances, whereas NMR data are best represented by intervals due to their associated measurement error [2], a remark is in order. Certain interatomic quantities are known reasonably precisely: notably, covalent bonds and covalent angles [43]. Moreover, because of the scaling between interval width and NMR machinery resolution scope, certain NMR intervals can be considered as finite sets of values [39]. This allows one to define a virtual order on protein backbones such that each consecutive quadruplet of atoms induces a 4-clique with the property that at most one of its edges is weighted by a finite set of values (instead of a single value) [23]. Under these conditions, the resulting graph can be seen as a collection of finitely many $^{\mathsf{K}}$DMDGP graphs, each of which is rigid.

The present work falls into a sequence of works about solving realizability problems using the BP approach. In this paper we make an important theoretical contribution about the structure of the solution set of $^{\mathsf{K}}$DMDGP instances: given any solution x, the others can all be obtained as "partial reflection" operators applied to x. This can also be used to speed up the time to find X, as shown in [33], where we apply some of the ideas of this paper to the problem of finding 3D realizations of protein backbones.

The rest of this paper is organized as follows. We recap some rigidity definitions and results (Sect. 2) and sketch the Branch-and-Prune (BP) algorithm for finding $^{\mathsf{K}}$DMDGP graph realizations (Sect. 3). We then give formal descriptions of the $^{\mathsf{K}}$DMDGP and of the BP algorithm (Sect. 4). Next, we study the geometrical properties of the BP search tree (Sect. 5), and prove that the number of solutions of YES instances of the $^{\mathsf{K}}$DMDGP is a power of two with probability one (Sect. 6). We exhibit a (zero measure) family of counterexamples to the "power of two" conjecture in Sect. 7, and finally we extend our results to the application-specific setting of protein conformation (Sect. 8).

2. Rigidity

Two distance spaces (S, d) and (S', d') in $\mathbb{R}^K$ are *congruent* when there exists an affine operator $T : \mathbb{R}^K \to \mathbb{R}^K$ such that the restriction of T to S is a bijection $S \to S'$ and $d(p, q) = d'(Tp, Tq)$ for all $p, q \in S$. Such operators are also called *isometries*. A *framework* is a pair (G, x) where $G = (V, E, d)$ is a simple undirected graph weighted by $d : E \to \mathbb{R}_+$ and $x : V \to \mathbb{R}^K$ is a valid realization of G. A *displacement* of a framework (G, x) is a continuous function $y : [0, 1] \to \mathbb{R}^{K|V|}$ such that $y(0) = x$ and $y(t)$ is a valid realization of G for all $t \in [0, 1]$. A *flexing* of a framework (G, x) is a displacement y of (G, x) such that $y(t)$ is incongruent to x for any $t \in (0, 1]$. A framework is *flexible* if it has a flexing, otherwise it is *rigid*. The *rigidity matrix* R of a framework (G, x) has $|E|$ rows and $K|V|$ columns; the row indexed by $\{u, v\} \in E$ has exactly $2K$ nonzero components, namely $x_{uk} - x_{vk}$ in the columns indexed by (u, k) and $x_{vk} - x_{uk}$ in the columns indexed by (v, k) (for $k \leq K$). The *complete rigidity matrix* $\bar{R}$ refers to the case where G is the full clique on V. The rigidity matrix is used to give a stronger definition of rigidity (namely that of infinitesimal rigidity) which is used in statics [47, 48]. Intuitively, infinitesimally rigid structures are resilient to collapse when certain forces are applied to them. Since for the kinds of frameworks considered in this paper rigidity and infinitesimal rigidity are equivalent [46, p. 23], we only limit the discussion to rigidity.

As defined above, rigidity is a property of frameworks rather than of graphs. It turns out, however, that if a graph has a certain type of rigid framework, then all its frameworks are rigid. More precisely, if x is a valid realization whose components are all algebraically independent over $\mathbb{Q}$ (i.e. there is no polynomial over $\mathbb{Q}$ having all the components as roots), then x is called a *generic realization*: if (G, x) is rigid, then all generic realizations of G are rigid, and we can then say that G is a *rigid graph*. This statement implicitly assumes that G is not weighted: in other words, a graph is rigid if its rigidity properties depend on the graph topology itself rather than the edge weight function, apart perhaps from a set of edge weight functions which have at most the cardinality of algebraic relations over $\mathbb{Q}$. Since the latter is a countable cardinality, whereas the set of all edge weight functions is uncountable, that is another way of saying that almost all edge weight functions yield a rigid framework if the graph is rigid. Having said that, the definition of generic realization given above is too stringent. Typically, rigid graphs might fail to yield rigid frameworks for

edge weight functions which cause $(K+1)$-cliques to be embedded in a space of dimension $< K$ (e.g. a 3-cycle embedded in $\mathbb{R}^2$ as three collinear points). Graver's definition of a generic realization [13] makes this concept precise: each nontrivial minor of the complete rigidity matrix must be nonzero. This implies that there is no need for calling in the very strong requirement of full algebraic independence over $\mathbb{Q}$: it suffices to make sure that the components of the realization giving rise to a rigid framework do not satisfy the polynomial equations engendered by the nontrivial minors of the complete rigidity matrix. If G has such a realization, then the graph is rigid. Also see Sect. 3.1.

Graphs with a unique realization are known as *uniquely* or *globally* rigid. Global rigidity recently generated several theoretical advances, such as a polynomial algorithm for realizing such graphs using the duality theory of Semidefinite Programming (SDP) [45] and the equivalence between unique K-localizability and generic universal rigidity [49] (among others). Unique K-localizability and universal rigidity are strong forms of unique realizability, which also involve uniqueness in higher dimensions. All these interesting properties arise in the study of localization of sensor networks, and are based on two application-driven requirements: unique realizability and the presence of a set of vertices (called *beacons* or *anchors*) whose realization is known *a priori*. Since at any time the network has exactly one localization to be determined, graphs should be dense enough so that they guarantee some unique realizability property. Moreover, most sensor networks are linked to other communication networks by means of routers (representing the anchors) whose position in space is fixed and known. When there are sufficiently many anchors in general positions, one can guarantee that any valid realization of the network can only take place in the full K space, and not in any smaller dimensional one. The application of rigidity to protein fails to provide either anchors or unique realizability guarantees. Quite on the contrary, it is interesting, from a biological point of view, to have a list of possible bio-polymers that a given set of distances can realize.

2.1. Henneberg type I graphs

Several important results on the rigidity of frameworks date from the end of the XIX century [8, 41], motivated by the construction of buildings whose supporting structures consisted of bars and joints. It appears from the literature of the period that verifying rigidity was an inductive process: one would start with a rigid structure and then add rigid components to it, so that the resulting structure would also be rigid. Henneberg [17] was the first to formalize two inductive steps for verifying rigidity. The first of these (known as *Henneberg type I* step [46, 18]) can be paraphrased (and generalized) as follows: if there is an order on V such that the first K vertices have a known realization, and such that every subsequent vertex is adjacent to at least K predecessors, then the graph almost certainly has a rigid realization in $\mathbb{R}^K$. This idea was already present in the works of Saviotti [40, Sect. XI, p. 57 and Fig. 30, pl.XV]. A similar order, which asks for the first K vertices to be a clique and each subsequent vertex to be adjacent to at least K predecessors, is called *discretization vertex order*, and the problem of finding such an order given a graph is known as

the Discretization Vertex Order Problem (DVOP) [20]. The DVOP is **NP**-complete, but only because the K-clique problem trivially reduces to it. Once the initial K-clique is known, a greedy procedure can find the order or decide it does not exist. The DVOP is therefore in **P** for fixed K. Because in Henneberg type I orders a partial realization is given for the first K vertices, it follows that the initial K-clique is known in advance; finding such an order in a graph is therefore also in **P**. Graphs with a Henneberg type I vertex order are also called *Henneberg type I graphs*. It is shown in [34] that graphs with a discretization vertex order are rigid; the realization problem for such graphs is known as the Discretizable Distance Geometry Problem (DDGP) and is **NP**-hard [34]. We remark that requiring that each vertex should have at least $K+1$ adjacent predecessors (instead of only K) yields a *K-trilateration order*: graphs with such orders are called *K-trilateration graphs* and can be realized in polynomial time [12]. It is interesting that the difference between the definitions of Henneberg type I orders and K-trilateration orders is as small as possible, and yet marks the distinction between an easy and a hard corresponding realization problem.

In the rest of this paper we are concerned with a particular type of Henneberg type I orders, namely those that ensure that the next vertex is adjacent to exactly K *immediate* predecessors, and perhaps also to other (not necessarily immediate) ones. We call these *KDMDGP orders*: they are important in using distances to find the conformation of proteins in space. As mentioned above, this restriction on the order does not make the corresponding realization problem any easier from the worst-case computational complexity standpoint [22]. It does guarantee rigidity, however, and it allows us to devise a recursive procedure for finding all incongruent realizations. Moreover, as we shall see, it also allows us to determine that the number of these realizations is a power of two for almost all edge weight functions.

3. Sphere intersections and Branch-and-Prune

Henneberg type I graphs are rigid because they enjoy the *Sphere Intersection Property* (SIP). Let $v \in V$ and suppose the first $v-1$ vertices have already been realized in $\mathbb{R}^K$: we know then that the v-th vertex is adjacent to at least K predecessors. Consider K of these. Since they precede v in the order, their position in $\mathbb{R}^K$ is already known by induction. Their distance to v is known because they are adjacent to v. Therefore, the position for v can only belong to the intersection of K spheres centered at these K adjacent predecessors. Since the intersection of K spheres in $\mathbb{R}^K$ generally consists of either 0 or 2 points (see Fig. 2), no vertex can ever be placed in an uncountable set of positions in $\mathbb{R}^K$, as the definition of flexing would imply. Here, the term *generally* has a meaning similar to the one given by Graver's definition of generic rigidity [13]: the set of edge weight functions for which the statement fails to hold has countable cardinality. Since such sets have Lebesgue measure 0 in $\mathbb{R}^K$, we also speak of this statement *holding with probability 1*. In this particular case, the set of K-tuples of spheres for which the SIP does not hold has Lebesgue measure

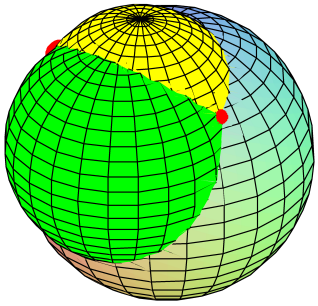


Figure 2: General case for the intersection P of three spheres in $\mathbb{R}^3$.

0 in the set of all possible K-tuples of spheres (see Fig. 3). See Sect. 3.1 for a

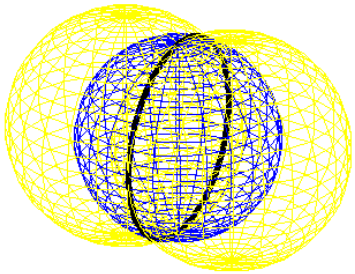


Figure 3: When three adjacent predecessors are collinear in $\mathbb{R}^3$, the sphere intersection (the thick black circle) might have uncountable cardinality.

more complete discussion about probability 1 statements.

The SIP can be exploited algorithmically by a recursive algorithm which will test in turn each of the two possible positions for the current vertex. If they yield valid partial realizations, then the algorithm calls itself recursively and tries to place the next vertex. This algorithm is called *Branch-and-Prune* (BP) [28]. The branching occurs because at each rank in the vertex order, two positions are possible. Pruning occurs when a vertex is adjacent to more than K adjacent predecessors: when a proposed position for the v-th vertex is inconsistent with some adjacent predecessor, recursion on the next vertex does not occur. The BP algorithm was originally only defined for $^{\mathsf{K}}$DMDGP graphs [21, 22], but was then extended to several other settings: for Henneberg type I graphs [34], for certain types of interval-weighted graphs related to proteins [32, 26, 23], and for the purpose of overcoming a technical limitation of NMR machinery, which generally only provides reliable distance measures for pairs of hydrogen atoms

[35, 24, 25, 27]. A publically available BP implementation is described in [38]. The current computational state-of-the-art for the BP algorithm is attained with a parallel BP implementation [37, 36], which can realize a protein backbone of 10^4 atoms (183444 distances) in $\mathbb{R}^3$ in 1.57s of CPU time on a cluster of 64 nodes.

It is easy to see that the BP yields a worst-case exponential behaviour, occurring when each vertex has exactly K adjacent predecessors. In such a case, the BP search tree is a full binary tree of height at most $|V|$ and width $2^{|V|-K}$ attained at the last level. Paths of length $|V|$ from the root to a leaf node encode realizations of the input graph and hence denote YES instances of the $^{\mathsf{K}}$DMDGP, whereas a tree with height strictly less than $|V|$ certifies a NO instance. In practice, however, the BP outperforms its continuous search competitors in both efficiency and reliability [22]. One particularly useful feature of BP is that, because the search is complete, it finds the set X of *all* incongruent realizations for a given graph, whereas most other DGP algorithms only find *one* realization. As already remarked, this is useful in biology because it allows one to list all the bio-polymers that are consistent with a particular set of distances.

In all our computational tests on $^{\mathsf{K}}$DMDGP instances, we observed that the number of incongruent realizations is a power of two: this comes to no surprise in the exponential worst case mentioned above, but there is no apparent reason why this should be the case when adjacent predecessors also include other vertices than the K immediate predecessors; and, indeed, in Sect. 7 we exhibit a set of counterexamples to the conjecture that all YES instances of the $^{\mathsf{K}}$DMDGP have power of two solutions. Yet, the computational trend remained unexplained. The main contribution of this paper is a proof that the set of YES instances of the $^{\mathsf{K}}$DMDGP such that $|X|$ is a power of two has Lebesgue measure 1 in the set of all YES instances of the $^{\mathsf{K}}$DMDGP. The statement is based on the assumption that we consider edge weight functions whose range consists of real numbers. We also partially extend this result to graphs which are more realistic protein models.

