Adv. Appl. Clifford Algebras (2019) 29:9

0188-7009/010001-15
published online December 3, 2018
https://doi.org/10.1007/s00006-018-0925-0

Advances in Applied Clifford Algebras

Oriented Conformal Geometric Algebra and the Molecular Distance Geometry Problem

Carlile Lavor* and Rafael Alves

Abstract. The problem of 3D protein structure determination using distance information from nuclear magnetic resonance (NMR) experiments is a classical problem in distance geometry. NMR data and the chemistry of proteins provide a way to define a protein backbone order such that the distances related to the pairs of atoms $\{i-3,i\},\{i-2,i\},\{i-1,i\}$ are available, implying a combinatorial method to solve the problem, called branch-and-prune (BP). There are two main steps in BP algorithm: the first one (the branching phase) is to intersect three spheres centered at the positions for atoms $i-3, i-2, i$, with radius given by the atomic distances $d_{i-3,i}, d_{i-2,i}, d_{i-1,i}$, respectively, to obtain two possible positions for atom i; and the second one (the pruning phase) is to check if additional spheres (related to distances $d_{j,i}$, $j < i-3$) can be used to select one of the two possibilities for atom i. Differently from distances $d_{i-2,i}, d_{i-1,i}$ (associated to bond lenghts and bond angles), distances $d_{j,i}$, $j \leq i-3$, may not be precise. BP algorithm has difficulties to deal with uncertainties, and this paper proposes the oriented conformal geometric algebra to take care of intersection of spheres when their centers and radius are not precise.

Keywords. Oriented conformal geometric algebra, Distance geometry, Branch and prune algorithm, 3D protein structure.

1. The Molecular Distance Geometry Problem

The origin of distance geometry (DG) is due to Menger [34], in 1928, when he characterized geometric concepts using the idea of distance. The fundamental problem of DG is how to determine spatial positions for a set of points, whose distances among some of them are known [30,32,33].

There are many applications of DG [4,5,36] and one of the most important is related to molecular geometry [24], where the problem is the calculation of 3D protein structures using Nuclear Magnetic Resonance (NMR)

This article is part of the Topical Collection on Homage to Prof. W.A. Rodrigues Jr edited by Jayme Vaz Jr..

*Corresponding author.

experiments, which provide distances between hydrogen atoms that are close enough [10,40].

We can formally define the problem, called the *Molecular Distance Geometry Problem* (MDGP) [30], using a graph $G = (V, E, d)$, where V represents the set of atoms, E represents the set of atom pairs for which a distance is available, and $d : E \to (0, \infty)$ is the function that assigns non-negative real numbers to each pair $\{u, v\} \in E$ (we say that G is *simple* and *undirected* when, respectively, if $\{u, v\} \in E$ then $u \neq v$, and $\{u, v\} = \{v, u\}$):

Definition 1.1. Given a simple undirected graph $G = (V, E, d)$ whose edges are weighted by $d : E \to (0, \infty)$, find a function $x \colon V \to \mathbb{R}^3$ such that

$$\forall \{u, v\} \in E, \ ||x_u - x_v||_2 = d_{u,v}, \tag{1.1}$$

where $x_u = x(u)$, $x_v = x(v)$, and $d_{u,v} = d(\{u, v\})$.

Using NMR data and the chemistry of proteins [25], the MDGP can be solved iteratively using a method called *Branch-and-Prune* (BP) [28,29]. BP algorithm is based on a vertex order $v_1, ..., v_n \in V$ such that [7,15,35] (we denote x_i instead of x_{v_i} and $d_{i,j}$ instead of $d_{v_i v_j}$):

1. For v_1, v_2, v_3, there exist $x_1, x_2, x_3 \in \mathbb{R}^3$ satisfying equations (1.1);
2. For $i > 3$, there exist $v_{i-3}, v_{i-2}, v_{i-1}$ such that

$$\{\{v_{i-3}, v_i\}, \{v_{i-2}, v_i\}, \{v_{i-1}, v_i\}\} \subset E \tag{1.2}$$

and

$$d_{i-3,i-2} + d_{i-2,i-1} > d_{i-3,i-1}. \tag{1.3}$$

MDGP instances with this order are called the *Discretizable Molecular Distance Geometry Problem* (DMDGP) [20,21]. For other approaches that use different starting triplets of atoms, see [14].

From Property 1 above, the DMDGP solution set can avoid solutions obtained by rotations and translations and, from Property 2, the position for v_4 can be obtained solving the system

$$||x_4 - x_3||_2 = d_{3,4}, \tag{1.4}$$
$$||x_4 - x_2||_2 = d_{2,4},$$
$$||x_4 - x_1||_2 = d_{1,4}, \tag{1.5}$$

which can result in up to two possible solutions. For each position determined for v_4, we obtain other two for v_5, and so on, implying that the DMDGP search space is finite, having 2^{n-3} possible solutions [21,31].

For some $i > 4$, we may also have $\{v_j, v_i\} \in E$, $j < i-3$, adding another equation to the system related to v_i:

$$||x_i - x_{i-1}||_2 = d_{i-1,i},$$
$$||x_i - x_{i-2}||_2 = d_{i-2,i},$$
$$||x_i - x_{i-3}||_2 = d_{i-3,i},$$
$$||x_i - x_j||_2 = d_{j,i}.$$

If the points $x_{i-1}, x_{i-2}, x_{i-3}, x_j \in \mathbb{R}^3$ are not in the same plane, we obtain a unique solution x_i^* for v_i, supposing $||x_i^* - x_j||_2 = d_{j,i}$. However, it may

happen that both possible positions for v_i are infeasible with respect to additional distances $d_{j,i}$, $j < i-3$. In this case, it is necessary to consider the other possible position for v_{i-1} and repeat the procedure [21].

