



A simple and optimal algorithm for strict circular seriation

Mikhael Carmona, Victor Chepoi, Guyslain Naves, Pascal Pr  a

► To cite this version:

Mikhael Carmona, Victor Chepoi, Guyslain Naves, Pascal Pr  a. A simple and optimal algorithm for strict circular seriation. 2024. hal-03714735v2

HAL Id: hal-03714735

<https://hal.science/hal-03714735v2>

Preprint submitted on 2 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destin  e au d  p  t et    la diffusion de documents scientifiques de niveau recherche, publi  s ou non,   manant des   tablissements d'enseignement et de recherche fran  ais ou   trangers, des laboratoires publics ou priv  s.

A simple (and optimal?) algorithm for circular seriation

M. CARMONA^{a,b}, V. CHEPOI^a, G. NAVES^a, AND P. PRÉA^{a,b}

^aLIS, Aix-Marseille Université, CNRS, and Université de Toulon
Marseille, France

^bÉcole Centrale Marseille, Marseille, France

{mikhael.carmona,victor.chepoi,guyslain.naves,pascal.prea}@lis-lab.fr

December 4, 2023

Abstract. Recently, Armstrong, Guzmán, and Sing Long (2021), presented an optimal $O(n^2)$ time algorithm for strict circular seriation (called also the recognition of strict quasi-circular Robinson spaces). In this paper, we give a very simple $O(n \log n)$ time algorithm for computing a compatible circular order for strict circular seriation. When the input space is not known to be strict quasi-circular Robinson, our algorithm is complemented by an $O(n^2)$ time verification of compatibility of the returned order. This algorithm also works for recognition of other types of strict circular Robinson spaces known in the literature. We also prove that the circular Robinson dissimilarities (which are defined by the existence of compatible orders on one of the two arcs between each pair of points) are exactly the pre-circular Robinson dissimilarities (which are defined by a four-point condition).

1. INTRODUCTION

A major issue in classification and data analysis is to visualize simple geometrical and relational structures between objects based on their pairwise distances. The classical *seriation problem* (called also the matrix reordering problem), introduced by Robinson [22], asks to find a simultaneous ordering (or permutation) of the rows and the columns of the dissimilarity matrix so that its values increase monotonically in the rows and the columns when moving away from the main diagonal in both directions. The permutation which leads to a matrix with such a property is called a *compatible order* and dissimilarity matrices admitting a compatible order are called *Robinson matrices*. The Robinson matrices can be thus characterized by the existence of a compatible order $<$ and the 3-point condition $d(x, z) \geq \max\{d(x, y), d(y, z)\}$ for any three points x, y, z such that $x < y < z$. If this inequality is strict, then such a matrix is called *strict Robinson*.

Due to the importance in seriation and classification, the algorithmic problem of recognizing Robinson dissimilarities/matrices attracted the interest of many authors. The existing recognition algorithms can be classified into *combinatorial* and *spectral*. All combinatorial algorithms use the correspondence between Robinson dissimilarities and interval hypergraphs. The main difficulty arising in recognition algorithms is the existence of several compatible orders. The first recognition algorithm by Mirkin and Rodin [18] consists in testing if the hypergraph of balls is an interval hypergraph and runs in $O(n^4)$ time. Chepoi and Fichet [6] gave a simple divide-and-conquer algorithm running in $O(n^3)$ time. Seston [23] presented another $O(n^3)$ time algorithm, by using threshold graphs. In [24], he improved the complexity of his algorithm to $O(n^2 \log n)$ by using a sorting of the data and PQ-trees. Finally, in 2014, Préa and Fortin [20] presented an algorithm running in optimal $O(n^2)$ time. The efficiency of the algorithm of [20] is due to the use of the PQ-trees of Booth and Lueker [3] as a data structure for encoding all compatible orders. Even if optimal, the algorithm of [20] is far from being simple. Subsequently, two new recognition algorithms were proposed by Laurent and Seminaroti: in [14] they presented an algorithm of complexity $O(\alpha \cdot n)$ based on classical LexBFS traversal and divide-and-conquer (where α is the depth of the recursion tree, which is at most the number of distinct elements of the input matrix), and in [15] they presented an $O(n^2 \log n)$ algorithm, which extends LexBFS to weighted matrices and is used as a multisweep traversal. More recently, in [5] we gave a simple and practical $O(n^2)$ divide-and-conquer

algorithm based on decompositions of dissimilarity spaces into modules. The spectral approach was introduced by Atkins et al. [2] and was subsequently used in numerous papers. The method is based on the computation of the second smallest eigenvalue and its eigenvector of the Laplacian of a similarity matrix and also uses PQ-trees to represent the compatible orders.