3.1. Statements holding with probability 1

In the following, we assume that the probability of any point of $\mathbb{R}^K$ belonging to any given subset of $\mathbb{R}^K$ having Lebesgue measure zero is equal to zero. Based on this assumption, when we state "$(\forall p \in P\ F(p))$ with probability 1" for a certain well-formed formula F with a free variable ranging over a set P having a strictly positive Lebesgue measure, we really mean that there exists a Lebesgue measurable subset $Q \subseteq P$, with Lebesgue measure equal to that of P, such that $\forall p \in Q\ F(p)$. Equivalently, statements holding with probability 1 should be taken to mean that the set of $^{\mathsf{K}}$DMDGP instances and partial realizations x for which the statements do *not* hold has Lebesgue measure zero in the set of all $^{\mathsf{K}}$DMDGP instances and partial realizations.

In this paper, zero Lebesgue measure sets are associated to cases that occur whenever pairs of real numbers (such as components of vectors) happen to be equal: we remark that affine subspaces of $\mathbb{R}^K$ defined by linear equations all have zero Lebesgue measure in $\mathbb{R}^K$. For example, the set of all pairs of points

$y, y' \in \mathbb{R}^K$ having t equal components (say, the first t), is described by the t linear equations $\forall j \leq t\ (y_j = y'_j)$ and therefore has zero Lebesgue measure in the set of all pairs of points in $\mathbb{R}^K$. In other words, when uniformly sampling pairs of points randomly from $\mathbb{R}^K$, the event that they should end up having at least one equal component has probability zero. This is similar to the spirit of Graver's definition of genericity using the nontrivial minors of the complete rigidity matrix [13]. The guiding principle in most of the results in this paper is that whenever a logical case of a proof requires two (or more) real values, sampled in a uniform distribution over a set of positive Lebesgue measure, to be equal, then this case has probability 0 to occur. The sampling usually refers to choosing instances randomly from the $^{\mathsf{K}}$DMDGP set, and more precisely to the real weights assigned to the edges of G. This directly translates to a uniform sampling over vectors of $\mathbb{R}^K$ (the components of the realizations associated to each instance).

We remark that the notion we consider is different from the usual genericity notion employed in rigidity theory [6], which requires distances to be algebraically independent over $\mathbb{Q}$. In order to solve the $^{\mathsf{K}}$DMDGP practically, we deal with instances whose edge weights are encoded in floating points, which are certainly not algebraically independent over $\mathbb{Q}$ (since they are themselves a subset of $\mathbb{Q}$). For example, Lemma 4.2 would not hold under algebraic independence (the intersection of $K+1$ "generic spheres" in $\mathbb{R}^K$ is empty), but it holds under our weaker requirement.

4. Formal definitions: the $^{\mathsf{K}}$DMDGP and the BP algorithm

For a set $U = \{x_i \in \mathbb{R}^K \mid i \leq K+1\}$ of points in $\mathbb{R}^K$, let D be the symmetric matrix whose (i,j)-th component is $\|x_i - x_j\|^2$ for all $i, j \leq K+1$ and let D' be D bordered by a left $(0, 1, \ldots, 1)^\top$ column and a top $(0, 1, \ldots, 1)$ row (both of size $K+2$). Then the Cayley-Menger formula states that the volume $\Delta_K(U)$ of the K-simplex on U is given by $\Delta_K(U) = \sqrt{\frac{(-1)^{K+1}}{2^K (K!)^2} |D'|}$. The strict simplex inequalities are given by $\Delta_K(U) > 0$. For $K = 2$, these reduce to strict triangle inequalities. We remark that only the distances of the simplex edges are necessary to compute $\Delta_K(U)$, rather than the actual points in U; the needed information can be encoded as a $(K+1)$-clique with these distances as edge weights.

Let $n = |V|$ and $m = |E|$. For all $v \in V$, let $N(v) = \{u \in V \mid \{u, v\} \in E\}$ be the star of vertices around v (also called the adjacencies of v); for a directed graphs (V, A), where $A \subseteq V \times V$, we denote the outgoing star by $N^+(v) = \{u \in V \mid (v, u) \in A\}$. For an order $<$ on V, let $\gamma(v) = \{u \in V \mid u < v\}$ be the set of predecessors of v, and let $\rho(v) = |\gamma(v)| + 1$ be the rank of v in $<$. For $V' \subseteq V$, we denote by $G[V']$ the subgraph of G induced by V'. For a finite set M, let $\mathcal{P}(M)$ be its power set. For a sequence $x = (x_1, \ldots, x_n)$ and a subset $U \subseteq \{1, \ldots, n\}$ we let $x[U]$ be the subsequence of x indexed by U. If x is an initial subsequence of y, then y is an *extension* of x. We denote a clique on q vertices by $\mathbf{K}_q$.

GENERALIZED DISCRETIZABLE MOLECULAR DISTANCE GEOMETRY PROBLEM ($^{\mathsf{K}}$DMDGP). Given an undirected graph $G = (V, E)$, an edge weight function $d : E \to \mathbb{R}_+$, an integer $K > 0$, a subset $V_0 \subseteq V$ with $|V_0| = K$, a partial realization $\bar{x} : V_0 \to \mathbb{R}^K$ valid for $G[V_0]$, and a total order $<$ on V such that for each v with $\rho(v) > K$ there is a set $U_v \subseteq N(v) \cap \gamma(v)$ with the following properties:

$$\{v \in V \mid \rho(v) \leq K\} = V_0; \tag{2}$$

$$\forall u \in U_v \quad (\rho(v) - K \leq \rho(u) \leq \rho(v) - 1); \tag{3}$$

$$\forall v \in V \setminus V_0 \ (G[U_v] = \mathbf{K}_K \ \wedge \ \Delta_{K-1}(U_v) > 0), \tag{4}$$

decide whether there is a valid extension $x : V \to \mathbb{R}^K$ of $\bar{x}$.

Condition (2) requires the first K vertices to induce a K-clique in G. Condition (3) requires the K immediate predecessors of v to be also adjacent to it. Condition (4), which was not mentioned in the informal discussion in Sect. 1-2, prevents the realizations x of G from being flexible because of an algebraic dependence on the components of x. Specifically, requiring the Cayley-Menger determinants to be strictly positive implies that no K-clique determined by immediate predecessor will embedded as a volume 0 simplex in $\mathbb{R}^{K-1}$ (Fig. 3 shows the effect of a triangle embedded in a line instead of in a plane: the SIP fails to hold).

Two further remarks are in order.

- Whenever vertex v is being considered in the BP algorithm, all its predecessors have already been placed. Hence, all of the distances between all predecessors are already known; thus, the BP can also solve instances for which $G[U_v]$ is not the full K-clique, although they are not formally in the $^{\mathsf{K}}$DMDGP.

- Edge weights are real numbers in $^{\mathsf{K}}$DMDGP instances: this naturally makes the Turing machine model difficult to apply. We observe, however, that it is not known whether the DGP is in **NP**, since the certifying realizations might have irrational components even in the case when the weights are restricted to be integer (just take an equilateral triangle with unit weights). In practice, our algorithms work with floating point numbers, so whenever we say "exact" we mean "exact in theory". In practice, solutions will be floating point approximations of the exact solutions. This, however, holds for all existing methods targeting DGP-type problems.

4.1. Branching and Pruning

Let G be a $^{\mathsf{K}}$DMDGP instance. Consider $v \in V$ with rank $\rho(v) = i > K$, let $G^v = G[\gamma(v) \cup \{v\}]$ and x be a valid realization of $G[\gamma(v)]$. We characterize the number of extensions of x valid for G^v in the following lemmata (which also hold for the DDGP). Lemma 4.1 (resp. 4.2) essentially state that, under the given conditions, $G[\{v\} \cup (N(v) \cap \gamma(v))]$ is a rigid (resp. uniquely rigid) graph. The results in this section are not new, but we list the proofs here because these Lemmata form the basic toolbox for what is to follow.

4.1 Lemma
If $|N(v) \cap \gamma(v)| = K$ then there are at most two distinct extensions of x that are valid for G^v. If one valid extension exists, then with probability 1 there are exactly two distinct valid extensions.

Proof. Since $|N(v) \cap \gamma(v)| = K$, $U_v = N(v) \cap \gamma(v)$ and v is at the intersection of exactly K spheres in $\mathbb{R}^K$ (each centered at x_u with radius d_{uv}, where $u \in U_v$). The position $z \in \mathbb{R}^K$ of v must then satisfy:

$$\forall u \in U_v \quad \|z - x_u\| = d_{uv} \Rightarrow \|z\|^2 - 2x_u \cdot z + \|x_u\|^2 = d_{uv}^2. \tag{5}$$

As in [11], we choose an arbitrary $w \in U_v$, say $w = \max_< U_v$, and subtract from the row of Eq. (5) indexed by w the other rows of (5), obtaining the system:

$$\left.\begin{array}{rcl} \forall u \in U_v \smallsetminus \{w\} \quad 2(x_u - x_w) \cdot z & = & (\|x_u\|^2 - d_{uv}^2) - (\|x_w\|^2 - d_{wv}^2) \\ \|z\|^2 - 2x_w \cdot z + \|x_w\|^2 & = & d_{wv}^2. \end{array}\right\} \tag{6}$$

System (6) consists of a set of $K-1$ linear equations and a single quadratic equation in the K-vector z. We write the linear equations as the system $Az = b$, where the (u,j)-th component of A is $2(x_{uj} - x_{wj})$, the u-th component of b is $\|x_u\|^2 - \|x_w\|^2 - d_{uv}^2 + d_{wv}^2$, A is a $(K-1) \times K$ matrix and $b \in \mathbb{R}^{K-1}$. By the strict simplex inequalities, A has full rank (for otherwise the linear dependence condition $\sum_{u \neq w} \xi_u (x_u - x_w) = 0$, for some coefficients ξ_u, implies that x_w is in the span of $\{x_u \mid u \in U_v\}$, and hence that $\Delta_{K-1}(U_v) = 0$); so without loss of generality assume that the square matrix B formed by the first $K-1$ columns of A is invertible. Let z_B be the vector consisting of the first $K-1$ components of z; then the linear part (first $K-1$ equations) of (6) yields $z_B = B^{-1}(b - Nz_K)$ as a function $z_B(z_K)$ of z_K, where $N = 2(x_{uK} - x_{wK} \mid u \in U_v \smallsetminus \{w\}) \in \mathbb{R}^{K-1}$. After replacement of z_B in (6) with $z_B(z_K)$, we obtain the following quadratic equation in z_K:

$$(\|\bar{N}\|^2 + 1) z_K^2 - 2((\bar{b} + x_{wB})\bar{N} + x_{wK}) z_k + (\|x_{wB} - \bar{b}\|^2 + x_{wK}^2 - d_{wv}^2) = 0, \tag{7}$$

where $\bar{b} = B^{-1}b$ and $\bar{N} = B^{-1}N$. If the discriminant of (7) is negative then no extension of $\bar{x}$ to v is possible and the result follows. If the discriminant is nonnegative, (7) has solutions z'_K, z''_K yielding points $z' = (z_B(z'_K), z'_K)$ and $z'' = (z_B(z''_K), z''_K) \in \mathbb{R}^K$, which are distinct with probability 1 because the discriminant is zero with probability 0. The extended realizations, distinct with probability 1, are given by (x, z') and (x, z''). □

4.2 Lemma
If $|N(v) \cap \gamma(v)| > K$ then, with probability 1, there is at most one extension of x.

Proof. Consider a subset $S \subseteq N(v) \cap \gamma(v)$ such that $|S| = K+1$ and $S \supseteq U_v$. Either there is at least one point x_v such that (x, x_v) is a realization of $G[S \cup \{v\}]$ that is valid w.r.t. the system:

$$\forall u \in S \quad \sum_{k \leq K} (x_{vk}^2 - 2x_{uk}x_{vk} + x_{uk}^2) = d_{uv}^2, \tag{8}$$

or the system has no solution. In the latter case, the result follows, so we assume now that there is a point x_v satisfying (8). Since the points x_u are known for all $u \in S$, (8) is a quadratic system with K variables and $K+1$ equations. As in the proof of Lemma 4.1, we derive an equivalent linear system from (8). Since d satisfies the strict simplex inequalities on U_v with probability 1 and $S \supseteq U_v$, by [7] $\{x_u \mid u \in S\}$ are not co-planar and the system has exactly one solution, as claimed. □

4.3 Lemma
With the notation of Lemma 4.1, if $\bar{x}$ is a valid realization for $G[U_v]$, then z'' is a reflection of z' with respect to the hyperplane through the K points of $\bar{x}$.

Proof. Any sphere in $\mathbb{R}^K$ is symmetric with respect to any hyperplane through its center; so the intersection of up to K spheres in $\mathbb{R}^K$ is symmetric with respect to the hyperplane containing all the centers. □

4.4 Remark
Reflections with respect to hyperplanes are isometries, and can therefore be represented by linear operators. If $a \in \mathbb{R}^K$ is the unit normal vector to a hyperplane H containing the origin, then the reflection operator R_0 w.r.t. H can be expressed in function of the standard basis by the matrix $I - 2aa^\top$, where I is the $K \times K$ identity matrix [5]. Let H be a hyperplane with equation $a^\top x = a_0$ (with $a_0 \neq 0$) and a_i, for some $1 \leq i \leq K$, be the nonzero coefficient of smallest index in a. Then, the reflection operator R acting on a point $p \in \mathbb{R}^K$ w.r.t. H is given by $R(p) = R_0(p - \frac{a_0}{a_i}e_i) + \frac{a_0}{a_i}e_i$, where $e_i \in \mathbb{R}^K$ is the unit vector with 1 at index i and 0 elsewhere: we first we translate p so that we can reflect it using R_0 w.r.t. the translation of H containing the origin, then we perform the inverse translation of the reflection.