Geometrically, the requirements (1.2) and (1.3) of the DMDGP definition mean that, at each iteration of the BP algorithm, we have to intersect three spheres centered at the positions for vertices $v_{i-3}, v_{i-2}, v_{i-1}$ with radius $d_{i-3,i}, d_{i-2,i}, d_{i-1,i}$, respectively, resulting in two possible positions for v_i.

Distances $d_{i-1,i}$ and $d_{i-2,i}$ can be considered precise values, since they are related to bond lengths and bond angles of a protein, but if distances $d_{i-3,i}$ are provided by NMR experiments, they contain uncertainties [11].

In [22], imprecision in distances $d_{i-3,i}$ were represented as *interval distances* $[\underline{d}_{i-3,i}, \overline{d}_{i-3,i}]$, where $\underline{d}_{i-3,i} \leq d_{i-3,i} \leq \overline{d}_{i-3,i}$, and an extension of the BP algorithm, called iBP, were proposed, where the idea is to sample values from $[\underline{d}_{i-3,i}, \overline{d}_{i-3,i}]$ in order to solve a system like (1.4) related to v_i. The main drawback of this approach is that choosing many values, the search space increases exponentially, and for small samples, a solution may not be found [1,8,9,16,37,39]. When the data are not precise, it is difficult to do sphere intersection using linear algebra, because uncertainties in distance values imply uncertainties also in the centers of the spheres, not only in their radius.

Using Geometric Algebra (GA), it is possible to avoid sampling process and calculate intersection of spheres considering the uncertainties associated to their centers and radius. To the best of our knowledge, the first mathematical relationship between DG and GA was established by Dress and Havel, in 1993 [13]. However, a strong connection between GA and the DMDGP appeared only in 2015 [23]. In [2,3], for the first time, the Conformal GA (CGA) [17,26] was consistently applied to solve discrete versions of DG.

In [2,3], CGA was used to model uncertainties in the DMDGP, avoiding the sampling strategy and eliminating the heuristic characteristics of iBP. This was done for the *branching phase* of iBP, and this paper explains how the Oriented CGA (OCGA) [6] can be used in the *pruning phase* of iBP.

Next section explains how CGA replaces the classical approach to the DMDGP and Section 3 provides the original contribution of this paper, describing how OCGA can be integrated in iBP. New research directions are presented in Section 4.

2. Conformal Geometric Algebra for Branching

We will follow the arguments given in [2,3] to explain how CGA deals with uncertainties in the DMDGP.

If $d_{1,4}$ is an interval distance $[\underline{d}_{1,4}, \overline{d}_{1,4}]$, we have to intersect two spheres with one spherical shell resulting in two arcs, instead of two points in $\mathbb{R}^3$ (Fig. 1). Thus, we first obtain the points from the intersection of the spheres centered at the positions for v_1, v_2, v_3 with radius $\underline{d}_{1,4}, d_{2,4}, d_{3,4}$, resulting in $\underline{P}_4^0$ and $\underline{P}_4^1$, and with radius $\overline{d}_{1,4}, d_{2,4}, d_{3,4}$, resulting in $\overline{P}_4^0$ and $\overline{P}_4^1$ (Fig. 1). These points can be obtained from the *point pairs* generated by $\underline{S}_{1,4} \wedge S_{2,4} \wedge$

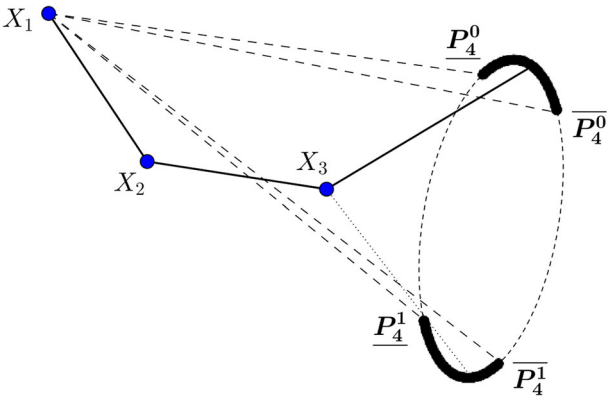

FIGURE 1. Interval distance $d_{1,4}$ generates two arcs for X_4

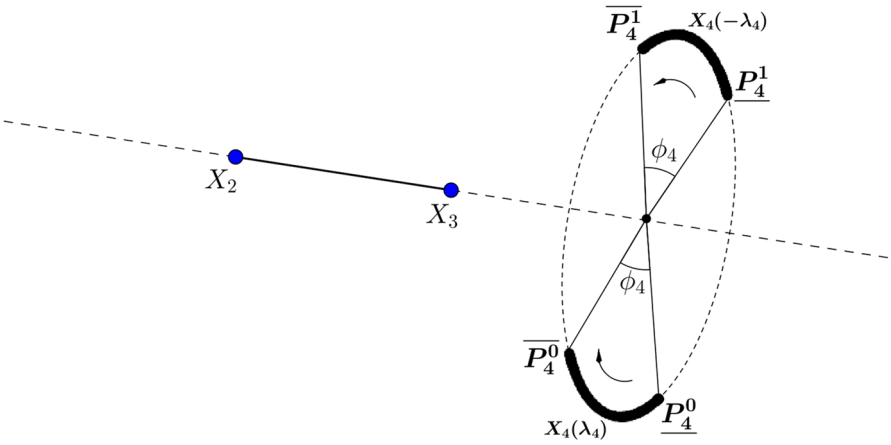

FIGURE 2. Rotation axis defined by X_2 and X_3

$S_{3,4}$ and $\overline{S}_{1,4} \wedge S_{2,4} \wedge S_{3,4}$ [12], where underline and overline indicate the use of $\underline{d}_{1,4}$ and $\overline{d}_{1,4}$, respectively, and $S_{i,j}$ is the sphere centered at the position of vertex v_i, denoted by X_i, with radius $d_{i,j}$:

$$S_{i,j} = X_i - 0.5 d_{i,j}^2 e_\infty.$$

With the starting and the ending point of an arc, we can define a *rotor* acting on that. For v_4, the rotation axis of its rotor is defined by X_2 and X_3, denoted by z_4, and the rotation angle ϕ_4 (in radians) is the angle corresponding to the arcs $\underline{P_4^0}\overline{P_4^0}$ and $\underline{P_4^1}\overline{P_4^1}$ (Fig. 2). The associated rotor R_4 is given by

$$R_4 = \cos\left(\tfrac{\lambda_4}{2}\right) - \sin\left(\tfrac{\lambda_4}{2}\right) z_4^*, \ 0 \leq \lambda_4 \leq \phi_4,$$

where $z_4 = X_2 \wedge X_3 \wedge e_\infty$ (z_4^* is the dual of z_4), and the two possible arcs are described by

$$X_4^0(\lambda_4) = R_4 \underline{P_4^0} R_4^{-1}$$

and

$$X_4^1(-\lambda_4) = R_4 \underline{P_4^1} R_4^{-1}.$$

We use negative values in $X_4^1(-\lambda_4)$ to invert the orientation in $\underline{P_4^1}\overline{P_4^1}$ (Fig. 2).

For $i > 4$, we can easily obtain (all the values ϕ_i, for $i > 3$, can be computed *a priori* based on the DMDGP definition)

$$X_i^0(\lambda_i) = R_i \underline{P_i^0} R_i^{-1},$$
$$X_i^1(-\lambda_i) = R_i \underline{P_i^1} R_i^{-1},$$

where

$$R_i = \cos\left(\frac{\lambda_i}{2}\right) - \sin\left(\frac{\lambda_i}{2}\right) z_i^*,\ 0 \leq \lambda_i \leq \phi_i, \tag{2.1}$$

and

$$z_i = X_{i-2} \wedge X_{i-1} \wedge e_\infty.$$

Note that expressions for $X_i^0(\lambda_i)$ and $X_i^1(-\lambda_i)$ consider that X_{i-2} and X_{i-1} are fixed points. However, the effect of changing the points in the arcs must be taken into account in order to avoid the sampling process (more details in the next section). This was done in [3], resulting in

$$X_i^b(\lambda_4, \ldots, \lambda_i) = (R_i \cdots R_4)\, \underline{P_i^b}\, \left(R_4^{-1} \cdots R_i^{-1}\right),$$

where

$$R_i = \cos\left(\frac{\lambda_i}{2}\right) - \sin\left(\frac{\lambda_i}{2}\right) z_i^*,\ 0 \leq \lambda_i \leq \phi_i,$$
$$z_i = (R_i \cdots R_4)\left(\underline{P_{i-2}^b} \wedge \underline{P_{i-1}^b} \wedge e_\infty\right)\left(R_4^{-1} \cdots R_i^{-1}\right),$$

and $\underline{P_i^b}$ is one of the points obtained from the intersection $\underline{S}_{i-3,i} \wedge S_{i-2,i} \wedge S_{i-1,i}$, for $i = 4, \ldots, n$, with $b \in \{0, 1\}$ (these values are defined when i BP chooses one of the branches in the search tree).

3. Oriented Conformal Geometric Algebra for Pruning

For each v_i, $i > 4$, iBP verifies if there are *pruning edges* $\{v_j, v_i\} \in E$, $j < i - 3$. If this is not the case, it can choose $\underline{P_i^0}$ or $\underline{P_i^1}$ and "ignore" the pruning phase, since both points can be used to continue the search. Since there is no pruning edge for v_4, let us consider v_5 with interval distances $[\underline{d}_{2,5}, \overline{d}_{2,5}]$ and $[\underline{d}_{1,5}, \overline{d}_{1,5}]$, for example (see Fig. 8).

Using the classical approach [22], we have to sample values from the interval $[\underline{d}_{2,5}, \overline{d}_{2,5}]$ to solve the system

$$||x_5 - x_4||_2 = d_{4,5},$$
$$||x_5 - x_3||_2 = d_{3,5},$$
$$\underline{d}_{2,5} \leq ||x_5 - x_2||_2 \leq \overline{d}_{2,5}.$$

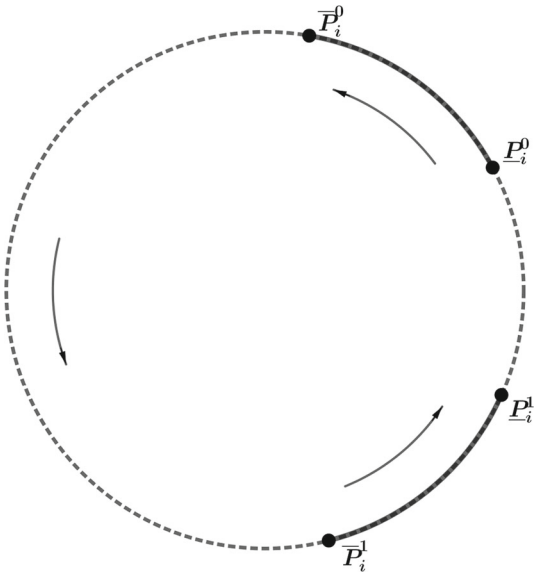

FIGURE 3. Oriented circle $C_i = \underline{P}_i^0 \wedge \overline{P}_i^0 \wedge \overline{P}_i^1$