A natural generalization of Robinson dissimilarities and compatible orders is to consider a circular order instead of a linear one. This is often referred to as the *circular seriation problem*. For origins of circular seriation see the papers [9, 10, 11] and for recent applications of circular seriation in planar tomographic reconstruction, gene expression, DNA replication and 3D conformation, see the papers [7, 17, 16]. For a spectral approach to circular seriation, see the papers [8, 13, 21]. At the difference of the classical seriation, where the notion of Robinson dissimilarity is a well-established standard, in circular seriation there are several non-equivalent notions of circular Robinson dissimilarities. Hubert, Arabie, and Meulman [10] were the first to define circular Robinson dissimilarities via a 4-point condition. Brucker and Osswald [4] undertaken a systematic study of various definitions of circular Robinson dissimilarities from the point of view of classification and combinatorics. In Robinson dissimilarity spaces, the sets of balls, of 2-balls (intersections of two balls), and of clusters (maximal cliques in the threshold graphs) are all interval hypergraphs. Hypercycles, introduced and investigated by Quillot [19], are the circular analogs of interval hypergraphs and are the hypergraphs whose hyperedges can be represented as circular intervals (arcs). In case of circular Robinson dissimilarities, requiring that the ball, the 2-ball, or the cluster hypergraphs are hypercycles lead to three different classes of dissimilarities. Their structural properties have been thoroughly studied by Brucker and Osswald [4]. The dissimilarities whose ball hypergraph is a hypercycle is the most general one and was characterized via a simple 4-point condition: there exists a circular order \prec such that if $x \prec y \prec z \prec t$, then $d(x, z) \geq \min\{d(x, y), d(x, t)\}$. We call such dissimilarities quasi-circular Robinson. Recently, Armstrong, Guzmán, and Sing Long [1] presented an optimal $O(n^2)$ time algorithm for the recognition of strict quasi-circular Robinson dissimilarities, i.e., for which the ball hypergraph is a hypercycle and $x \prec y \prec z \prec t$ implies that $d(x, z) > \min\{d(x, y), d(x, t)\}$. Among other tools, their algorithm uses PQ-trees. Finally, to characterize the dissimilarities for which the 2-ball hypergraphs are hypercycles, Brucker and Osswald [4] introduced the notion of pre-circular Robinson dissimilarities: these are the dissimilarities which admit a circular order \prec such that $x \prec y \prec z \prec t$ implies that $d(x, z) \geq \min\{\max\{d(x, y), d(y, z)\}, \max\{d(x, t), d(t, z)\}\}$.

In this paper, we give a very simple $O(n \log n)$ time algorithm which builds a compatible circular order for all versions of strict circular Robinson dissimilarities, introduced and investigated in the papers [1, 4, 10]. Then the adjunction of a verification step gives an optimal $O(n^2)$ time algorithm to recognize these dissimilarities. Our second main result is proving that the pre-circular Robinson dissimilarities are exactly the dissimilarities for which there exists a circular order \prec such that for each pair (x, y) , the restriction of d to one of the two arcs between x and y is a Robinson dissimilarity (in the usual sense) and \prec is its compatible order. To our knowledge, prior to our work no results of this kind for circular seriation were known. Our result shows that in fact pre-circular Robinson spaces should be called circular Robinson spaces. Finally, the simplicity of our algorithm led us to other structural properties of strict circular Robinson spaces, in particular we show that they admit only one or two compatible circular orders and their opposites.

2. PRELIMINARIES

2.1. Dissimilarity spaces. Let $X = \{x_1, \dots, x_n\}$ be a set of n elements, called *points*. A *dissimilarity* on X is a symmetric function d from X^2 to the nonnegative real numbers such that $d(x, y) = d(y, x) \geq 0$ and $d(x, y) = 0$ if and only if $x = y$. Then $d(x, y)$ is called the *distance* between x, y and (X, d) is called a *dissimilarity space*. The *ball* (respectively, the *sphere*) of radius $r \geq 0$ centered at $x \in X$ is the set $B_r(x) := \{y \in X : d(x, y) \leq r\}$ (respectively, $S_r(x) := \{y \in X : d(x, y) = r\}$). The *eccentricity* of a point x is $r_x := \max\{d(x, y) : y \in X\}$. Given a point $x \in X$, a point $y \in X$ is called a *furthest neighbor* of x if $d(x, y) = r_x$. Denote by F_x the set

of all furthest neighbors of x and note that $F_x = S_{r_x}(x)$. The *distance* between two subsets A, B of X is $d(A, B) = \min\{d(a, b) : a \in A, b \in B\}$.