4.2. The BP search tree

We denote the BP binary search tree by $\mathcal{T} = (\mathcal{V}, \mathcal{A})$. The tree is directed from the root to the leaf nodes, which are triplets $\alpha = (x(\alpha), \lambda(\alpha), \mu(\alpha))$. Informally, $x(\alpha)$ is a realization from the root to node α, $\lambda(\alpha)$ is 0 or 1 according as to whether α is a left or right subnode of its parent, and μ is $\boxplus$ if $x(\alpha)$ is feasible and $\boxminus$ otherwise.

More formally, for $\alpha \in \mathcal{T}$ we denote by $\mathsf{p}(\alpha)$ the unique path from the root node r of $\mathcal{T}$ to α, and by α^- the unique parent node of α (unless $\alpha = \mathsf{r}$, in which case we define $\mathsf{r}^- = \mathsf{r}$). The symbol $x(\alpha)$ is defined recursively to denote an extension of the realization $x^- = x(\alpha^-)$ found on $\mathsf{p}(\alpha^-)$. The symbol $\lambda(\alpha) \in \{0, 1\}$ distinguishes whether α is a "left" or a "right" subnode of α^-. More precisely, let α be a node at level i in $\mathcal{T}$, $v = \rho^{-1}(i)$, $\bar{x}$ be a partial realization of $G[U_v]$, and $a_v^\top x = a_{v0}$ be the equation of the hyperplane through the points of $\bar{x}$, which is $(K-1)$-dimensional by (4). Assuming that $u = \rho^{-1}(i-1)$ and

$a_v \in \mathbb{R}^K$ is oriented so that $a_v \cdot a_u \geq 0$; then:

$$\lambda(\alpha) = \begin{cases} 0 & \text{if} \quad a_v^\top x(\alpha)_i \leq a_{v0} \\ 1 & \text{if} \quad a_v^\top x(\alpha)_i > a_{v0}. \end{cases} \tag{9}$$

Lastly, $\mu(\alpha) = \boxplus$ if $x(\alpha)$ is a valid extension of x^-, in which case the node is said to be *feasible*, and $\mu = \boxminus$ otherwise. This allows us to retrieve the set X of all valid realizations of G by simply traversing $\mathcal{T}$ backwards from the leaf nodes marked $\boxplus$ up to r.

Although the notation we introduced to describe the BP algorithm looks overly formal and complicated, it allows us to give rigorous proofs of subsequent material.

We remark that Alg. 1 differs from the original BP formulation [28] because it applies to K dimensions and explicitly stores several details of the binary search tree.

4.5 Lemma
At termination of Alg. 1, X contains all valid realizations of G extending $\bar{x}$.

Proof. Z (in Step 17) exists with probability 1 by Lemma 4.1. Every realization in X is valid because of Steps 17 and 19-20. No other valid extension of $\bar{x}$ exists because of Lemmata 4.1-4.2. □

The realizations in X are incongruent apart perhaps from between 1 and $n-K$ reflections along the hyperplane defined by $\bar{x}_1, \ldots, \bar{x}_v$. The exact number of these reflections depends on the order on V and the edges in G.

We now partition $\mathcal{V}$ in pairwise disjoint subsets $\mathcal{V}_1, \ldots, \mathcal{V}_n$ where for all $i \leq n$ the set $\mathcal{V}_i$ contains all the nodes of $\mathcal{V}$ at level i of the tree $\mathcal{T}$. We show in Prop. 4.6 that no level of $\mathcal{T}$ has two distinct feasible nodes having respectively one and two feasible subnodes.

4.6 Proposition
With probability 1, there is no level $i \leq n$ having two distinct feasible nodes $\beta, \theta \in \mathcal{V}_i$ such that $|\{\alpha \in N^+(\beta) \mid \mu(\alpha) = \boxplus\}| = 1$ and $|\{\alpha \in N^+(\theta) \mid \mu(\alpha) = \boxplus\}| = 2$.

Proof. We show that for all $i \leq n$ the event of having two distinct nodes $\beta, \theta \in \mathcal{V}_i$, with $\rho^{-1}(i) = v$, such that β has one feasible subnode and θ has two has probability 0. Consider $T_v = N(v) \cap \gamma(v)$: if $|T_v| = K$ then, by Lemma 4.1, β should have exactly two feasible subnodes with probability 1. Since by hypothesis it only has one, the event $|T_v| = K$ occurs with probability 0. Since $|T_v| \geq K$ by (4), the event $|T_v| > K$ occurs with probability 1. Thus by Lemma 4.2 there is, with probability 1, at most one valid realization extending the partial realization at v, which means that the two feasible subnodes of θ represent the same realization, an event that occurs with probability 0. □

We remark that Prop. 4.6 also holds for the DDGP, provided U_v is chosen in Alg. 1 as any subset of $N(v) \cap \gamma(v)$ satisfying the constraints of Eq. (4).

Algorithm 1 The Branch and Prune algorithm.

Require: Partial realization $\bar{x}$ of first K vertices of G
Ensure: Set X of valid realizations of G
Let $\alpha = (\bar{x}_1, 0, \boxplus)$ and $\alpha' = (\bar{x}_1, 1, \boxminus)$
Initialize $\mathcal{V} = \{\alpha, \alpha'\}$ and $\mathcal{A} = \{(\mathsf{r}, \alpha), (\mathsf{r}, \alpha')\}$
for $1 < i \leq K$ **do**
  Let $\alpha = (\bar{x}_i, 0, \boxplus)$, $\alpha' = (\bar{x}_i, 1, \boxminus)$, $\beta = (\bar{x}_{i-1}, 0, \boxplus)$
  Let $\mathcal{V} \leftarrow \mathcal{V} \cup \{\alpha, \alpha'\}$ and $\mathcal{A} \leftarrow \mathcal{A} \cup \{(\beta, \alpha), (\beta, \alpha')\}$
end for
BranchAndPrune$(K+1,\ (\bar{x}_K, 0, \boxplus))$
Let $X = \{x(\theta) \mid \theta \in \mathcal{V} \wedge |N^+(\theta)| = 0 \wedge \mu(\theta) = \boxplus\}$
stop

function BranchAndPrune$(i,\ \beta)$:
if $i > n \vee \mu = \boxminus$ **then**
  return
end if
Let $v = \rho^{-1}(i)$
Compute the equation $a_v^\top x = a_{v0}$ of the hyperplane through $x[U_v]$
Let $Z = \{z', z''\}$ be extensions of $x(\beta)$ to v, and $Z' = Z$
for $z \in Z$ **do**
  if $\exists\{u, v\} \in E\ \|x(\beta)_u - z\| \neq d_{uv}$ **then**
    Let $Z = Z \setminus \{z\}$
  end if
end for
if $Z = \{z', z''\}$ **then**
  if $a_v^\top z' \leq a_{v0}$ **then**
    Let $\alpha = (z', 0, \boxplus)$, $\alpha' = (z'', 1, \boxplus)$
  else
    Let $\alpha = (z'', 0, \boxplus)$, $\alpha' = (z', 1, \boxplus)$
  end if
else if $Z = \{z\}$ **then**
  if $a_v^\top z \leq a_{v0}$ **then**
    Let $\alpha = (z, 0, \boxplus)$, $\alpha' = (Z' \setminus \{z\}, 1, \boxminus)$
  else
    Let $\alpha = (z, 1, \boxplus)$, $\alpha' = (Z' \setminus \{z\}, 0, \boxminus)$
  end if
else
  return
end if
Let $\mathcal{V} \leftarrow \mathcal{V} \cup \{\alpha, \alpha'\}$ and $\mathcal{A} \leftarrow \mathcal{A} \cup \{(\beta, \alpha), (\beta, \alpha')\}$
for $\theta \in N^+(\beta)$ such that $\mu(\theta) = \boxplus$ **do**
  BranchAndPrune$(i+1,\ \theta)$
end for
return

Lastly, we emphasize the fact that for all $\ell \in \{i, \dots, n\}$ and for all $\alpha \in \mathcal{V}_\ell$ the set $\mathsf{p}(\alpha) \cap \mathcal{V}_i$ has a unique element, as it contains the unique node at level i on the path from α to the BP tree root node.

5. Geometry in BP Trees

The most important result of this section is that, for any valid realization $y \in X$, if the BP tree branches at level $i = \rho(v)$ on the path to y and both branches continue to the last level, then the realization obtained by reflecting all the points of y past the $(i-1)$-st vertex through the hyperplane defined by $y[U_v]$ is also valid with probability 1. We remark that the results in this section only apply to the $^{\mathsf{K}}$DMDGP (not to the DDGP, as shown in the counterexample of Fig. 8).

We need to emphasize those BP branchings which carry on to feasible leaf nodes along both branches. For $y \in X$ and a vertex $v \in V \setminus V_0$ we denote $\Upsilon(y, v)$ the following property:

$\Upsilon(y, v)$: *there are feasible leaf nodes* $\beta, \beta' \in \mathcal{V}_n$ *such that* $x(\beta) = y$, $\mathsf{p}(\beta) \cap \mathcal{V}_{\rho(v)-1} = \mathsf{p}(\beta') \cap \mathcal{V}_{\rho(v)-1}$ *and* $\mathsf{p}(\beta) \cap \mathcal{V}_{\rho(v)} \neq \mathsf{p}(\beta') \cap \mathcal{V}_{\rho(v)}$.

In other words, $\Upsilon(y, v)$ holds whenever $\mathsf{p}(\beta) \cap \mathcal{V}_{\rho(v)-1}$ contains a feasible node with two feasible subnodes leading to different valid realizations.

5.1. Partial reflection operators

With $\Upsilon(y, v)$ true, we let R^v be the Euclidean reflection operator with respect to the hyperplane through $y[U_v]$ (as discussed in Remark 4.4). Define

$$\tilde{R}^v = I^{\rho(v)-1} \times (R^v)^{n-\rho(v)},$$

i.e., $\tilde{R}^v y = (y_1, \dots, y_{\rho(v)-1}, R^v y_{\rho(v)}, \dots, R^v y_n)$ for any realization y. This is a *partial reflection* of y which only acts on vertices past rank $\rho(v) - 1$.

We remark that the matrix representing R^v could change depending on y. Since we wish R^v to represent the effect of a reflection at level v rather than the reflection matrix itself, we introduce the following technicality. Consider an equivalence relation on the set of all possible (not necessarily feasible) realizations $V \to \mathbb{R}^K$ extending $\bar{x}$. Let $E_{\mathsf{d}} \subseteq E$ be the set of edges $\{u, v\} \in E$ such that $|\rho(v) - \rho(u)| \leq K$, and G_{d} be the subgraph of G defined by E_{d}. Two realizations y, y' are equivalent if both are valid for G_{d}. Now R^v can be formally defined as the set of matrices representing R^v over all realizations of the same equivalence class. This definition carries over to $\tilde{R}^v$.

The following is an easy corollary to Lemma 4.3, and states, with the formal notation introduced for Alg. 1, that two feasible subnodes of a feasible node are associated with partial realizations whose last components are K-vectors which are reflections of each other.

5.1 Corollary
Let $\alpha \in \mathcal{V}_{i-1}$ *for some* $i > 1$, $v = \rho^{-1}(i)$ *and* $N^+(\alpha) = \{\eta, \beta\}$ *with* $\mu(\eta) = \mu(\beta) = \boxplus$. *Then* $x(\eta)_v = R^v x(\beta)_v$.

5.2. Probability of conditional events

In most subsequent results, we assume the considered $^{\mathsf{K}}$DMDGP instance to be a YES one, so that probabilities are conditional to this event.

The following remark is in order.

5.2 Remark
If $\Upsilon(y, v)$ holds for some $y \in X$ and $v \in V \smallsetminus V_0$, then by definition there are feasible leaf nodes in the BP tree, which implies that the considered $^{\mathsf{K}}$DMDGP instance is YES.

An important consequence of Remark 5.2 is that all statements assuming $\Upsilon(y, v)$ and claiming a result with probability 1 implicitly also assume that the probability is conditional to the event of the $^{\mathsf{K}}$DMDGP instance being a YES one. In particular, since the instance is YES, certain points *must* be placed at certain distances with probability 1 (notably, at distances satisfying the equations (1)), for otherwise the instance would be NO. This is evident in Prop. 5.4, Cor. 5.6, Cor. 5.7, and Thm. 5.9, where we state that certain real scalars and vectors must belong to certain finite sets with probability 1. The sense of these assertions, in this context, is that the Lebesgue measure of the set of YES instances not satisfying the result is zero in the set of all YES instances.

5.3. Reflections in the BP tree

We build towards the main result of the section; we start with a technical lemma which relates the position of reflected points in the realizations with the "left/right" λ-components of the corresponding BP node triplets.