That is, some values $d_{2,5}^1, d_{2,5}^2, ..., d_{2,5}^k \in [\underline{d}_{2,5}, \overline{d}_{2,5}]$ should be selected. Since $\{v_1, v_5\} \in E$, before considering vertex v_6, at least one of such values, *e.g.* $d_{2,5}^*$, must be associated with a solution x_5^* to the above system, *i.e.*,

$$\begin{aligned} ||x_5^* - x_4||_2 &= d_{4,5}, \\ ||x_5^* - x_3||_2 &= d_{3,5}, \\ ||x_5^* - x_2||_2 &= d_{2,5}^*, \end{aligned}$$

such that

$$\underline{d}_{1,5} \leq ||x_5^* - x_1||_2 \leq \overline{d}_{1,5}.$$

However, if the first pruning edge is $v_{1,6}$ ($\{v_1, v_5\} \notin E$), we do not know (during the calculations for v_5) how refined the sample from $[\underline{d}_{2,5}, \overline{d}_{2,5}]$ must be (*i.e.* how big is k?) in order to obtain a position for v_6 that satisfies

$$\underline{d}_{1,6} \leq ||x_6 - x_1|| \leq \overline{d}_{1,6}.$$

To avoid the sampling strategy during the pruning phase of *i*BP, it is necessary to take care of the *orientation* of the new arcs obtained when additional spherical shells related to the pruning edges must be considered. Oriented CGA (OCGA) can do that [6].

OCGA is an extension of the Oriented Projective Geometry, developed by Stolfi [38], for problems in computer graphics and computer vision.

First, let us define an orientation for the circle obtained from the intersection $S_{i-2,i} \wedge S_{i-1,i}$ (Fig. 3), given by

$$C_i = \underline{P_i^0} \wedge \overline{P_i^0} \wedge \overline{P_i^1}.$$

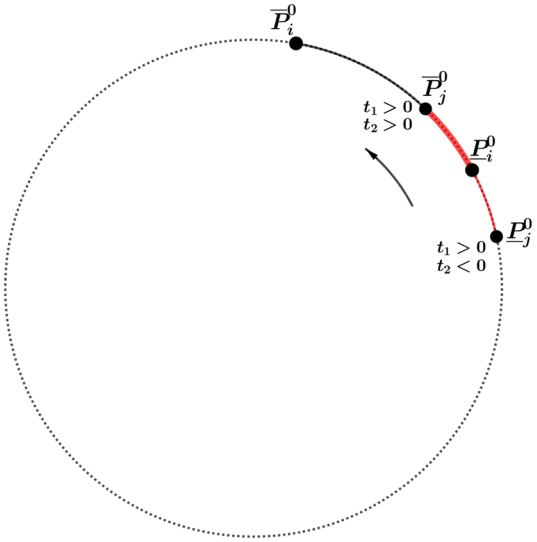

FIGURE 4. The new arc is $\underline{P_i^0}\overline{P_j^0}$

C_i is a trivector in the conformal space. So, its dual C_i^* is a bivector, orthogonal to the plane that contains the circle, which implies that the line

$$C_i^* \wedge e_\infty$$

is oriented according to C_i. Since C_i^* is the dual circle, given by the intersection of the dual spheres centered at X_{i-2} and X_{i-1}, the line $C_i^* \wedge e_\infty$ has the same direction of $\pm (X_{i-2} \wedge X_{i-1} \wedge e_\infty)$, but it carries the orientation of C_i. In practice, this means we are deciding correctly between $X_{i-2} \wedge X_{i-1} \wedge e_\infty$ and $X_{i-1} \wedge X_{i-2} \wedge e_\infty$ to be the rotation axis.

The related rotor R_i is then defined in a different way, compared to (2.1), using the normalized bivector dual to the rotation axis $C_i^* \wedge e_\infty$,

$$R_i = \cos\left(\frac{\lambda_i}{2}\right) - \left(\frac{\lambda_i}{2}\right) z_i, \ 0 \leq \lambda_i \leq \phi_i,$$

where

$$z_i = \frac{(C_i^* \wedge e_\infty)^*}{||C_i^* \wedge e_\infty||}.$$

Let us suppose now that for v_i, $i > 4$, there is a pruning edge $\{v_j, v_i\} \in E$, $j < i - 3$, with an interval distance $[\underline{d}_{j,i}, \overline{d}_{j,i}]$. Denoting by $\underline{P_j^0}\overline{P_j^0}$ and $\underline{P_j^1}\overline{P_j^1}$ the arcs obtained from the intersections $\underline{S}_{j,i} \wedge S_{i-2,i} \wedge S_{i-1,i}$ and $\overline{S}_{j,i} \wedge S_{i-2,i} \wedge S_{i-1,i}$, respectively, and using the tests defined in [6] to compare the orientation of two circles (where $t_i > 0$ indicates they have the same orientation), we can recognize all the possible cases, as illustrated in Figs. 4,5,6,7. This means that we have to calculate the following values (t_1