2.2. Compatible orders and Robinson spaces. A partial order on X is called *linear* if any two elements of X are comparable. A dissimilarity d and a linear order $<$ on X are called *compatible* if $x < y < z$ implies $d(x, z) \geq \max\{d(x, y), d(y, z)\}$. If d and $<$ are compatible, then d is also compatible with the linear order $<^{\text{op}}$ opposite to $<$. If a dissimilarity space (X, d) admits a compatible order, then d is said to be *Robinson* and (X, d) is called a *Robinson space*. Equivalently, (X, d) is Robinson if its distance matrix $D = (d(p_i, p_j))$ can be symmetrically permuted so that its elements do not decrease when moving away from the main diagonal along any row or column. Such a dissimilarity matrix D is said to have the *Robinson property*. From the definition of a Robinson dissimilarity follows that d is Robinson if and only if there exists an order $<$ on X such that all balls $B_r(x)$ of (X, d) are intervals of $<$. Moreover, this property holds for all compatible orders. Basic examples of Robinson dissimilarities are the ultrametrics, thoroughly used in phylogeny. Recall, that d is an *ultrametric* if $d(x, y) \leq \max\{d(x, z), d(y, z)\}$ for all $x, y, z \in X$. Another example of a Robinson space is provided by the standard *line-distance* between n points $p_1 < \dots < p_n$ of \mathbb{R} . Any line-distance has exactly two compatible orders: the order $p_1 < \dots < p_n$ defined by the coordinates of the points and its opposite.

2.3. Compatible circular orders and circular Robinson spaces. Informally speaking, a circular order on a finite set X is obtained by arranging the points of X on a circle C . Formally, a *circular order* is a ternary relation β on X where $\beta(u, v, w)$ expresses that the directed path from u to w passes through v . This relation is total, asymmetric, and transitive, which can be formulated in terms of Huntington's axioms [12]: for any four points u, v, w, x of X ,

- (CO1) if u, v, w are distinct, then $\beta(u, v, w)$ or $\beta(w, v, u)$,
- (CO2) $\beta(u, v, w)$ and $\beta(w, v, u)$ is impossible,
- (CO3) $\beta(u, v, w)$ implies $\beta(v, w, u)$,
- (CO4) $\beta(u, v, w)$ and $\beta(u, w, x)$ imply $\beta(u, v, x)$.

It follows from this definition that only triplets of distinct points can be in the relation β and that the reverse relation β^{op} defined by $\beta^{\text{op}}(u, v, w)$ exactly when $\beta(w, v, u)$ is also a circular order. Since X is finite, the circular orders on X are just the orientations of the circle C with points of X located on C . We will suppose that β corresponds to the counterclockwise order of C and β^{op} to the clockwise order of C . Given a circular order β and three distinct points u, v, w , we will write $u \triangleleft v \triangleleft w$ if $\beta(u, v, w)$ holds. For a sequence of points x_1, x_2, \dots, x_k containing at least three distinct points, we will write $x_1 \triangleleft x_2 \triangleleft \dots \triangleleft x_k$ if for any $i < j < k$ with x_i, x_j, x_k distinct, $\beta(x_i, x_j, x_k)$ holds. Notice that in that case, if $x_i = x_j$, then $x_k = x_i$ for each $k \in \{i, \dots, j\}$, and that $x_2 \triangleleft x_3 \triangleleft \dots \triangleleft x_k \triangleleft x_1$.