5.3 Lemma
Let $\alpha \in \mathcal{V}_{i-1}$ for some $i > 1$ such that $N^+(\alpha) = \{\eta', \beta'\}$, let $u = \rho^{-1}(i)$, let v with $\rho(v) = \ell$ be such that $\ell > i$, and consider two feasible nodes $\eta, \beta \in \mathcal{V}_\ell$ such that $\{\eta'\} = \mathsf{p}(\eta) \cap \mathcal{V}_i$ and $\{\beta'\} = \mathsf{p}(\beta) \cap \mathcal{V}_i$. Further, $\forall\, i \leq j \leq \ell$ let $w = \rho^{-1}(j)$, $\mathsf{p}(\eta) \cap \mathcal{V}_j = \{\eta''\}$ and $\mathsf{p}(\beta) \cap \mathcal{V}_j = \{\beta''\}$. Then, with probability 1, the following statements are equivalent:

(*i*) $\forall\, i \leq j \leq \ell \quad x(\beta'')_w = R^u x(\eta'')_w$;

(*ii*) $\forall\, i \leq j \leq \ell \quad \lambda(\eta'') = 1 - \lambda(\beta'')$.

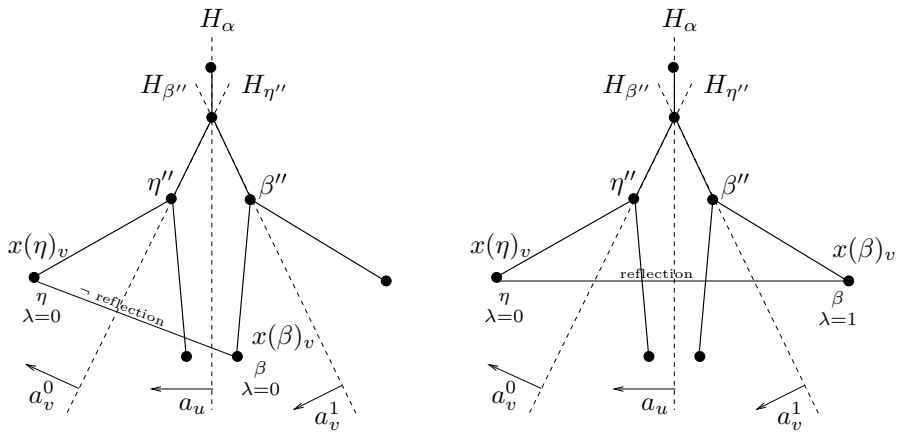
Proof. Let ${a_v^0}^\top x = a_{v0}^0$, ${a_v^1}^\top x = a_{v0}^1$ be the equations of the hyperplanes H_η, H_β defined respectively by $x(\eta)[U_v]$ and $x(\beta)[U_v]$, with the normals oriented as explained in Remark 4.4. We prove by induction on $\ell - i$ that the following statement is equivalent to (*i*) and (*ii*):

(*iii*) for all $i \leq j \leq \ell$, $x(\beta'')_w = R^u x(\eta'')_w$ and $a_u \cdot a_w^0 = a_u \cdot a_w^1$, where a_w^0 and a_w^1 are the normal vectors of the hyperplanes $H_{\eta''}$ and $H_{\beta''}$ oriented as usual.

The induction starts because, if $\ell = i$, then (i), (ii), and (iii) hold simultaneously. Indeed, $\eta = \eta'$ and $\beta = \beta'$, hence $x(\beta)_v = R^u x(\eta)_v$ (Lemma 4.3) and $\lambda(\eta) = 1 - \lambda(\beta)$ (Alg. 1, Steps 25 and 27). In addition, we have $H_\eta = R^u H_\beta$, therefore $|a_u \cdot a_v^0| = |a_u \cdot a_v^1|$. Because the orientation of a_v^0, a_v^1 is such that $a_u \cdot a_v^0, a_u \cdot a_v^1 \geq 0$, the result holds. Assume that the equivalence stated above holds for level $\ell - 1$, we show that it is still the case at level ℓ. In the sequel, denote $t = \rho^{-1}(\ell - 1)$.

$(iii) \Rightarrow (i)$: trivial.

$(i) \Rightarrow (ii)$. Suppose for all $i \leq j < \ell$, $x(\beta'')_w = R^u x(\eta'')_w$ and $\lambda(\eta'') = 1 - \lambda(\beta'')$ (by the induction hypothesis, both statements are equivalent). Hence, $H_{\eta''} = R^u H_{\beta''}$ holds for all j, because the K points generating the hyperplanes either belong to H_α, or are reflections of each other. This is true in particular if we choose $\eta'', \beta'' \in \mathcal{V}_{\ell-1}$. In addition, if we use the induction hypothesis $(i) \Rightarrow (iii)$, we have $a_u \cdot a_t^0 = a_u \cdot a_t^1$, so a_t^0, a_t^1 are directed similarly w.r.t a_u, and hence $\lambda(\eta) = 1 - \lambda(\beta)$ if and only if $x(\beta)_v = R^u x(\eta)_v$ (see Fig. 4).



(a) ${a_v^0}^\top x(\eta)_v > a_{v0}^0$ and ${a_v^1}^\top x(\beta)_v > a_{v0}^1$ (b) ${a_v^0}^\top x(\eta)_v > a_{v0}^0$ and ${a_v^1}^\top x(\beta)_v < a_{v0}^1$

Figure 4: Proof of Lemma 5.3: Case (4a) shows the contradiction deriving from $\lambda(\eta) = \lambda(\beta) = 0$ (or $x(\beta)_v \neq R^u x(\eta)_v$), and case (4b) the situation that actually occurs.

$(ii) \Rightarrow (iii)$. Suppose for all $i \leq j \leq \ell$, $\lambda(\eta'') = 1 - \lambda(\beta'')$. By the previous result, we also know that for all $i \leq j \leq \ell$, $x(\beta'')_w = R^u x(\eta'')_w$. It remains to prove that $a_u \cdot a_v^0 = a_u \cdot a_v^1$, i.e. that the angles θ_v^0 and θ_v^1 formed by these vectors have the same cosine. Notice once again that $H_\eta = R^u H_\beta$. By induction, we know that the angles θ_t^0, θ_t^1 formed by a_u and respectively a_t^0, a_t^1, have same cosine. With probability 1, the hyperplanes H_η, H_β are not parallel, hence their normal vectors cannot be identical, therefore, $\theta_t^0 = -\theta_t^1$ (see the illustration on Fig. 5). Denote θ^0, θ^1 the angles formed respectively by a_t^0 and a_v^0, and by a_t^1 and a_v^1. We also have, $H_{\eta''} = R^u H_{\beta''}$, where $\eta'', \beta'' \in \mathcal{V}_{\ell-1}$, hence the normal vectors of these 4 hyperplanes are also symmetric, which implies $\theta^0 = -\theta^1$ or $\theta^0 = \pi - \theta^1$. By the definition of a_v^0 and a_v^1 (page 12), since the scalar products are positive,

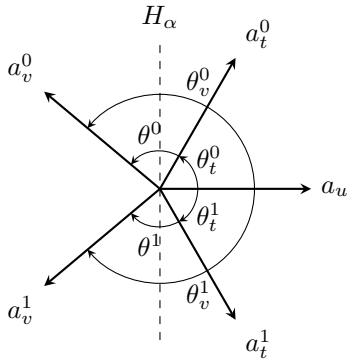


Figure 5: Proof of Lemma 5.3: illustration of the fact that $a_u \cdot a_v^0 = a_u \cdot a_v^1$.

$-\pi/2 \leq \theta^0, \theta^1 \leq \pi/2$, thus $\theta^0 = -\theta^1$. Therefore, $\theta_v^0 = \theta_t^0 + \theta^0 = -\theta_t^1 - \theta^1 = -\theta_v^1$, which concludes this part of the proof. □

The following important proposition considers vertices $u, v \in V$ such that $|\rho(v) - \rho(u)| = K + 1$ and states that there are only two possible values of $\|y_u - y_v\|$ if y is to be a feasible realization.

5.4 Proposition
Consider a subtree $\mathcal{T}'$ of $\mathcal{T}$ consisting of $K+2$ consecutive levels $i-K-1, \ldots, i$ (where $i \geq 2K+1$), rooted at a single node η and such that all nodes at all levels are marked $\boxplus$. Let $p = 2^{K+1}$ and consider the set $Y' = \{y^j \mid j \leq p\}$ of partial realizations of G at the leaf nodes $\{\alpha_j \mid j \leq p\}$ of $\mathcal{T}'$, such that $\forall j \leq p$ $(y^j = x(\alpha_j))$. Let $u = \rho^{-1}(i-K-1)$ and $v = \rho^{-1}(i)$. Then with probability 1 there are two distinct positive reals r, r' such that $\|y_u^j - y_v^j\| \in \{r, r'\} = H^{uv}$ for all $j \leq p$.

Proof. Fig. 6 shows a graphical proof sketch for $K = 2$. With a slight abuse of notation, for a vertex $w \in V$ in this proof we denote by R^w the set of all reflections at level w. We order the α_j nodes (and the corresponding y^j) so that the action of R^v on $(\alpha_1, \ldots, \alpha_p)$ is the permutation $\prod_{j \bmod 2 = 1}(j, j+1)$. Let $t = \rho^{-1}(i-1)$. Since all nodes are feasible, $\|y_v^j - y_t^j\| = d_{tv}$ and $\|y_u^j - y_t^j\| = d_{ut}$ for all $j \leq p$ (we remark that $\{t, v\}$ and $\{u, t\}$ must be in E by the $^{\mathsf{K}}$DMDGP definition). With probability 1, the segments through y_u^j and y_t^j (where $j \leq p$) do not respectively lie within the hyperplanes defining the reflections R^v; and the same holds for the segments through y_t^j and y_v^j. Thus, there is a set Q of positive reals $r_1, \ldots, r_p$ s.t. for all $j \leq p$ with $j \bmod 2 = 1$ we have $\|y_u^j - y_v^j\| = r_j$ and $\|y_u^{j+1} - y_v^{j+1}\| = r_{j+1}$, which shows $|Q| \leq p = 2^{K+1}$. By Lemma 5.3, the action of R^t on $(\alpha_1, \ldots, \alpha_p)$ is the permutation $\prod_{j \bmod 4 = 1}(j, j+3)(j+1, j+2)$: this implies that $r_j = r_{j+3}$ and $r_{j+1} = r_{j+2}$ for all $j \bmod 4 = 1$, which shows $|Q| \leq p/2 = 2^K$. Inductively, for a vertex w s.t. $i - K \leq \rho(w) \leq i-1$ the action

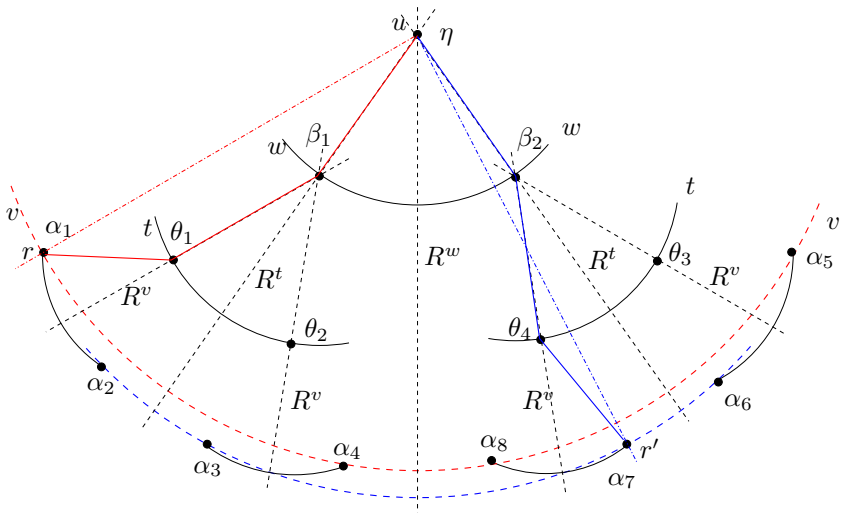


Figure 6: Proof of Prop. 5.4 in $\mathbb{R}^2$. The arrangement of three segments gives rise, in general, to two distances r, r' between root and leaves.

of R^w is $\prod_{j \bmod 2^{j-\rho(w)+1}} (j, j+2^{i-\rho(w)+1}-1)(j+1, j+2^{i-\rho(w)+1}-2)\cdots(j+2^{i-\rho(w)}-1, j+2^{i-\rho(w)})$, which implies that $|Q| \leq 2^{K+1-i+\rho(w)}$. Therefore $\rho(w) = i-K$ proves that $|Q| \leq 2$. The case $|Q| = 1$ can only occur if y_u^j, y_t^j and y_v^j are collinear for all $j \leq p$, an event that occurs with probability 0. □

Prop. 5.4 is useful in order to show that certain configurations of nodes within $\mathcal{T}$ can only occur with probability 0.

5.5 Example
Consider a subtree $\mathcal{T}'$ of $\mathcal{T}$ like the one in Fig. 6 embedded in $\mathbb{R}^2$, and suppose that all nodes at level u, w, t are marked $\boxplus$, and further that only one node within α_1, α_2 is feasible, only one node within α_3, α_4 is feasible, only one node within α_7, α_8 is feasible, and α_5, α_6 are both infeasible. This must be due to a distance $d_{u'v}$ with $u' \leq u$. Consider now a circle C completely determined by its center at $y_{u'}^1 = x(\alpha_1)_u$ and its radius $d_{u'v}$; if C also contains the points at the nodes $\alpha_1, \alpha_4, \alpha_8$ or the points at the nodes $\alpha_2, \alpha_3, \alpha_7$ then we must have $u' = u$, in which case also one of α_5, α_6 will be feasible (against the hypothesis). And the probability that C should contain the points at the nodes $\alpha_1, \alpha_3, \alpha_8$ or $\alpha_2, \alpha_4, \alpha_7$ is zero. Hence $\mathcal{T}'$ can only occur with probability 0.

We now generalize Prop. 5.4 to vertices $u, v \in V$ with arbitrary rank difference.