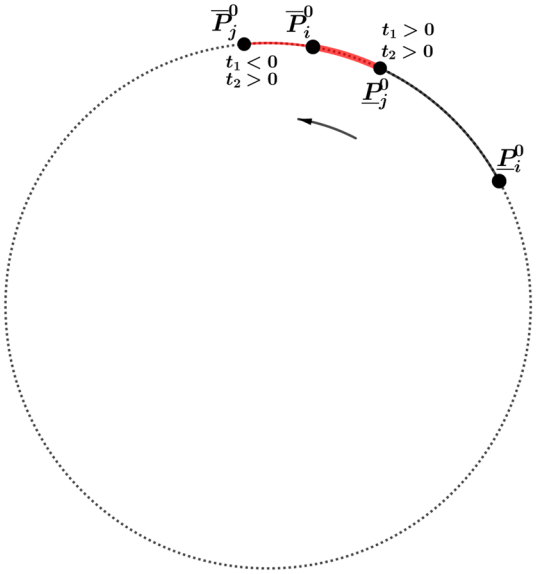

FIGURE 5. The new arc is $\underline{P_j^0}\overline{P_i^0}$

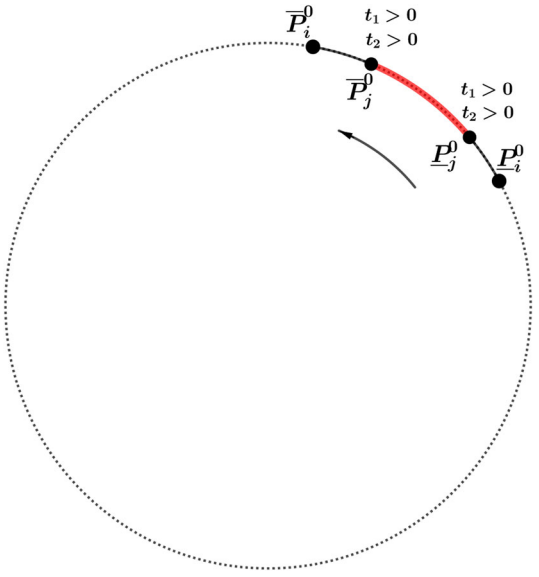

FIGURE 6. The new arc is $\underline{P_j^0}\overline{P_j^0}$

and t_2) for $\underline{P_j^0}$ and $\overline{P_j^0}$ (the same procedure for $\underline{P_j^1}$ and $\overline{P_j^1}$):

$$t_1 = \left(\underline{P_j^0} \wedge \overline{P_i^0} \wedge \overline{P_i^1}\right) C_i, \;\; t_2 = \left(\underline{P_i^0} \wedge \underline{P_j^0} \wedge \overline{P_i^1}\right) C_i$$

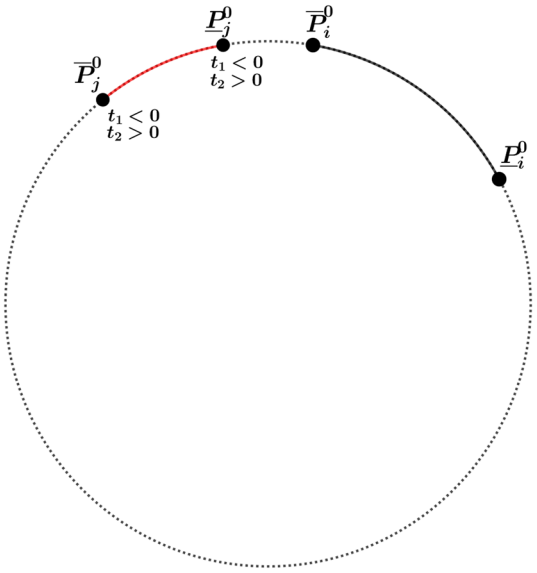

FIGURE 7. No intersection

and

$$t_1 = \left(\overline{P_j^0} \wedge \overline{P_i^0} \wedge \overline{P_i^1}\right) C_i, \;\; t_2 = \left(\underline{P_i^0} \wedge \overline{P_j^0} \wedge \overline{P_i^1}\right) C_i,$$

where

$$C_i = \underline{P_i^0} \wedge \overline{P_i^0} \wedge \overline{P_i^1}.$$

The situation in Fig. 6 is different from the others, since both points ($\underline{P_j^0}$ and $\overline{P_j^0}$) are inside the arc $\underline{P_i^0}\overline{P_i^0}$. Consider the oriented circle defined by

$$\underline{P_j^0} \wedge \overline{P_j^0} \wedge P,$$

where P is any point on the circle, but outside the arc $\underline{P_i^0}\overline{P_i^0}$. Without loss of generality, let us consider P as a point obtained when a rotor is applied on $\overline{P_j^0}$, in such a way to be outside from the arc $\underline{P_i^0}\overline{P_i^0}$. Since

$$\underline{P_j^0} \wedge \overline{P_j^0} \wedge P = P \wedge \underline{P_j^0} \wedge \overline{P_j^0},$$

there is no change in the orientation of the trivector $\underline{P_j^0} \wedge \overline{P_j^0} \wedge P$, which implies that

$$\left(\underline{P_j^0} \wedge \overline{P_j^0} \wedge P\right)\left(\underline{P_j^0} \wedge \overline{P_j^0} \wedge P'\right) > 0,$$

for different rotation angles generating P and P'.

3.1. Example

Let us consider a DMDGP instance with the vertex order $v_1, v_2, v_3, v_4, v_5, v_6$ and the following associated distances:

$$d_{i-1,i} = 1,\ i = 2, ..., 6,$$
$$d_{i-2,i} = \sqrt{3},\ i = 3, ..., 6,$$

$$d_{1,4} = 2.15,\ d_{2,5} \in [2.20, 2.60],\ d_{3,6} \in [2.40, 2.60],$$
$$d_{1,5} \in [2.45, 2.55].$$

This is the same example presented in [3], where the pruning phase was done "visually". Now, we can solve it formally, using OCGA. All the calculations were done using GAALOP: http://www.gaalop.de/.

After fixing the positions for v_1, v_2, v_3, and also for v_4 (since $d_{1,4}$ is a precise value),

$$x_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},\ x_2 = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix},\ x_3 = \begin{bmatrix} -1.5 \\ \frac{\sqrt{3}}{2} \\ 0 \end{bmatrix},\ x_4 = \begin{bmatrix} -1.311 \\ 1.552 \\ 0.702 \end{bmatrix},$$

we obtain the arcs $\underline{P_5^0}\overline{P_5^0}$ and $\underline{P_5^1}\overline{P_5^1}$, from the intersection of spheres $\underline{S}_{2,5} \wedge S_{3,5} \wedge S_{4,5}$ and $\overline{S}_{2,5} \wedge S_{3,5} \wedge S_{4,5}$, respectively:

$$\underline{P_5^0} = e_0 - 0.409e_1 + 1.981e_2 + 0.753e_3 + 2.329e_\infty,$$
$$\underline{P_5^1} = e_0 - 1.502e_1 + 1.350e_2 + 1.663e_3 + 3.422e_\infty,$$

and

$$\overline{P_5^0} = e_0 - 1.386e_1 + 2.525e_2 + 0.484e_3 + 4.266e_\infty,$$
$$\overline{P_5^1} = e_0 - 2.046e_1 + 2.144e_2 + 1.033e_3 + 4.966e_\infty.$$