We say that a nonempty proper subset A of X is an *arc* of a circular order β on X if there are no four distinct points $u, v \in A$ and $x, y \in X \setminus A$ such that $u \triangleleft x \triangleleft v \triangleleft y$. If A is an arc, then so is $X \setminus A$. For two points $x, y \in X$, consider the arcs $X_{xy}^\beta = \{t \in X : \beta(x, t, y)\} \cup \{x, y\}$ and $X_{yx}^\beta = \{t \in X : \beta(y, t, x)\} \cup \{x, y\}$. Notice that $X_{xy}^\beta \cup X_{yx}^\beta = X$ and $X_{xy}^\beta \cap X_{yx}^\beta = \{x, y\}$. Moreover, if $x \triangleleft y \triangleleft z$, then $X_{xy}^\beta \subset X_{xz}^\beta$ and $X_{yz}^\beta \subset X_{xz}^\beta$. For $x, y \in X$ and $Z \subset X$, we write $x \triangleleft y \triangleleft Z$ if for all $z \in Z$ we have $x \triangleleft y \triangleleft z$. Arcs are for circular orders what intervals are for linear orders. The intersection of two arcs is not necessarily an arc but we have the following elementary observation:

Lemma 1. *Let A and B be two arcs of a circular ordered set (X, β) . If there exists $x \in X \setminus (A \cup B)$, then $A \cap B$ is an arc. If there exist $x \in B \setminus A$, then $A \setminus B$ is an arc or is empty.*

Let (X, d) be dissimilarity space, β be a circular order on X and $x, y, z, t \in X$ such that $x \triangleleft y \triangleleft z \triangleleft t$.

- The points x, y, z, t are *one-side Robinson*, and we denote it by $\text{cR}(x, y, z, t)$, if $d(x, z) \geq \min\{\max\{d(x, y), d(y, z)\}, \max\{d(x, t), d(t, z)\}\}$.

- The points x, y, z, t are *strictly one-side Robinson*, and we denote it by $\text{scR}(x, y, z, t)$, if $d(x, z) > \min\{\max\{d(x, y), d(y, z)\}, \max\{d(x, t), d(t, z)\}\}$.
- The points x, y, z, t are *quasi one-side Robinson*, and we denote it by $\text{qcR}(x, y, z, t)$, if $d(x, z) \geq \min\{d(y, z), d(t, z)\}$.
- The points x, y, z, t are *strictly quasi one-side Robinson*, and we denote it by $\text{sqcR}(x, y, z, t)$, if $d(x, z) > \min\{d(y, z), d(t, z)\}$.

Notice that the conditions $\text{cR}(x, y, z, t)$ and $\text{qcR}(x, y, z, t)$ trivially hold if $x = y < z < t$ or $x < y < z = t$. For $x, y, z, t \in X$ such that $x < y < z < t$, the following implications hold:

$$\begin{array}{ccc} \text{scR}(x, y, z, t) & \implies & \text{cR}(x, y, z, t) \\ \Downarrow & & \Downarrow \\ \text{sqcR}(x, y, z, t) & \implies & \text{qcR}(x, y, z, t) \end{array}$$

A dissimilarity space (X, d) is called *pre-circular Robinson* if there exists a circular order β , which is said to be a *compatible order*, such that for all $x, y, z, t \in X$, if $x < y < z < t$ then $\text{cR}(x, y, z, t)$ holds. The *strictly pre-circular Robinson*, *quasi-circular Robinson*, and *strictly quasi-circular Robinson* spaces are defined in a similar way by using conditions $\text{scR}(x, y, z, t)$, $\text{qcR}(x, y, z, t)$, and $\text{sqcR}(x, y, z, t)$, respectively. Finally, a dissimilarity space (X, d) is called *circular Robinson* if there exists a circular order β , called a *compatible order* such that for all $x, y \in X$, either (X_{xy}^β, d) or (X_{yx}^β, d) is a Robinson space and the restriction of $<$ to the arc X_{xy}^β or respectively X_{yx}^β is a (linear) compatible order. If (X, d) is a circular (respectively, pre-circular) Robinson space with a compatible circular order β , we say that (X, d) is a circular (respectively, pre-circular) Robinson space. Notice also that for all definitions, if a circular order β is compatible, then β^{op} is also compatible. A set X of n points on a circle C in \mathbb{R}^2 endowed with the arc distance or with the chord (i.e., Euclidean) distance is a basic example of a strict circular Robinson space.

Circular Robinson spaces introduced by Hubert et al. [10] generalize classical Robinson spaces and are particular pre-circular Robinson spaces (for their definition, see [4, 10]). That Robinson spaces are circular Robinson can be seen by arranging the points of X on a circle C according to a compatible order of (X, d) . Then for all $x, y \in X$, if $x < y$ in the compatible order, then d is Robinson on the arc X_{xy}^β . (Strictly) circular Robinson spaces are (strictly) quasi-circular Robinson spaces. However not any circular order β satisfying $\text{sqcR}(x, y, z, t)$ for all quadruplets $x < y < z < t$ also satisfies $\text{scR}(x, y, z, t)$. Such an example is provided in Figure 1.