5.6 Corollary
Consider a subtree $\mathcal{T}'$ of $\mathcal{T}$ consisting of $K+h+1$ consecutive levels $i-K-h, \ldots, i$ (where $i \geq 2K+h$ and $h \geq 1$), rooted at a single node η and such

that all nodes at all levels are marked $\boxplus$. *Let* $p = 2^{K+h}$ *and consider the set* $Y' = \{y^j \mid j \leq p\}$ *of partial realizations of* G *at the leaf nodes of* $\mathcal{T}'$. *Let* $u = \rho^{-1}(i - K - h)$ *and* $v = \rho^{-1}(i)$. *Then with probability 1 there is a set* $H^{uv} = \{r_j \mid j \leq 2^h\}$ *of* 2^h *distinct positive reals such that* $\|y_u^j - y_v^j\| \in H^{uv}$ *for all* $j \leq p$.

Proof. The proof of Prop. 5.4 can be generalized to span an arbitrary number of levels by induction on q. Two distances $r_{j_1}, r_{j_2} \in H^{uv}$ can only be equal by collinearity of some subsets of points, an event occurring with probability 0. □

The next results shows that, if $\{u, v\}$ is an edge of E with rank difference greater than K, the distance d_{uv} must belong to a certain finite set of values whenever the instance is a YES one.

5.7 Corollary
Let $y \in X$ *and* $v \in V \setminus V_0$ *such that* $\Upsilon(y, v)$ *holds. If* $\{u, w\} \in E$ *with* $u < v < w$ *and* $\rho(w) - \rho(u) > K$ *then* $d_{uw} \in H^{uw}$ *with probability 1.*

Proof. Since $\Upsilon(y, v)$ holds, then the $^{\mathsf{K}}$DMDGP instance is YES and there must exist at least two feasible nodes at level $\rho(w)$ in $\mathcal{T}$. If $d_{uw} \notin H^{uw}$ the probability for a given sphere to contain two arbitrary points in $\mathbb{R}^K$ is zero. Since the instance is a YES one, however, the BP algorithm does not prune all feasible nodes due to d_{uw}. By Cor. 5.6 the only remaining possibility (which therefore occurs with probability 1) is that $d_{uw} \in H^{uw}$. □

Next, reflecting single points on realizations yields points that lie on the partial reflection of the whole realization.

5.8 Corollary
Let $y \in X$ *and* $v \in V \setminus V_0$ *such that* $\Upsilon(y, v)$ *holds. If* $u \in V$ *with* $u > v$ *then* $R^v y_u$ *belongs to a valid extension of* $y[U_v]$.

Proof. If there is no edge $\{w, u\} \in E$ with $\rho(u) - \rho(w) > K$ the result follows by Cor. 5.1. Otherwise, by Cor. 5.7, $d_{wu} \in H^{wu}$. As in the proof of Prop. 5.4, all pairs of points that are feasible with respect to d_{wu} are reflections of each other with respect to R^v. □

Finally, we state the main result of the section: if a $^{\mathsf{K}}$DMDGP instance has a valid realization y and v is a vertex where a "valid branching" (in the sense of the $\Upsilon(y, v)$ assumption) takes place in the BP algorithm, then the partial reflection of y with respect to v is also a valid realization. This is surprising as the $\Upsilon(y, v)$ assumption only states that one of the branches at v leads to y, whilst the other might end up at any other valid realization; Thm. 5.9 actually shows that the partial reflection of y with respect to v is valid.

5.9 Theorem
Let $y \in X$ *and* $v \in V \setminus V_0$ *such that* $\Upsilon(y, v)$ *holds. Then* $\tilde{R}^v y \in X$ *with probability 1.*

Proof. We have to show that $\tilde{R}^v y$ is a valid realization for G. Partition E into three subsets E_1, E_2, E_3, where $E_1 = \{\{t,u\} \in E \mid t,u < v\}$, $E_2 = \{\{t,u\} \in E \mid t,u \geq v\}$ and $E_3 = \{\{t,u\} \in E \mid t < v \wedge u \geq v\}$. For E_1, by definition $\|(\tilde{R}^v y)_t - (\tilde{R}^v y)_u)\| = \|Iy_t - Iy_u\| = \|y_t - y_u\| = d_{tu}$ as claimed. For E_2, $\|(\tilde{R}^v y)_t - (\tilde{R}^v y)_u)\| = \|R^v y_t - R^v y_u\| = \|y_t - y_u\| = d_{tu}$ because R^v is an isometry. For E_3, we aim to show that $\|Iy_t - R^v y_u\| = d_{tu}$. Since $y \in X$, by Lemma 4.5 there is a feasible leaf node α with $x(\alpha) = y$. Because $\Upsilon(y,v)$, $\exists \eta \in \mathcal{V}_{\rho(v)-1}$ such that $x(\eta) = y[\gamma(v)]$ and $\mu(\beta) = \boxplus$ for all $\beta \in N^+(\eta)$. Let $\mathsf{p}(\alpha) \cap \mathcal{V}_{\rho(v)} = \{\beta\}$ for some $\beta \in N^+(\eta)$. Again by $\Upsilon(y,v)$, there is at least one feasible leaf node α' such that $\mathsf{p}(\alpha') \cap \mathcal{V}_{\rho(v)} = \{\beta'\}$ for some $\beta' \in N^+(\eta) \smallsetminus \{\beta\}$. Let $\{\omega\} = \mathsf{p}(\alpha) \cap \mathcal{V}_{\rho(u)}$ and $\{\omega'\} = \mathsf{p}(\alpha') \cap \mathcal{V}_{\rho(u)}$. Because ω' is feasible, $\|x(\omega')_t - x(\omega')_u\| = d_{tu}$; because η is an ancestor of both α and α' at level $\rho(v) - 1$ and $t < v$, $\mathsf{p}(\alpha') \cap \mathcal{V}_{\rho(t)} = \mathsf{p}(\alpha) \cap \mathcal{V}_{\rho(t)}$, which implies that $x(\omega')_t = x(\omega)_t = y_t$. Thus, $\|y_t - y_u\| = d_{tu} = \|y_t - x(\omega')_u\|$. Furthermore, because $\beta' \in \mathsf{p}(\omega') \cap \mathcal{V}_{\rho(v)}$, $x(\omega')$ extends $x(\beta')$. By Alg. 1, Steps 25 and 27, $\lambda(\beta) = 1 - \lambda(\beta')$. Because α is feasible, at every vertex $u' \in V$ such that $\rho(v) \leq \rho(u') < \rho(u)$ the node $\theta \in \mathsf{p}(\alpha) \cap \mathcal{V}_{\rho(u')}$ has $f \in \{1,2\}$ feasible subnodes; by Prop. 4.6, the node $\theta' \in \mathsf{p}(\alpha') \cap \mathcal{V}_{\rho(u')}$ also has f feasible subnodes. If $f = 2$, by Cor. 5.8 and Lemma 5.3 it is possible to choose α' so that $\lambda(\theta') = 1 - \lambda(\theta)$ with probability 1; if $f = 1$ then by Alg. 1, Steps 31 and 33, all feasible nodes inherit the same λ value as their parents, so $\lambda(\theta') = 1 - \lambda(\theta)$. By Lemma 5.3, $x(\omega')_u = R^v y_u$ with probability 1. Hence $\|y_t - R^v y_u\| = d_{tu}$ as claimed. □

6. Symmetry and number of solutions

Our strategy for proving that YES instances of the $^{\mathsf{K}}$DMDGP have power of two solutions with probability 1 is as follows. We map realizations $y \in X$ to binary sequences $\chi \in \{0,1\}^n$ describing the "branching path" in the tree $\mathcal{T}$. We define a symmetry operation on χ by flipping its tail from a given component i (this operation is akin to branching at level i, i.e. to a partial reflection applied to realizations). We show that the cardinality of the group of all such symmetries is a power of two by bijection with a set of binary sequences. Finally we prove that the cardinality of the symmetry group is the same as $|X|$.

For all leaf nodes $\alpha \in \mathcal{V}$ with $\mu(\alpha) = \boxplus$ let $\chi(\alpha) = (\lambda(\beta) \mid \beta \in \mathsf{p}(\alpha))$; since realizations in X are also in correspondence with leaf $\boxplus$-nodes of $\mathcal{T}$ by Alg. 1, Step 8, χ defines a relation on $X \times \{0,1\}^n$.

6.1 Lemma
With probability 1, the relation χ is a function.

Proof. For χ to fail to be well-defined, there must exist a realization x which is in relation with two distinct binary sequences χ', χ'', which corresponds to the discriminant of the quadratic equation in the proof of Lemma 4.1 taking value zero at some rank $> K$, which happens with probability 0. □

Let $\Xi = \{\chi(y) \mid y \in X\}$. For $y \in X$ let y^i be its subsequence $(x_1, \ldots, x_i)$.

We extend χ to be defined on all such subsequences by simply setting $\chi^i = (\chi(y)_1, \ldots, \chi(y)_i)$; $\chi(y)$ is valid if y is a valid realization.

Let $N = \{1, \ldots, n\}$ and g be the $n \times n$ binary matrix such that $g_{ij} = 1$ if $i \leq j$ and 0 otherwise (the upper triangular $n \times n$ all-1 matrix); let g_i be its i-th row vector and $\Gamma = \{g_i \mid i \in N\}$. Consider the elementwise modulo-2 addition in the set $\mathbb{F}_2^n$ (denoted $\oplus$): this endows $\mathbb{F}_2^n$ with an additive group structure with identity $e = (0, \ldots, 0)$ where each element is idempotent. Thus, $\mathcal{G} = (\mathbb{F}_2^n, \oplus) \cong C_2^n$. This group naturally acts on itself (and subsets thereof) using the same $\oplus$ operation. It is not difficult to prove that Γ is a set of group generators for $\mathcal{G}$ and a linearly independent set of the vector space $\mathcal{V}$ given by $\mathcal{G}$ with scalar multiplication over $\mathbb{F}_2$. For all $S \subseteq N$, let

$$g_S = \bigoplus_{i \in S} g_i,$$

and define a mapping $\phi : \mathcal{P}(N) \to \mathcal{G}$ given by $\phi(S) = g_S$.

6.2 Lemma
ϕ is injective.

Proof. We show that for all $S, T \subseteq N$, if $g_S = g_T$ then $S = T$.

$$\begin{array}{lcrcl}
 & & g_S & = & g_T \\
 & \Rightarrow & \bigoplus_{i \in S} g_i & = & \bigoplus_{i \in T} g_i \\
 & \Rightarrow & \bigoplus_{i \in S} g_i \oplus \bigoplus_{i \in T} g_i^{-1} & = & e \\
\text{idempotency} & \Rightarrow & \bigoplus_{i \in S} g_i \oplus \bigoplus_{i \in T} g_i & = & e \\
g_i \oplus g_i = g_i^2 & \Rightarrow & \bigoplus_{i \in S \triangle T} g_i \oplus \bigoplus_{i \in S \cap T} g_i^2 & = & e \\
\text{idempotency} & \Rightarrow & \bigoplus_{i \in S \triangle T} g_i & = & e \\
\text{linear independence} & \Rightarrow & S \triangle T & = & \emptyset \\
 & \Rightarrow & S & = & T.
\end{array}$$

This concludes the proof. □

The following result shows essentially that groups of partial reflections have power of two cardinality.

6.3 Lemma
For all $H \subseteq \Gamma$, $|\langle H \rangle| = 2^{|H|}$.

Proof. The restriction of function ϕ to $\mathcal{P}(H)$ is injective by Lemma 6.2. Furthermore, each element g of $\langle H \rangle$ can be written as $\bigoplus_{i \in S} g_i$ for some $S \subseteq H$ because H is a spanning set for the vector space H over $\mathbb{F}_2^n$, which is setwise equal to

the group $\langle H\rangle$. Thus ϕ is surjective too. Hence ϕ is a bijection between $\mathcal{P}(H)$ and $\langle H\rangle$, which yields the result. □

Let I be the set of levels of $\mathcal{T}$ for which from all nodes with two feasible subnodes there is a path going to a feasible leaf through both. Let $L = \{g_i \in \Gamma \mid i \in I\}$ and $\Lambda = \langle L\rangle$ be the subgroup of $\mathcal{G}$ of "allowed partial reflections" generated by L. In the following (the main result of this section) we relate partial reflections to χ representations of valid realizations. We show that any valid realization, in its χ representation, generates the whole set of valid realizations by means of the action of the group of allowed partial reflections.

6.4 Theorem
If $\Xi \neq \emptyset$, for all $\xi \in \Xi$ we have $\xi \oplus \Lambda = \Xi$ with probability 1.

Proof. ($\Rightarrow$) We show that $\xi \oplus \Lambda \subseteq \Xi$ with probability 1; because $\langle L\rangle = \Lambda$ it suffices to show that $\xi \oplus g_i \in \Xi$ for an arbitrary $g_i \in L$, i.e. that there exists a valid realization $w \in X$ such that $\chi(w) = \xi \oplus g_i$. Let $y \in \chi^{-1}(\xi)$ and $v = \rho^{-1}(i)$ such that $\Upsilon(y, v)$, and define $w = \widetilde{R}^v y$; by Thm. 5.9, $w \in X$. Let α' be the leaf node of $\mathcal{T}$ such that $x(\alpha') = y$; by Lemma 4.5, there is a leaf node β' such that $x(\beta') = w$. We have to show that for all $\ell \geq i$ the node $\beta \in \mathsf{p}(\beta') \cap \mathcal{V}_\ell$ is such that $\lambda(\beta) = 1 - \lambda(\alpha)$, where α is the node in $\mathsf{p}(\alpha') \cap \mathcal{V}_\ell$. We proceed by induction on ℓ. For $\ell = i$ this holds by Lemma 4.3. For $\ell > i$, the induction hypothesis allows us to apply Lemma 5.3 and conclude that the event $\lambda(\alpha) = 1 - \lambda(\beta)$ occurs with probability 1.