The pruning edge $\{v_1, v_5\}$ implies we have to calculte $\underline{S}_{1,5} \wedge S_{3,5} \wedge S_{4,5}$ and $\overline{S}_{1,5} \wedge S_{3,5} \wedge S_{4,5}$, giving the points

$$\underline{A_5^0} = e_0 - 0.6735e_1 + 2.299e_2 + 0.5132e_3 + 3.0012e_\infty,$$
$$\underline{A_5^1} = e_0 - 1.2602e_1 + 1.2827e_2 + 1.664e_3 + 3.0012e_\infty,$$

and

$$\overline{A_5^0} = e_0 - 0.7952e_1 + 2.3768e_2 + 0.47e_3 + 3.2512e_\infty,$$
$$\overline{A_5^1} = e_0 - 1.4069e_1 + 1.3173e_2 + 1.6696e_3 + 3.2512e_\infty.$$

Now, we define the oriented circle C_5,

$$C_5 = \underline{P_5^0} \wedge \overline{P_5^0} \wedge \overline{P_5^1},$$

and perform the tests described above for $\underline{A_5^0}, \underline{A_5^1}, \overline{A_5^0}, \overline{A_5^1}$:

$$\begin{aligned}
\underline{A_5^0}: \quad & t_1 = (\underline{A_5^0} \wedge \overline{P}_5^0 \wedge \overline{P}_5^1)C_5 = \ 0.468 > 0 \\
& t_2 = (\underline{P_5^0} \wedge \underline{A_5^0} \wedge \overline{P}_5^1)C_5 = \ 0.531 > 0 \\
\underline{A_5^1}: \quad & t_1 = (\underline{A_5^1} \wedge \overline{P}_5^0 \wedge \overline{P}_5^1)C_5 = 0.965 > 0 \\
& t_2 = (\underline{P_5^0} \wedge \underline{A_5^1} \wedge \overline{P}_5^1)C_5 = -1.425 < 0 \\
\overline{A_5^0}: \quad & t_1 = (\overline{A_5^0} \wedge \overline{P}_5^0 \wedge \overline{P}_5^1)C_5 = \ 0.359 > 0 \\
& t_2 = (\underline{P_5^0} \wedge \overline{A_5^0} \wedge \overline{P}_5^1)C_5 = \ 0.650 > 0 \\
\overline{A_5^1}: \quad & t_1 = (\overline{A_5^1} \wedge \overline{P}_5^0 \wedge \overline{P}_5^1)C_5 = 0.878 > 0
\end{aligned}$$

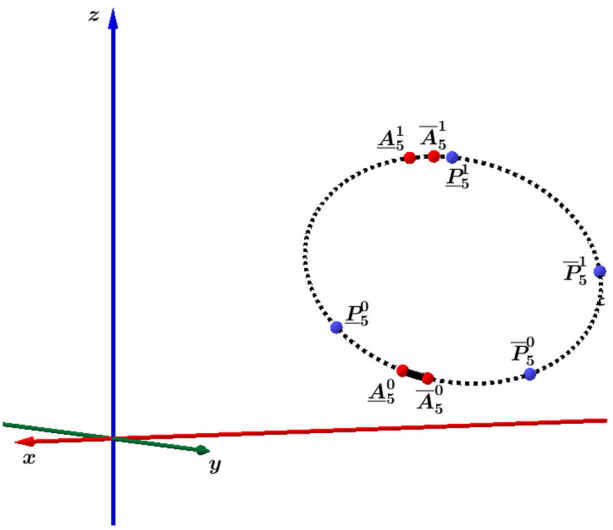

FIGURE 8. One arc is reduced and the other is pruned

$$t_2 = (\underline{P_5^0} \wedge \overline{A_5^1} \wedge \overline{P}_5^1)C_5 = -1.389 \ < 0$$

From the results of the tests (see Fig. 8), arc $\underline{P_5^1}\ \overline{P_5^1}$ is pruned and arc $\underline{P_5^0}\overline{P_5^0}$ is reduced to a new one, given by $\underline{A_5^0}\overline{A_5^0}$. The rotation axis z_5 is defined by

$$z_5 = \left(\left(\underline{A_5^0} \wedge \overline{A_5^0} \wedge \overline{P_5^0}\right)^* \wedge e_\infty\right)^*,$$

the new angle ϕ_5 (associated to $\underline{A_5^0}\overline{A_5^0}$) is recomputed, the rotor R_5 is given by

$$R_5 = \cos\left(\tfrac{\lambda_5}{2}\right) - \sin\left(\tfrac{\lambda_5}{2}\right) z_5, \ 0 \leq \lambda_5 \leq 0.174,$$

and, finally, after pruning at vertex v_5, without sampling values from $[2.20, 2.60]$, we obtain

$$X_5^0(\lambda_5) = R_5 \underline{A_5^0} R_5^{-1}.$$

Doing the calculations, we get

$$x_5 = \begin{bmatrix} -1.7602\sin^2(\frac{\lambda_5}{2}) - 1.3085\cos(\frac{\lambda_5}{2})\sin(\frac{\lambda_5}{2}) - 0.6735\cos^2(\frac{\lambda_5}{2}) \\ 1.4920\sin^2(\frac{\lambda_5}{2}) + 0.9673\cos(\frac{\lambda_5}{2})\sin(\frac{\lambda_5}{2}) + 2.2990\cos^2(\frac{\lambda_5}{2}) \\ 1.5939\sin^2(\frac{\lambda_5}{2}) - 0.5935\cos(\frac{\lambda_5}{2})\sin(\frac{\lambda_5}{2}) + 0.5133\cos^2(\frac{\lambda_5}{2}) \end{bmatrix}.$$