FIGURE 1. A strict quasi-circular Robinson space $(X = \{x, y, z, t\}, d)$ with a compatible circular order β . The quadruplet $x < y < z < t$ satisfies $\text{sqcR}(x, y, z, t)$ but not $\text{scR}(x, y, z, t)$. Notice that the circular order obtained by reversing z and t , i.e., such that $x < y < t < z$, satisfies the condition scR for all quadruplets.

3. PRE-CIRCULAR ROBINSON SPACES ARE CIRCULAR ROBINSON

In this section, we characterize pre-circular Robinson spaces. Instead of relying directly on the condition $\text{cR}(x, y, z, w)$, some proofs will use the following consequence of the definition of pre-circular Robinson spaces, stating intuitively that for any pair u, w , one of the arcs $X_{u,w}^\beta, X_{w,u}^\beta$ has only chords shorter than $d(u, w)$.

Lemma 2. *Let (X, d) be a pre-circular Robinson space with a compatible circular order β and points $u \triangleleft y \triangleleft y' \triangleleft w \triangleleft z \triangleleft z'$, where u and w are distinct. Then $d(u, w) \geq \min\{d(y, y'), d(z, z')\}$. Moreover, if (X, d) is strictly pre-circular Robinson, then this inequality is strict.*

Proof. We present the proof for the non-strict case, the strict case being slightly simpler. For sake of contradiction, assume $u \triangleleft y \triangleleft y' \triangleleft w \triangleleft z \triangleleft z'$ is a counterexample with a minimum number of distinct points, implying that $d(u, w) < d(y, y')$ and $d(u, w) < d(z, z')$. By $\text{cR}(u, y, y', w)$,

$$\begin{aligned} d(u, y') &\geq \min\{\max\{d(u, y), d(y, y')\}, \max\{d(y', w), d(w, u)\}\} \\ &\geq \min\{d(y, y'), d(w, u)\} \\ &= d(w, u). \end{aligned}$$

If $d(u, y') = d(w, u)$, then $u \triangleleft y \triangleleft w \triangleleft w \triangleleft z \triangleleft z'$ is a counterexample. If $d(u, y') > d(w, u)$, then $u \triangleleft u \triangleleft y' \triangleleft w \triangleleft z \triangleleft z'$ is a counterexample. By the minimality of our counterexample, we conclude that either $u = y$ or $w = y'$. Symmetrically, either $z = w$ or $z' = x$. We also have that $\{y, y'\} \neq \{u, w\}$. Hence, let $y'' \in \{y, y'\} \setminus \{u, w\}$ and $z'' \in \{z, z'\} \setminus \{u, w\}$. By $\text{cR}(u, y'', w, z'')$,

$$\begin{aligned} d(u, w) &\geq \min\{\max\{d(u, y''), d(y'', w)\}, \max\{d(w, z''), d(z'', u)\}\} \\ &\geq \min\{d(y, y'), d(z, z')\}, \end{aligned}$$

since $\{y, y'\}$ is either $\{u, y''\}$ or $\{y'', w\}$, and $\{z, z'\}$ is either $\{w, z''\}$ or $\{z'', u\}$. This is in contradiction with the minimality of the counterexample. \square

As a consequence we have:

Lemma 3. *Let (X, d) be a pre-circular Robinson space with a compatible circular order β and x, y, z be three arbitrary points of X such that $x \triangleleft y \triangleleft z$. If $d(x, z) < \max\{d(x, y), d(y, z)\}$, then (X_{zx}^β, d) is a Robinson space.*