($\Leftarrow$) Now we show that $\Xi \subseteq \xi \oplus \Lambda$ with probability 1, i.e. for any $\eta \in \Xi$ there is $g \in \Lambda$ with $\xi \oplus g = \eta$. We proceed by induction on n, which starts when $n = K+1$: if $K+1 \notin I$ then $|\Xi| = 1$, $L = \emptyset$ and the theorem holds; if $K+1 \in I$ then $|\Xi| = 2$, $L = \{g_{K+1}\}$ and the theorem holds. Now let $n > K+1$; for all $j \in \{K+1, \ldots, n-1\}$ define $\Xi^j = \{\xi^j \mid \xi \in \Xi\}$ and $L^j = \{g_\ell \in \Gamma \mid \ell \in I \wedge \ell \leq j\}$. By the induction hypothesis, for all $\xi' \in \Xi^j$ ($\xi' \oplus \langle L^j\rangle = \Xi^j$). Now, either $n \notin I$ or $n \in I$; by Prop. 4.6, with probability 1 if $n \notin I$ then nodes in $\mathcal{V}_{n-1}$ can only have zero or one feasible subnode (let B_1^n be the set of all such feasible subnodes), and if $n \in I$ then nodes in $\mathcal{V}_{n-1}$ can only have zero or two feasible subnodes β (let B_2^n be the set of all such feasible subnodes). In the former case we let $\Xi^n = \{\xi(x(\beta)) \mid \beta \in B_1^n\}$ and $L^n = L^{n-1}$; in the latter we let $\Xi^n = \{\xi(x(\beta)) \mid \beta \in B_2^n\}$ and $L^n = L^{n-1} \cup \{g_n\}$. In both cases it is easy to verify that the theorem holds for Ξ^n, L^n: in the former case it follows by the induction hypothesis, and in the latter case it follows because $g_n = (0, \ldots, 0, 1)$, namely, if $\eta \in \Xi$ and $n \in I$ then take $\xi = \eta \oplus g_n$ (the result follows by idempotency of g_n). □

The main result of the paper is now simply a corollary of Thm. 6.4.

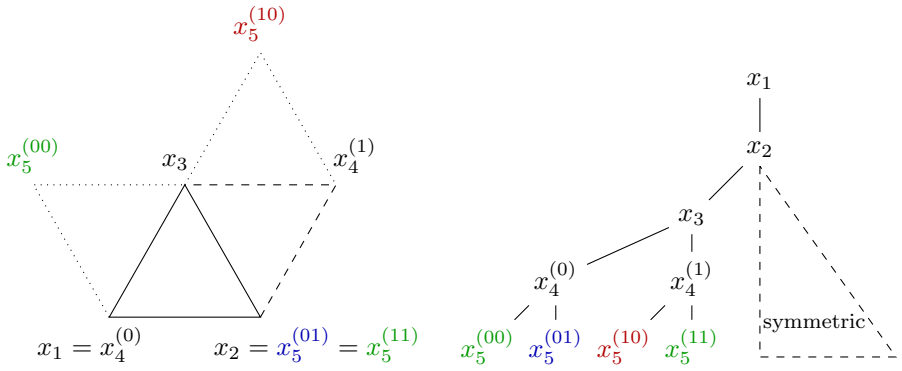
6.5 Corollary
If a $^{\mathsf{K}}$DMDGP instance is YES, $|X|$ is a power of two with probability 1.

Proof. By Lemma 6.1 χ is a function with probability 1. Let $x, x' \in X$ be distinct; then by Alg. 1, Steps 25, 27, 31, and 33, the map $\chi : X \to \Xi$ is injective. By definition of Ξ it is also surjective, hence $|X| = |\Xi|$. By Thm. 6.4 $|\Xi| = |\chi \oplus \Lambda|$ for all $\chi \in \Xi$ with probability 1. It is easy to show that $|\chi \oplus \Lambda| = |\Lambda|$, so by Lemma 6.3 $|X|$ is a power of two with probability 1. □

7. Counterexamples

7.1. Disproving the "power of two" conjecture

We first discuss a class of counterexamples to the conjecture that *all* $^{\mathsf{K}}$DMDGP instances have a number of solutions which is a power of two (also see Lemma 5.1 in [21]). All these counterexamples are hand-crafted and have the property that two distinct realizations x, x' have at least a level i where $x_i = x'_i$, which is an event which happens with probability 0. For any $K \geq 1$, let $n = K + 3$, $V = \{1, \ldots, n\}$, $E = \{\{i, j\} \mid 0 < i - j \leq K\} \cup \{\{1, n\}\}$ and $d_{ij} = 1$ for all $\{i, j\} \in E$. The first $n - 2 = K + 1$ vertices can be realized as the vertices of a regular simplex in dimension K; then either $x_{n-1} = x_1$ or x_{n-1} is the symmetric position from x_1 with respect to the hyperplane through $\{x_2, \ldots, x_{n-2}\}$. In the first case, the two positions for x_n are valid, in the second only $x_n = x_2$ is possible (see Fig. 7 for the 2-dimensional case), yielding a YES instance where $|X| = 6$.



(a) Positions of the points on the plane. (b) BP tree.

Figure 7: The counterexample in the case $K = 2$. Realizations $x_5^{(00)}$, $x_5^{(01)}$, and $x_5^{(11)}$ are valid, while $x_5^{(10)}$ is not.

7.2. Necessity of immediate predecessors

Lastly, Fig. 8 shows an example where the $(ii) \Rightarrow (i)$ implication of Lemma 5.3 fails for instances in DDGP$\setminus^{\mathsf{K}}$DMDGP. This shows that any generalization

of our result to the DDGP must be nontrivial. Let $V = \{1, \ldots, 6\}$ (the graph drawing is the same as the realization in $\mathbb{R}^2$). The points $5', 6'$ linked with dashed lines show alternative placements for the corresponding vertices. Let $U_5 = \{3, 4\}$ and $U_6 = \{1, 2\}$. The line through the points $3, 4$ does not provide a valid reflection mapping 6 to $6'$. This happens because U_6 does not consist of the two *immediate* predecessors of vertex 6.

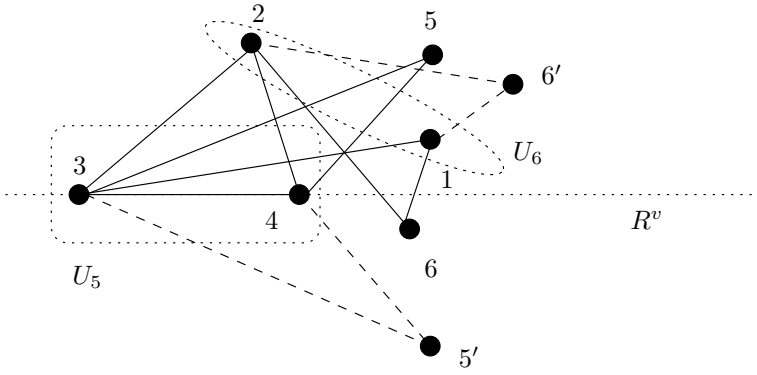


Figure 8: A counterexample to Lemma 5.3 applied to DDGP $\setminus {}^{\mathsf{K}}$DMDGP.

8. Application of the theory to protein conformation

The "*interval*BP algorithm" (iBP) is the adaptation of the BP algorithm to a more realistic setting for protein backbones [23]. For $v \in V$, we shall denote the vertex of rank $\rho(v) - h$ (for some $h < \rho(v)$) by $v - h$ (in other words, we identify the vertex labels with their own ranks in the order). We note that throughout this section K is fixed to 3; we still use the symbolic notation K to emphasize results that hold for whatever value of K.

As mentioned in the introduction, covalent bonds and angles are known reasonably precisely. These allow us to consider the distances $d_{v-2,v}$ and $d_{v-1,v}$ as precise for any v with rank greater than K. The data provided by NMR actually consists of a frequency measurement for a triplet (a, b, δ) where a, b are atom types (e.g. H, C, O and so on) and δ is a distance value [15]. In other words, NMR gives experimental evidence that the distance δ is to be expected between atoms of type a, b with a certain probability. NMR specialists then solve an assignment problem so that they can decide which pair(s) of atoms of type a, b should be assigned the distance δ. This procedure gives rise to errors, due to which all distances are actually reported as intervals rather than precise values. We emphasize two important observations:

1. By default, NMR experiments are rigged up to find the distances between pairs of *hydrogen* atoms only [43]. Distances between pairs of atoms of

different types are possible but they require more work, and are prone to more errors.

2. The NMR machinery has a resolution limit of around about one decimal digit (in Å units); if the distance interval is on a similar scale, only a finite number of values belonging to the intervals can actually be measured [39].

We initially exploited Observation 1 by defining certain vertex orders that only include hydrogens [27]; following a similar methodology, we later defined another vertex order (called the iBP order) which also takes into account Observation 2 [23]. More precisely, for each $v \in V$ of rank greater than 3, this order is such that:

- $d_{v-2,v}, d_{v-1,v}$ are single values in $\mathbb{R}_+$;
- $d_{v-3,v}$ is either a single value, or a finite set $\{d^1_{uv}, \ldots, d^q_{uv}\}$ (with $q > 1$) of values in $\mathbb{R}_+$,
- d_{uv} is either a single value, or an interval $[d^L_{uv}, d^U_{uv}]$ for all $\rho(u) < \rho(v) - 3$.

Essentially, this order is such that all interval distances are only used for pruning purposes, whilst the discretization of the search space occurs differently because $d_{v-3,v}$ could be a finite set of values.

8.1. Discretization with a finite set of values

In this section we go through the previous results and adapt them to the iBP setting insofar as $d_{v-3,v}$ might be a set instead of a single value. We still consider precise distances for pruning rather than intervals (this further extension will be tackled in Sect. 8.2). Before going into the details, the intuitive explanation as to why most results still hold (albeit with some modifications) is that one can see the iBP search tree as the union of q^k different BP trees, where k is the number of vertices $v \in V$ such that $d_{v-3,v}$ is a set. Thus the iBP tree inherits many of the properties of all the BP trees that compose it.

8.1.1. Changes to Sect. 4

Assume $v \in V$ is such that $d_{v-3,v}$ is a set of q values and every other distance is precise. Then Lemma 4.1 changes as follows.

8.1 Lemma
If $|N(v) \cap \gamma(v)| = 3$ then there are at most $2q$ distinct extensions of x that are valid for G^v. If one valid extension exists, then with probability 1 there are exactly $2q$ distinct valid extensions.

Proof. We shall only point out the differences with the proof of Lemma 4.1. We choose the vertex $w \in U_v$ such that $w - 3$; then Eq. (5) becomes $\forall u \in U_v \smallsetminus \{w\}$ $(\|z - x_u\| = d_{uv}) \ \wedge \ (\|z - x_w\| \in d_{wv})$, so that the second line

in Eq. (6) becomes $\|z\|^2 - 2x_w \cdot z + \|x_w\|^2 \in d_{wv}^2$, where d_{wv}^2 denotes the set $\{(d_{wv}^j)^2 \mid j \leq q\}$. We rewrite this as q different quadratic equations

$$\|z\|^2 - 2x_w \cdot z + \|x_w\|^2 = d_{wv}^j \tag{10}$$

for each $j \leq q$. From each of these, by Lemma 4.1 we obtain two distinct values for z with probability 1, which concludes the proof. □

Lemma 4.2 involves adjacent but not immediate predecessors. The statement is unchanged, but the proof is slightly different.

8.2 Lemma
If $|N(v) \cap \gamma(v)| > 3$ then, with probability 1, there is at most 1 extension of x.

Proof. Since $U_v \subseteq N(v) \cap \gamma(v)$, only considering vertices in U_v we are in the situation of Lemma 8.1 and we have $2q$ possible extensions of x. Now consider an adjacent predecessor u of v which is not in U_v: with probability 1, the distances $\|x_u - z^{1j}\|$, $\|x_u - z^{1j}\|$ are all distinct (for all $j \leq q$). Hence, at most one can be equal to d_{uv}. □

Lemma 4.3 can be adapted in the following way.

8.3 Lemma
Let $\bar{x}$ be any valid realization of $G[U_v]$ and $\{z^{1j}, z^{2j} \mid j \leq q\} \subseteq \mathbb{R}^3$ be the $2q$ positions for vertex v given by Lemma 8.1, with z^{1j}, z^{2j} arising from the Eq. 10 indexed by j. Then z^{2j} is a reflection of z^{1j} with respect to the hyperplane through the 3 positions of $\bar{x}$.

The proof is the same as that of Lemma 4.3, for each $j \leq q$.

The change to these Lemmata impacts Alg. 1 in several respects. The most important change is that whenever $d_{v-3,v}$ is a set and v has exactly 3 adjacent predecessors, the BP tree node α at level $v-1$ has exacty $2q$ subnodes α^{ij} (for $i \in \{1,2\}$ and $j \leq q$) at level v. Accordingly, for all $i \in \{1,2\}$ and $j \leq q$, $\lambda(\alpha^{ij})$ is defined as follows:

$$\lambda(\alpha^{ij}) = \begin{cases} -j & \text{if} \quad a_v^\top z^{ij} \leq a_{v0} \\ j & \text{if} \quad a_v^\top z^{ij} > a_{v0} \end{cases} \tag{11}$$

The set Z in Step 17 of Alg. 1 is $\{z^{ij} \mid i \in \{1,2\} \wedge j \leq q\}$. The test at Step 23 is $|Z| = 2q$. The body of the corresponding **if** changes in the following way:

for $j \leq q$ **do**
  if $a_v^\top z^{1j} \leq a_{v0}$ **then**
    Let $\alpha^{1j} = (z^{1j}, -j, \boxplus)$, $\alpha^{2j} = (z^{2j}, j, \boxplus)$
  else
    Let $\alpha^{1j} = (z^{1j}, j, \boxplus)$, $\alpha^{2j} = (z^{2j}, -j, \boxplus)$
  end if
end for

It is easy to show that the full iBP search tree is no longer a binary one: every time $d_{v-3,v}$ is a set, level each node at level $v-1$ has $2q$ subnodes. The total number of iBP nodes depends on the incidence of set distances within the $^{\mathsf{K}}$DMDGP vertex order.