For vertex v_6, we intersect the spheres centered at $X_3, X_4, X_5^0(0)$, with radius $\underline{d}_{3,6}, d_{4,6}, d_{5,6}$, resulting in

$$\underline{P_6^0} = e_0 - 1.149e_1 + 3.224e_2 + 0.2784e_3 + 5.896e_\infty,$$
$$\underline{P_6^1} = e_0 + 0.2279e_1 + 2.344e_2 + 0.768e_3 + 3.068e_\infty,$$

and with radius $\overline{d}_{3,6}, d_{4,6}, d_{5,6}$, resulting in

$$\overline{P_6^0} = e_0 - 0.907e_1 + 3.221e_2 + 0.9275e_3 + 6.03e_\infty,$$
$$\overline{P_6^1} = e_0 - 0.1484e_1 + 2.737e_2 + 1.197e_3 + 4.473e_\infty.$$

The oriented circle C_6 is defined by

$$C_6 = \underline{P_6^0} \wedge \overline{P_6^0} \wedge \overline{P_6^1}$$

and the angle ϕ_6 (associated to $\underline{P_6^0}\overline{P_6^0}$ and $\underline{P_6^1}\overline{P_6^1}$) is given by $\phi_6 = 0.8231$.

Now, we have to consider the effect of rotor R_5, *i.e.*

$$\begin{aligned} C_6(\lambda_5) &= \left(R_5 \underline{P_6^0} R_5^{-1}\right) \wedge \left(R_5 \overline{P_6^0} R_5^{-1}\right) \wedge \left(R_5 \overline{P_6^1} R_5^{-1}\right) \\ &= R_5 \left(\underline{P_6^0} \wedge \overline{P_6^0} \wedge \overline{P_6^1}\right) R_5^{-1} \\ &= R_5 C_6 R_5^{-1}, \end{aligned}$$

which implies that

$$R_6 = \cos\left(\tfrac{\lambda_6}{2}\right) - \sin\left(\tfrac{\lambda_6}{2}\right) z_6, \ 0 \leq \lambda_6 \leq \phi_6,$$

and

$$z_6 = \frac{(C_6^*(\lambda_5) \wedge e_\infty)^*}{||C_6^*(\lambda_5) \wedge e_\infty||}.$$

Thus, the positions for v_6 are given by

$$X_6^0(\lambda_5, \lambda_6) = R_6 R_5 \underline{P_6^0} R_5^{-1} R_6^{-1}$$

and

$$X_6^1(\lambda_5, -\lambda_6) = R_6 R_5 \underline{P_6^1} R_5^{-1} R_6^{-1}.$$

4. Conclusions

The branching phase of iBP algorithm can be described using the language of CGA to eliminate the heuristic characteristic of the classical strategy. This paper explains how Oriented CGA can be used to model the other phase of iBP, the pruning phase. This is an important step in order to solve one of the open problems in Distance Geometry, proposed in [27]. Next challenges are related to the design of efficient computational codes to solve instances with real protein NMR data.

Acknowledgements

We would like to thank the Brazilian research agencies CNPq, CAPES, and FAPESP, for their financial support, and also Leo Dorst, for his comments and suggestions.

References

[1] Agra, A., Figueiredo, R., Lavor, C., Maculan, N., Pereira, A., Requejo, C.: Feasibility check for the distance geometry problem: an application to molecular conformations. Int. Trans. Oper. Res. **24**, 1023–1040 (2017)

[2] Alves, R., Lavor, C.: Geometric algebra to model uncertainties in the discretizable molecular distance geometry problem. Adv. Appl. Clifford Algebra **27**, 439–452 (2017)

[3] Alves, R., Lavor, C., Souza, C., Souza, M.: Clifford algebra and discretizable distance geometry. Math. Methods Appl. Sci. **41**, 3999–4346 (2018)

[4] Billinge, S., Duxbury, P., Gonçalves, D., Lavor, C., Mucherino, A.: Assigned and unassigned distance geometry: applications to biological molecules and nanostructures. 4OR **14**, 337–376 (2016)

[5] Billinge, S., Duxbury, P., Gonçalves, D., Lavor, C., Mucherino, A.: Recent results on assigned and unassigned distance geometry with applications to protein molecules and nanostructures. Ann. Oper. Res. **271**, 161–203 (2018)

[6] Cameron, J., Lasenby, J.: Oriented conformal geometric algebra. Adv. Appl. Clifford Algebra **18**, 523–538 (2008)

[7] Cassioli, A., Gunluk, O., Lavor, C., Liberti, L.: Discretization vertex orders in distance geometry. Discrete Appl. Math. **197**, 27–41 (2015)

[8] Cassioli, A., Bordeaux, B., Bouvier, G., Mucherino, A., Alves, R., Liberti, L., Nilges, M., Lavor, C., Malliavin, T.: An algorithm to enumerate all possible protein conformations verifying a set of distance constraints. BMC Bioinform. **16**, 16–23 (2015)

[9] Costa, T., Bouwmeester, H., Lodwick, W., Lavor, C.: Calculating the possible conformations arising from uncertainty in the molecular distance geometry problem using constraint interval analysis. Inform. Sci. **415–416**, 41–52 (2017)

[10] Crippen, G., Havel, T.: Distance Geometry and Molecular Conformation. Wiley, New York (1988)

[11] Donald, B.: Algorithms in Structural Molecular Biology. MIT Press, Cambridge (2011)

[12] Dorst, L., Fontijne, D., Mann, S.: Geometric Algebra for Computer Science: An Object-Oriented Approach to Geometry. Morgan Kaufman, San Mateo (2007)