Proof. Let $y', y'' \in \{x, y, z\}$ be such that $d(x, z) < d(y', y'')$ and $x \triangleleft y' \triangleleft y'' \triangleleft z$. Pick any points $u, v, w \in X_{zx}^\beta$ such that $z \triangleleft u \triangleleft v \triangleleft w \triangleleft x$ (we may have $u = z$ or $w = x$) and suppose by way of contradiction that $d(u, w) < \max\{d(u, v), d(v, w)\}$, let $v', v'' \in \{u, v, w\}$ be such that $u \triangleleft v' \triangleleft v'' \triangleleft w$ and $d(u, w) < d(v', v'')$. If $d(u, w) \leq d(x, z)$, then $d(u, w) < d(y', y'')$, and by Lemma 2 on $w \triangleleft y' \triangleleft y'' \triangleleft u \triangleleft v' \triangleleft v''$, this is a contradiction. If $d(u, w) > d(x, z)$, then $d(x, z) < d(v', v'')$, and by Lemma 2 on $x \triangleleft y' \triangleleft y'' \triangleleft z \triangleleft v' \triangleleft v''$ we obtain again a contradiction. \square

Now, we can prove our first main result:

Theorem 1. *A dissimilarity space (X, d) is pre-circular Robinson if and only if (X, d) is circular Robinson.*

Proof. To prove the theorem, first suppose that (X, d) is a circular Robinson space and β is a compatible circular order on X . Pick any $x, y, z, t \in X$ such that $x \triangleleft y \triangleleft z \triangleleft t$. By definition of β , \triangleleft is a compatible linear order on the arc X_{xz}^β or X_{zx}^β . In the first case, since $y \in X_{xz}^\beta$, we have $d(x, z) \geq \max\{d(x, y), d(y, z)\}$. In the second case, since $t \in X_{zx}^\beta$, we have $d(x, z) \geq \max\{d(x, t), d(t, z)\}$. Consequently, $d(x, z) \geq \min\{\max\{d(x, y), d(y, z)\}, \max\{d(x, t), d(t, z)\}\}$, establishing that (X, d) is a pre-circular Robinson space.

Conversely, let (X, d) be a pre-circular Robinson space and β be a compatible circular order. Pick any pair of points a, b of X . If (X_{ab}^β, d) is not a Robinson space, then there exists three points

$x, y, z \in X_{ab}^\beta$ such that $x \triangleleft y \triangleleft z$ and $d(x, z) < \max\{d(x, y), d(y, z)\}$. By Lemma 3, (X_{zx}^β, d) is a Robinson space. Since $X_{ba}^\beta \subset X_{zx}^\beta$, this proves that (X_{ba}^β, d) is Robinson, establishing that (X, d) is a circular Robinson space and β is a compatible circular order. \square

As a consequence, we say that (X, d) is a *strictly circular Robinson space* if it is a strictly pre-circular Robinson space.

4. PROPERTIES OF QUASI-CIRCULAR AND STRICT CIRCULAR ROBINSON SPACES

In this section, we present several properties of (strict) quasi-circular and circular Robinson spaces. We also show how to verify if a dissimilarity space is (strictly) quasi-circular Robinson or (strictly) circular Robinson with respect to a given circular order.

4.1. Properties of (strictly) quasi-circular Robinson spaces. In this subsection, we recall or present some properties of (strictly) quasi-circular Robinson spaces. We start with the following characterization of quasi-circular Robinson spaces of [4]:

Proposition 1. [4] *A dissimilarity space (X, d) is quasi-circular Robinson if and only if there exists a circular order β such that any $x \in X$ and $r \in \mathbb{R}^+$, the ball $B_r(x)$ and its complement $X \setminus B_r(x) = \{t \in X : d(x, t) > r\}$ are arcs of β .*

Proof. First suppose that (X, d) is a quasi-circular Robinson space and β is a compatible circular order. Let $y \triangleleft u \triangleleft z \triangleleft v$ be four distinct points with $y, z \in B_r(x)$. We may assume $v \triangleleft x \triangleleft y$ by symmetry. By qcR(x, u, z, v), $r \geq d(x, z) \geq \min\{d(x, u), d(z, v)\}$, that is either $u \in B_r(x)$ or $v \in B_r(x)$, proving that $B_r(x)$ is an arc. Conversely, suppose that there exists a circular order β such that each ball $B_r(x)$ is an arc of β . Pick arbitrary points $x, y, z, t \in X$ such that $x \triangleleft y \triangleleft z \triangleleft t$. Let $r = d(x, z)$. Since $B_r(z)$ is an arc of β and $\beta(x, y, z), \beta(z, t, x)$ hold, either y or t must belong to the ball $B_r(z)$. Consequently, $d(x, z) \geq \min\{d(y, z), d(t, z)\}$, establishing qcR(x, y, z, t). \square