Lemma 4.5 does not change. Prop. 4.6 changes as follows.

8.4 Proposition
With probability 1, there is no level $i \leq n$ having two distinct feasible nodes $\beta, \theta \in \mathcal{V}_i$ such that $|\{\alpha \in N^+(\beta) \mid \mu(\alpha) = \boxplus\}| = 1$ and $|\{\alpha \in N^+(\theta) \mid \mu(\alpha) = \boxplus\}| = 2q$.

The proof is trivially adapted from that of Prop. 4.6.

8.1.2. Changes to Sect. 5

The statement of Cor. 5.1 changes as follows.

8.5 Corollary
Let $\alpha \in \mathcal{V}_{i-1}$ for some $i > 1$, $v = \rho^{-1}(i)$ and $N^+(\alpha) = \{\eta^{ij} \mid i \in \{1,2\} \wedge j \leq q\}$ with $\mu(\eta^{ij}) = \boxplus$ for all $i \in \{1,2\}$, $j \leq q$. Then $x(\eta^{1j})_v = R^v x(\eta^{2j})_v$.

Only Condition (ii) changes in Lemma 5.3:

> (ii) $\forall\, i \leq j \leq \ell$ if d_{uw} is a single value, then Condition (ii) in Lemma 5.3 holds; if d_{uw} is a set, $\lambda(\eta'') = -\lambda(\beta'')$.

The proof is an easy adaptation of that of Lemma 5.3: every time λ is mentioned, it suffices to verify whether the corresponding distance d_{uw} is a value or a set, and use the correct definition (either Eq. (9) or Eq. (11)). Prop. 5.4 changes its statement as follows.

8.6 Proposition
Consider a subtree $\mathcal{T}'$ of $\mathcal{T}$ consisting of $K+2$ consecutive levels $i-K-1, \ldots, i$ (where $i \geq 2K+1$), rooted at a single node η and such that all nodes at all levels are marked $\boxplus$. Let $u = \rho^{-1}(i-K-1)$, $w = \rho^{-1}(i-K)$, $t = \rho^{-1}(i-1)$ and $v = \rho^{-1}(i)$. If p is the number of leaf nodes of $\mathcal{T}'$, let $Y' = \{y^j \mid j \leq p\}$ be the set of partial realizations of G at the such leaf nodes.

1. *If d_{ut} and d_{wv} are single values, Lemma 5.4 holds.*
2. *If d_{ut} is a set of q values and d_{wv} is a single value or vice versa, let $p = 2^{K+1}q$; with probability 1 there is a set $H^{uv} \subseteq \mathbb{R}_+$ with $|H^{uv}| = 2q$ such that $\|y_u^j - y_v^j\| \in H^{uv}$ for each $j \leq p$.*
3. *If both d_{ut}, d_{wv} are sets of q values, let $p = 2^{K+1}q^2$; with probability 1 there is a set $H^{uv} \subseteq \mathbb{R}_+$ with $|H^{uv}| = 2q^2$ such that $\|y_u^j - y_v^j\| \in H^{uv}$ for each $j \leq p$.*

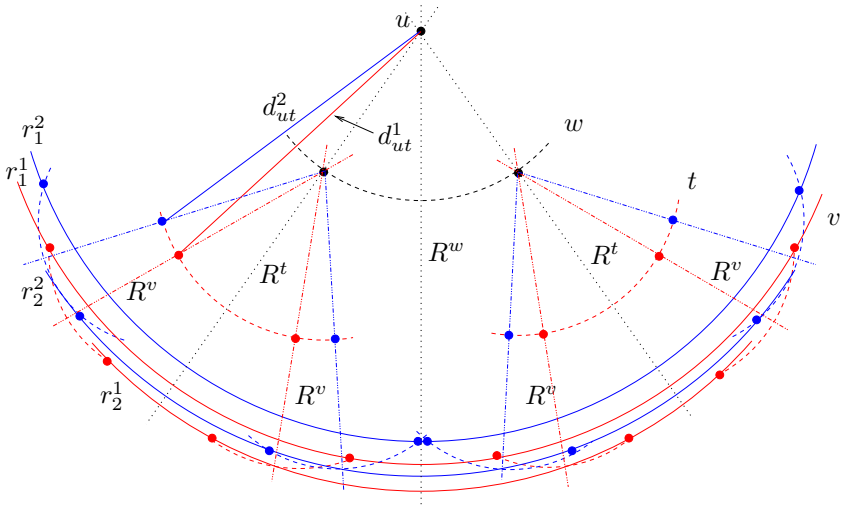


Figure 9: Proof of Prop. 5.4 adapted to the iBP setting in $\mathbb{R}^2$, with d_{ut} having cardinality 2 and d_{wv} being a single value. The arcs of circles r^j_h ($j \leq 2$, $h \in \{1,2\}$) centered at u mark the four possible distances between the position of u and that of v.

Proof. The situation in $\mathbb{R}^2$ is shown in Fig. 9 for the case where d_{ut} is a set of 2 values and d_{wv} is a single value. Only two edges, namely $\{u,t\}$ and $\{w,v\}$, can be weighted by sets of distance values in the subtree $\mathcal{T}'$. If $\{u,t\}$ is weighted by a set but $\{w,v\}$ is not, then at level t we have $2q$ as many nodes as in the previous level; if $\{w,v\}$ is weighted by a set but $\{u,t\}$ is not, then at level v we have $2q$ as many nodes as in level t; if both are weighted by sets, then at both levels the increase in the number of nodes is $2q$-fold. At any level $s \in \{t,v\}$ where there is an increase of $2q$ nodes w.r.t. the previous level, each pair of nodes α^{1j}, α^{2j} with $\lambda(\alpha^{1j}) = -\lambda(\alpha^{2j})$ (for $j \leq p$) is such that $x(\alpha^{1j})_s$ and $x(\alpha^{2j})_s$ are reflections through R^w by the adaptation of Lemma 5.3 (above). Hence each such pair yields a subtree of $\mathcal{T}'$ which, by Prop. 5.4, with probability 1 allows $\|y^j_u - y^j_v\|$ to only take values in a set $S \subseteq \mathbb{R}_+$ such that $|S| = 2$. Observe that any pair of such sets S will have non-empty intersection with probability 0. The result follows. □

Cor. 5.6 can be extended to the iBP setting by remarking that the set H^{uv} (appearing in its statement) has cardinality $2^h q^k$, where k is the number of vertices t between u and v such that $d_{t-3,t}$ is a set. Cor. 5.7-5.8 are unchanged.

The statement of Thm. 5.9 is unchanged, but the proof needs to be adapted. The statement $\lambda(\beta) = 1 - \lambda(\beta')$ must be changed to $\lambda(\beta) = -\lambda(\beta')$ by the adaptation of Alg. 1 to the iBP given above in this section. The statement $f \in \{1,2\}$ must be changed to "$f \in \{1,2\}$ if $d_{u'-3,u'}$ is a single value, and $f \in \{1,2q\}$ if $d_{u'-3,u'}$ is a set". The motivation for θ' to also have f subnodes is by Prop. 4.6 if $d_{u'-3,u'}$ is a single value, and by Prop. 8.4 if $d_{u'-3,u'}$ is a set.

Between the cases $f = 2$ and $f = 1$, the case $f = 2q$ must be added, with the comment that it is possible to choose α' so that $\lambda(\theta') = -\lambda(\theta)$ with probability 1 (by Lemma 5.3 adapted to the iBP setting). The case $f = 1$ must be changed so that if $d_{u'-3,u'}$ is a set, then $\lambda(\theta') = -\lambda(\theta)$.

8.1.3. Number of solutions

Since the iBP branches over $2q$ possibilities at some levels, in general the number of possible realization is no longer a power of two, but rather a number $2^\ell q^k$ for some integers ℓ, k. Because of this fundamental difference, we shall not attempt to adapt each result in Sect. 6 to the iBP setting, but rather propose a somewhat different development.

For all $v \in V$, let $\psi_v = \{0\}$ whenever $v \leq K$, $\psi_v = \{0, 1\}$ whenever $d_{v-3,v}$ is a single value, and $\psi_v = \{0, \ldots, 2q-1\}$ whenever $d_{v-3,v}$ is a set. Accordingly, we define χ to map realizations in X to sequences indexed by V, in such a way that $\chi_v \in \psi_v$ for all $v \in V$. More precisely, if α is an iBP tree node such that $x(\alpha)_v = y_v$ and $d_{v-3,v}$ is a set, then $\chi_v = \lambda(\alpha) - 1$ whenever $\lambda(\alpha) > 0$, and $\chi_v = q - \lambda(\alpha) - 1$ whenever $\lambda(\alpha) < 0$. With this definition, the set $\Xi = \{\chi(y) \mid y \in X\}$ can, with probability 1, be endowed with an Abelian group structure: $\chi^1 \oplus \chi^2 = (\chi^1_v \oplus \chi^2_v \mid v \in V)$, where $\oplus$ denotes addition modulo $|\psi_v|$. We remark that the group $\mathcal{G} = \langle \Xi, \oplus \rangle$ is a subgroup of $\bar{\mathcal{G}} = \prod_{v \in V} C_{|\psi_v|}$. Furthermore, since the underlying set of $\mathcal{G}$ is Ξ, $\mathcal{G}$ acts on itself: this yields a regular action, which is transitive by definition. Thus, for any $\chi \in \Xi$, we have $\chi \oplus \Xi = \Xi$.

For all $v \in V$ such that $\rho(v) > 3$, if $d_{v-3,v}$ is a single value, let $\iota_v = 1$; if $d_{v-3,v}$ is a set, let $\iota_v = q$. Consider the set $\bar{\Lambda}$ of the following elements of $\bar{\mathcal{G}}$:

- $\pi_v = (0, \ldots, 0, \iota_v, \ldots, \iota_n)$ for all v such that $\rho(v) > 3$;
- $\sigma_v = (0, \ldots, 0, 1_v, 0, \ldots, 0)$, where the 1 is in the $\rho(v)$-th position, for all v such that $d_{v-3,v}$ is a set.

It is easy to show that $\bar{\Lambda}$ generates the whole of $\bar{\mathcal{G}}$. Now, similarly to Sect. 6, we find a subset Λ of $\bar{\Lambda}$ which generates $\mathcal{G}$. More precisely:

- if $v \in V$ is such that $\rho(v) > 3$, $d_{v-3,v}$ is a single value, and for all iBP nodes at level v with two feasible subnodes, both subnodes are on paths continuing to leaf nodes, then $\pi_v \in \Lambda$;
- if $v \in V$ is such that $\rho(v) > 3$, $d_{v-3,v}$ is a set and for all iBP nodes at level v with $2q$ feasible subnodes, both subnodes are on paths continuing to leaf nodes, then $\pi_v, \sigma_v \in \Lambda$.

We are now in a position to prove a theorem whose significance is similar to that of Thm. 6.4.

8.7 Theorem
With probability 1, $\langle \Lambda, \oplus \rangle = \mathcal{G}$.

Proof. That $\langle \Lambda, \oplus \rangle \subseteq \mathcal{G}$ follows by definition: π_v encodes a valid branching, and the elements $\pi_v \oplus (\sigma_v \oplus \cdots \sigma_v)$ represent the $2q$ branches in the case of set distances. The other direction is similar to the ($\Leftarrow$) direction of the proof of Thm. 6.4. Replace "zero or two feasible subnodes" by "zero, two or $2q$ feasible subnodes", and $\{g_n\}$ by $\{\pi_n\}$ if there are two feasible subnodes at level n, or by $\{(\pi_n)^k \oplus (\sigma_v)^h \mid k \in \{0,1\} \wedge 0 \leq h \leq q-1\}$ otherwise, where $\alpha^f = \sum_{s<f} \alpha$ for any integer f and $\alpha \in \mathcal{G}$. □

Thm. 8.7, like Thm. 6.4, can be used to compute the whole of X having just one $x \in X$ and the partial reflection operators in $\mathcal{G}$.

Let M be the number of vertices v with $\rho(v) > 3$ such that $d_{v-3,v}$ is a set. It is easy to see that the cardinality of $\langle \Lambda, \oplus \rangle$ is $2^\ell q^k$ for some $\ell \leq n-3$ and $k \leq M$, since it is a subgroup of $\bar{\mathcal{G}}$. By Thm. 8.7, with probability 1, that is the same cardinality of $\mathcal{G}$, and therefore of Ξ.

8.2. Pruning by intervals

If $z \in \mathbb{R}^K$ is a candidate position for vertex v, and u is a non-immediate adjacent predecessor of v (i.e., such that $\rho(u) < \rho(v) - 3$), then d_{uv} may be the interval $[d^L_{uv}, d^U_{uv}]$, and the test on Line 19 of Alg. 1 becomes $\|x(\beta)_u - z\| \notin [d^L_{uv}, d^U_{uv}]$.