[13] Dress, A., Havel, T.: Distance geometry and geometric algebra. Found. Phys. **23**, 1357–1374 (1993)

[14] Fidalgo, F., Gonalves, D., Lavor, C., Liberti, L., Mucherino, A.: A symmetry-based splitting strategy for discretizable distance geometry problems. J. Glob. Optim. **71**, 717–733 (2018)

[15] Gonçalves, D., Mucherino, A.: Discretization orders and efficient computation of Cartesian coordinates for distance geometry. Optim. Lett. **8**, 2111–2125 (2014)

[16] Gonçalves, D., Mucherino, A., Lavor, C., Liberti, L.: Recent advances on the interval distance geometry problem. J. Glob. Optim. **69**, 525–545 (2017)

[17] Hestenes, D.: Old wine in new bottles: a new algebraic framework for computational geometry. In: Corrochano E. B., Sobczyk G. (eds.) Geometric Algebra with Applications in Science and Engineering. Birkhäuser, Boston (2001)

[18] Hildenbrand, D.: Foundations of Geometric Algebra Computing. Springer, Berlin Heidelberg (2012)

[19] Lavor, C., Xambó-Descamps, S., Zaplana, I.: A Geometric Algebra Invitation to Space-Time Physics Robotics and Molecular Geometry. SpringerBriefs in Mathematics. Springer, Berlin Heidelberg (2018)

[20] Lavor, C., Liberti, L., Maculan, N., Mucherino, A.: Recent advances on the discretizable molecular distance geometry problem. Eur. J. Oper. Res. **219**, 698–706 (2012)

[21] Lavor, C., Liberti, L., Maculan, N., Mucherino, A.: The discretizable molecular distance geometry problem. Comput. Optim. Appl. **52**, 115–146 (2012)

[22] Lavor, C., Liberti, L., Mucherino, A.: The interval BP algorithm for the discretizable molecular distance geometry problem with interval data. J. Glob. Optim. **56**, 855–871 (2013)

[23] Lavor, C., Alves, R., Figueiredo, W., Petraglia, A., Maculan, N.: Clifford algebra and the discretizable molecular distance geometry problem. Adv. Appl. Clifford Algebra **25**, 925–942 (2015)

[24] Lavor, C., Liberti, L., Lodwick, W., Mendonça da Costa, T.: An Introduction to Distance Geometry applied to Molecular Geometry. SpringerBriefs in Computer Science. Springer, Berlin Heidelberg (2017)

[25] Lavor, C., Liberti, L., Donald, B., Worley, B., Bardiaux, B., Malliavin, T., Nilges, M.: Minimal NMR distance information for rigidity of protein graphs. Discrete Applied Mathematics (2018) **(to appear)**

[26] Li, H., Hestenes, D., Rockwood, A.: Generalized Homogeneous Coordinates for Computational Geometry. In: Sommer, G. (ed.) Geometric Computing with Clifford Algebra, pp. 25–58. Springer, Berlin Heidelberg (2001)

[27] Liberti, L., Lavor, C.: Open Research Areas in Distance Geometry. In: Pardalos, P., Migdalas, A. (eds.) Open Problems in Optimization and Data Analysis. Springer, Berlin Heidelberg (2018). (to appear)

[28] Liberti, L., Lavor, C., Maculan, N.: A branch-and-prune algorithm for the molecular distance geometry problem. Int. Trans. Oper. Res. **15**, 1–17 (2008)

[29] Liberti, L., Lavor, C., Mucherino, A., Maculan, N.: Molecular distance geometry methods: from continuous to discrete. Int. Trans. Oper. Res. **18**, 33–51 (2010)

[30] Liberti, L., Lavor, C., Maculan, N., Mucherino, A.: Euclidean distance geometry and applications. SIAM Rev. **56**, 3–69 (2014)

[31] Liberti, L., Masson, B., Lee, J., Lavor, C., Mucherino, A.: On the number of realizations of certain Henneberg graphs arising in protein conformation. Discrete Appl. Math. **165**, 213–232 (2014)

[32] Liberti, L., Lavor, C.: Six mathematical gems from the history of distance geometry. Int. Trans. Oper. Res. **23**, 897–920 (2016)

[33] Liberti, L., Lavor, C.: Euclidean Distance Geometry. An Introduction. Springer, Berlin (2017)

[34] Menger, K.: Untersuchungen uber allgemeine Metrik. Math. Ann. **100**, 75–163 (1928)

[35] Mucherino, A., Lavor, C., Liberti, L.: The discretizable distance geometry problem. Optim. Lett. **6**, 1671–1686 (2012)

[36] Mucherino, A., Lavor, C., Liberti, L., Maculan, N. (eds.): Distance Geometry: Theory, Methods, and Applications. Springer, Berlin (2013)

[37] Souza, M., Lavor, C., Muritiba, A., Maculan, N.: Solving the molecular distance geometry problem with inaccurate distance data. BMC Bioinform. **14**, S71–S76 (2013)

[38] Stolfi, J.: Oriented Projective Geometry—A Framework for Geometric Computations. Academic Press, Cambridge (1991)

[39] Worley, B., Delhommel, F., Cordier, F., Malliavin, T., Bardiaux, B., Wolff, N., Nilges, M., Lavor, C., Liberti, L.: Tuning interval branch-and-prune for protein structure determination. J. Glob. Optim. **72**, 109–127 (2018)

[40] Wütrich, K.: Protein structure determination in solution by nuclear magnetic resonance spectroscopy. Science **243**, 45–50 (1989)

Carlile Lavor
University of Campinas (IMECC-UNICAMP)
Campinas SP 13081-970
Brazil
e-mail: clavor@ime.unicamp.br

Rafael Alves
Federal University of ABC (CMCC-UFABC)
Santo André SP 09210-580
Brazil
e-mail: rafasoalves@gmail.com

Received: September 14, 2018.
Accepted: November 21, 2018.