Let (X, d) be a quasi-circular Robinson space and β be a compatible circular order. For any point $x \in X$, recall that F_x consists of all furthest neighbors of x and r_x is the eccentricity of x . Let $M_x := X \setminus (F_x \cup \{x\})$. Since $M_x \cup \{x\}$ is the ball $B_{r_x}(x)$, where r is strictly smaller but sufficiently close to r_x , by Proposition 1, $M_x \cup \{x\}$ and F_x are complementary arcs of β . Consequently, the set M_x is partitioned into two arcs $L_x := \{t \in M_x : x \triangleleft t \triangleleft F_x\}$ and $R_x := \{t \in M_x : F_x \triangleleft t \triangleleft x\}$ (left and right arcs), where one of those arcs may be empty. Two points $y, y' \in M_x$ are called *x-separated* if they belong to distinct arcs L_x and R_x .

Lemma 4. *Let (X, d) be a quasi-circular Robinson space, β be a compatible circular order, and x any point of X . If $y, z \in L_x$ and $x \triangleleft y \triangleleft z$ or $y, z \in R_x$ and $z \triangleleft y \triangleleft x$, then $d(x, y) \leq d(x, z)$. Moreover, if (X, d) is strict quasi-circular, then $d(x, y) < d(x, z)$.*

Proof. Let $y, z \in L_x$ with $x \triangleleft y \triangleleft z$. Let $t \in F_x$. By qcR(x, y, z, t) and since $d(x, z) < d(x, t)$, we obtain that $d(x, y) \leq d(x, z)$. The proof for $y, z \in R_x$ is similar. \square

Lemma 5. *Let (X, d) be a strict quasi-circular Robinson space, β be a compatible circular order, and x any point of X . Then any sphere $S_r(x)$ contains at most two points. Furthermore, if $r < r_x$ and $S_r(x)$ consists of two points y, y' , then y and y' are x -separated.*

Proof. Suppose by way of contradiction that there exist three points $y, y', y'' \in S_r(x)$. We can suppose, with no loss of generality, that $x \triangleleft y \triangleleft y' \triangleleft y''$. But this contradicts qcR(y', y'', x, y). Since (X, d) is strict quasi-circular, by Lemma 4, $|S_r(x) \cap L_x| \leq 1$ and $|S_r(x) \cap R_x| \leq 1$, proving the second assertion. \square

4.2. Properties of (strictly) circular Robinson spaces. In this subsection, we present some properties of (strictly) circular Robinson spaces.

Proposition 2. *Let (X, d) be a circular Robinson space with a compatible order β . Then for all $x, y \in X$, $x' \in F_x$, $y' \in F_y$ with $|\{x, x', y, y'\}| \geq 3$, one of the following assertions holds:*

- (a) $x \leq y \leq x' \leq y'$,
- (b) $x \leq y' \leq x' \leq y$,
- (c) $\{y, y'\} \cap F_x \neq \emptyset$ or $\{x, x'\} \cap F_y \neq \emptyset$.

Moreover if (X, d) is strictly circular Robinson, then either (a) or (b) holds.

Proof. Suppose first that (X, d) is strictly circular Robinson. For sake of contradiction, assume none of these assertions holds. There are two cases depending on the order (up to reversal) of x, y, x', y' .

If $x \leq y \leq y' \leq x'$, by $\text{scR}(x, y, y', x')$ and $\text{scR}(y, y', x', x)$, and since $x' \in F_x$, $y' \in F_y$, we get:

$$\begin{aligned}
 d(x, x') &\geq d(x, y') \\
 &> \min\{\max\{d(x, y), d(y, y')\}, \max\{d(x, x'), d(x', y')\}\} \\
 &\geq \min\{d(y, y'), \max\{d(x, x'), d(x', y')\}\} \\
 &\geq \min\{d(y, y'), d(x, x')\}, \\
 d(y, y') &\geq d(y, x') \\
 &> \min\{\max\{d(y, y'), d(y', x')\}, \max\{d(y, x), d(x, x')\}\} \\
 &\geq \min\{\max\{d(y, y'), d(y', x')\}, d(x, x')\} \\
 &\geq \min\{d(y, y'), d(x, x')\}.
 \end{aligned}$$

From this, we get that $d(x, x') > d(y, y')$ and $d(y, y') > d(x, x')$, a contradiction.