It was shown in Prop. 5.4, Cor. 5.6, Prop. 8.6, and the adaptation of Cor. 5.6 to distance sets in Sect. 8.1.2 that with probability 1, for $u, v \in V$ with $\rho(v) - \rho(u) > K$, there are are sets $H^{uv} \subseteq \mathbb{R}_+$ of an even number of possible distance values that $\|y_u - y_v\|$ can take for any $y \in X$. Furthermore, with probability 1, for any distance value $r \in H^{uv}$ and partial realization $(y_1, \ldots, y_u)$ there are two possible extensions y, y' to a partial realization on $G[\gamma(v) \cup \{v\}]$ such that $\|y_u - y_v\| = r$; moreover, it turns out that we have $y_w = R^{u+1} y'_w$ for all $w \in V$ such that $\rho(u) < \rho(w) \leq \rho(v)$. By the results of Sect. 5, a single value $d_{uv} = r$ weighting an edge $\{u, v\} \in E$ prunes all possible realizations for v apart from those two for which $\|y_u - y_v\| = r$. Now, if d_{uv} is an interval, then it can contain any number of consecutive values in H^{uv}, and as such it can prune any even number of possible realizations for v (including zero), as shown graphically in Fig. 10.

It is therefore difficult to envisage an aprioristic calculation of the number of solutions $|X|$ of a given instance in this case. For YES instances, it is clear that the result obtained in Sect. 8.1.3 provides a lower bound to this number; and evidently the cardinality of $\bar{\mathcal{G}}$ provides an upper bound. Thus, there exist $\ell \leq n-3$ and $k \leq M$ (which depend on vertices v adjacent to edges $\{u, v\}$ where $u < v - 3$) such that:

$$2^\ell q^k \leq |X| \leq 2^{n-3} q^M. \tag{12}$$

8.3. Distance sets of different cardinalities

Although we assumed that, for all $v \in V$ such that $\rho(v) > 3$ and $d_{v-3,v}$ is a set of values, the cardinality of $|d_{v-3,v}|$ is always q, this may fail to be the case; we shall let $q_v = |d_{v-3,v}|$ for all such v (for the other $v \in V$, we define $q_v = 1$).

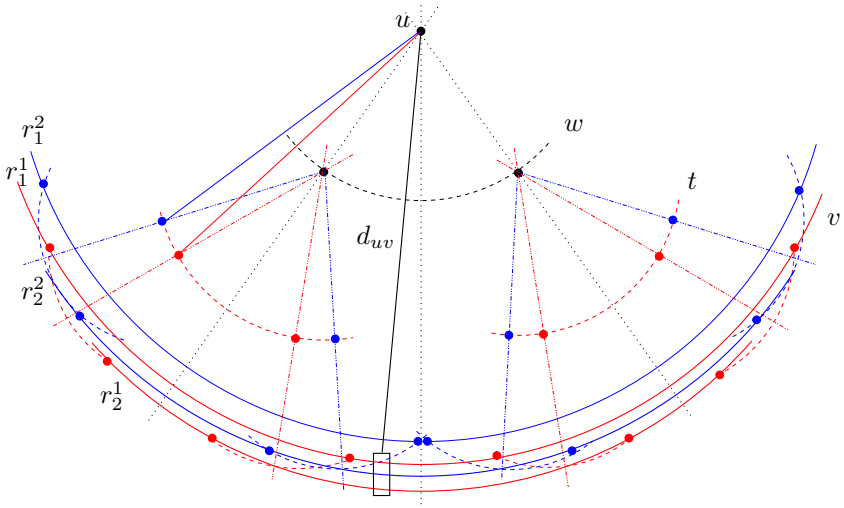


Figure 10: The interval distance d_{uv} makes 3 out of 4 possible values from H^{uv} feasible, namely r_1^1, r_1^2, r_2^2 (and hence 12 out of 16 possible positions for vertex v).

Furthermore, we let $\epsilon \in \{0,1\}^n$ such that $\epsilon_v = 0$ for all $\rho(v) \leq 3$ and all $\rho(v)$ such that $d_{v-3,v}$ is a single value. The other components of v will depend on whether v has K or more than K adjacent predecessors. It is not difficult to adapt the results of Sect. 8.1 in order to show that Eq. 12 becomes:

$$2^\ell \prod_{v \in V} q_v^{\epsilon_v} \leq |X| \leq 2^{n-3} \prod_{v \in V} q_v. \tag{13}$$

Both lower and upper bounds are modifications of Eq. (12), where q^k has been replaced by the product of the values q_v when v ranges over the relevant vertices of V.

9. Conclusion

This paper presents a mathematical theory of non-unique rigidity in connection to symmetry of certain graphs arising in protein conformation problems from NMR data. Further work will focus on a computational application of these results, where we shall exploit Thm. 6.4 in order to compute the set of *all* incongruent realizations of a given $^{\mathsf{K}}$DMDGP graph using just *one* realization thereof.

References

[1] Barvinok, A., 1995. Problems of distance geometry and convex properties of quadratic maps. Discrete and Computational Geometry 13, 189–202.

[2] Berger, B., Kleinberg, J., Leighton, T., 1999. Reconstructing a three-dimensional model with arbitrary errors. Journal of the ACM 46 (2), 212–235.

[3] Biswas, P., Lian, T., Wang, T., Ye, Y., 2006. Semidefinite programming based algorithms for sensor network localization. ACM Transactions in Sensor Networks 2, 188–220.

[4] Blumenthal, L., 1953. Theory and Applications of Distance Geometry. Oxford University Press, Oxford.

[5] Brady, T., Watt, C., 2006. On products of Euclidean reflections. American Mathematical Monthly 113, 826–829.

[6] Connelly, R., 2005. Generic global rigidity. Discrete Computational Geometry 33, 549–563.

[7] Coope, I., 2000. Reliable computation of the points of intersection of n spheres in $\mathbb{R}^n$. Australian and New Zealand Industrial and Applied Mathematics Journal 42, C461–C477.

[8] Cremona, L., 1872. Le figure reciproche nella statica grafica. G. Bernardoni, Milano.

[9] Crippen, G., Havel, T., 1988. Distance Geometry and Molecular Conformation. Wiley, New York.

[10] Dattorro, J., 2005. Convex Optimization and Euclidean Distance Geometry. $\mathcal{M}\epsilon\beta oo$, Palo Alto.

[11] Dong, Q., Wu, Z., 2003. A geometric build-up algorithm for solving the molecular distance geometry problem with sparse distance data. Journal of Global Optimization 26, 321–333.

[12] Eren, T., Goldenberg, D., Whiteley, W., Yang, Y., Morse, A., Anderson, B., Belhumeur, P., 2004. Rigidity, computation, and randomization in network localization. IEEE Infocom Proceedings, 2673–2684.

[13] Graver, J., 1991. Rigidity matroids. SIAM Journal on Discrete Mathematics 4, 355–368.

[14] Graver, J., Servatius, B., Servatius, H., 1993. Combinatorial Rigidity. American Mathematical Society.

[15] Gunther, H., 1995. NMR Spectroscopy: Basic Principles, Concepts, and Applications in Chemistry. Wiley, New York.

[16] Hendrickson, B., 1992. Conditions for unique graph realizations. SIAM Journal on Computing 21 (1), 65–84.

[17] Henneberg, L., 1911. Die Graphische Statik der starren Systeme. Teubner, Leipzig.

[18] John, A. L.-S., 2008. Geometric constraint systems with applications in cad and biology. Ph.D. thesis, University of Massachusetts at Amherst.

[19] Kang, R., Müller, T., 2011. Sphere and dot product representations of graphs. In: Proceedings of SCG11. ACM, pp. 308–314.

[20] Lavor, C., Lee, J., John, A. L.-S., Liberti, L., Mucherino, A., Sviridenko, M., 2012. Discretization orders for distance geometry problems. Optimization Letters 6, 783–796.

[21] Lavor, C., Liberti, L., Maculan, N., 2006. The discretizable molecular distance geometry problem. Tech. Rep. q-bio/0608012, arXiv.

[22] Lavor, C., Liberti, L., Maculan, N., Mucherino, A., 2012. The discretizable molecular distance geometry problem. Computational Optimization and Applications 52, 115–146.

[23] Lavor, C., Liberti, L., Mucherino, A., DOI:10.1007/s10898-011-9799-6. The *interval* Branch-and-Prune algorithm for the discretizable molecular distance geometry problem with inexact distances. Journal of Global Optimization.

[24] Lavor, C., Mucherino, A., Liberti, L., Maculan, N., 2009. An artificial backbone of hydrogens for finding the conformation of protein molecules. In: Proceedings of the Computational Structural Bioinformatics Workshop. IEEE, Washington D.C., USA, pp. 152–155.

[25] Lavor, C., Mucherino, A., Liberti, L., Maculan, N., 2009. Computing artificial backbones of hydrogen atoms in order to discover protein backbones. In: Proceedings of the International Multiconference on Computer Science and Information Technology. IEEE, Mragowo, Poland, pp. 751–756.

[26] Lavor, C., Mucherino, A., Liberti, L., Maculan, N., 2010. Discrete approaches for solving molecular distance geometry problems using NMR data. International Journal of Computational Biosciences 1, 88–94.

[27] Lavor, C., Mucherino, A., Liberti, L., Maculan, N., 2011. On the computation of protein backbones by using artificial backbones of hydrogens. Journal of Global Optimization 50, 329–344.

[28] Liberti, L., Lavor, C., Maculan, N., 2008. A branch-and-prune algorithm for the molecular distance geometry problem. International Transactions in Operational Research 15, 1–17.

[29] Liberti, L., Lavor, C., Mucherino, A., Maculan, N., 2010. Molecular distance geometry methods: from continuous to discrete. International Transactions in Operational Research 18, 33–51.

[30] Liberti, L., Masson, B., Lee, J., Lavor, C., Mucherino, A., 2011. On the number of solutions of the discretizable molecular distance geometry problem. In: Combinatorial Optimization, Constraints and Applications (COCOA11). Vol. 6831 of LNCS. Springer, New York, pp. 322–342.

[31] Menger, K., 1928. Untersuchungen über allgemeine Metrik. Mathematische Annalen 100, 75–163.

[32] Mucherino, A., Lavor, C., 2009. The branch and prune algorithm for the molecular distance geometry problem with inexact distances. In: Proceedings of the International Conference on Computational Biology. Vol. 58. World Academy of Science, Engineering and Technology, pp. 349–353.

[33] Mucherino, A., Lavor, C., Liberti, L., 2012. Exploiting symmetry properties of the discretizable molecular distance geometry problem. Journal of Bioinformatics and Computational Biology 10, 1242009(1–15).

[34] Mucherino, A., Lavor, C., Liberti, L., DOI:10.1007/s11590-011-0358-3. The discretizable distance geometry problem. Optimization Letters.

[35] Mucherino, A., Lavor, C., Liberti, L., Maculan, N., 2009. On the definition of artificial backbones for the discretizable molecular distance geometry problem. Mathematica Balkanica 23, 289–302.

[36] Mucherino, A., Lavor, C., Liberti, L., Talbi, E.-G., 2010. On suitable parallel implementations of the branch & prune algorithm for distance geometry. In: Proceedings of the Grid5000 Spring School. Lille, France.

[37] Mucherino, A., Lavor, C., Liberti, L., Talbi, E.-G., 2010. A parallel version of the branch & prune algorithm for the molecular distance geometry problem. In: ACS/IEEE International Conference on Computer Systems and Applications (AICCSA10). IEEE, Hammamet, Tunisia, pp. 1–6.

[38] Mucherino, A., Liberti, L., Lavor, C., 2010. `MD-jeep`: an implementation of a branch-and-prune algorithm for distance geometry problems. In: Fukuda, K., van der Hoeven, J., Joswig, M., Takayama, N. (Eds.), Mathematical Software. Vol. 6327 of LNCS. Springer, New York, pp. 186–197.

[39] Nilges, M., Macias, M., O'Donoghue, S., Oschkinat, H., 1997. Automated NOESY interpretation with ambiguous distance restraints: The refined NMR solution structure of the Pleckstrin homology domain from β-spectrin. Journal of Molecular Biology 269, 408–422.

[40] Saviotti, C., 1885. Nouvelles méthodes pour le calcul des travures réticulaires. In: Appendix to L. Cremona, "Les figures réciproques en statique graphique". Gauthier-Villars, Paris, pp. 37–100.

[41] Saviotti, C., 1888. La statica grafica: Lezioni. U. Hoepli, Milano.

[42] Saxe, J., 1979. Embeddability of weighted graphs in k-space is strongly **NP**-hard. Proceedings of 17th Allerton Conference in Communications, Control and Computing, 480–489.

[43] Schlick, T., 2002. Molecular modelling and simulation: an interdisciplinary guide. Springer, New York.

[44] Schoenberg, I., 1935. Remarks to Maurice Fréchet's article "Sur la définition axiomatique d'une classe d'espaces distanciés vectoriellement applicable sur l'espace de Hilbert". Annals of Mathematics 36 (3), 724–732.

[45] So, A. M.-C., Ye, Y., 2007. Theory of semidefinite programming for sensor network localization. Mathematical Programming B 109, 367–384.

[46] Tay, T.-S., Whiteley, W., 1985. Generating isostatic frameworks. Structural Topology 11, 21–69.

[47] Whiteley, W., 1984. Infinitesimally rigid polyhedra. I. Statics of frameworks. Transactions of the American Mathematical Society 285 (2), 431–465.

[48] Whiteley, W., 1988. Infinitesimally rigid polyhedra. II: Modified spherical frameworks. Transactions of the American Mathematical Society 306 (1), 115–139.

[49] Zhu, Z., So, A. M.-C., Ye, Y., 2010. Universal rigidity and edge sparsification for sensor network localization. SIAM Journal on Optimization 20 (6), 3059–3081.