If $x \leq y' \leq y \leq x'$, then by $\text{scR}(x, y', y, x')$ and using that $x' \in F_x$, $y' \in F_y$, we get:

$$\begin{aligned}
 \min\{d(x, x'), d(y, y')\} &\geq d(x, y) \\
 &> \min\{\max\{d(x, y'), d(y', y)\}, \max\{d(x, x'), d(x', y)\}\} \\
 &\geq \min\{d(y', y), d(x, x')\},
 \end{aligned}$$

a contradiction.

Suppose now that (X, d) is non-strictly circular Robinson. We follow the same argument. In the first case, instead of a contradiction, we get that $d(x, x') = d(y, y')$. Applying $\text{cR}(x, y, y', x')$ and $\text{cR}(y, y', x', x)$, we conclude that $y' \in F_x$ and $x' \in F_y$, implying (c). In the second case, we get that $d(x, y) = \min\{d(x, x'), d(y, y')\}$, implying either $y \in F_x$ or $x \in F_y$, that is (c). \square

Now, we determine the conditions that make a circular order β , compatible with a quasi-circular Robinson space (X, d) , not compatible with respect to circular Robinson property of (X, d) .

Proposition 3. *Let (X, d) be a (strict) quasi-circular Robinson space and β a compatible order, such that (X, d) is not (strict) circular Robinson with respect to β . Then there exist $x, y \in X$, $x' \in F_x$, $y' \in F_y$ such that $x \leq x' \leq y \leq y'$ or $x \leq y' \leq y \leq x'$. Moreover, in the non-strict case, we may also assume that $x, x' \notin F_y$ and $y, y' \notin F_x$.*

Proof. We first prove the strict case. Let $x \leq u \leq y \leq v$ be such that $\text{scR}(x, u, y, v)$ does not hold:

$$(1) \quad d(x, y) \leq \min\{\max\{d(x, u), d(u, y)\}, \max\{d(x, v), d(v, y)\}\}.$$

By $\text{sqcR}(x, u, y, v)$ and $\text{sqcR}(y, v, x, u)$, we get:

$$(2) \quad d(x, y) > \min\{d(x, u), d(x, v)\},$$

$$(3) \quad d(x, y) > \min\{d(y, u), d(y, v)\}.$$

Combining these inequations, we get:

$$\begin{aligned} \min\{d(x, u), d(x, v)\} &< \max\{d(x, u), d(u, y)\}, & \min\{d(x, u), d(x, v)\} &< \max\{d(x, v), d(v, y)\}, \\ \min\{d(y, u), d(y, v)\} &< \max\{d(x, u), d(u, y)\}, & \min\{d(y, u), d(y, v)\} &< \max\{d(x, v), d(v, y)\}, \end{aligned}$$

and then:

$$\begin{aligned} d(x, v) &< d(x, u) \vee d(x, u) < d(u, y), & d(x, u) &< d(x, v) \vee d(x, v) < d(v, y), \\ d(y, v) &< d(y, u) \vee d(y, u) < d(x, u), & d(y, u) &< d(y, v) \vee d(y, v) < d(x, v), \end{aligned}$$

which is equivalent to the disjunction of these two symmetric assertions:

- (i) $d(x, v) < d(x, u)$, $d(x, v) < d(v, y)$, $d(y, u) < d(y, v)$, and $d(y, u) < d(x, u)$,
- (ii) $d(x, v) > d(x, u)$, $d(x, v) > d(v, y)$, $d(y, u) > d(y, v)$, and $d(y, u) > d(x, u)$.

We may assume the first. Then $d(x, v) = \min\{d(x, v), d(x, u)\} < d(x, y) \leq \min\{d(x, u), d(v, y)\} \leq d(x, u)$ (by Inequalities (1) and (2)), which implies by the strict unimodality of distances from x that $F_x \subseteq X_{uy}^\beta$. Similarly, $d(y, u) < d(x, y) \leq d(y, v)$ which implies that $F_y \subseteq X_{vx}^\beta$. Consequently, if $x' \in F_x$ and $y' \in F_y$, then we get $x \prec x' \prec y \prec y'$, as expected.

In the non-strict case, Inequality (1) becomes a strict inequality, while Inequalities (2) and (3) become non-strict inequalities. Combining these inequations, we get the same conclusion as in the strict case. For example, in the first case we get that $d(x, v) = \min\{d(x, v), d(x, u)\} \leq d(x, y) < \min\{d(x, u), d(v, y)\} \leq d(x, u)$, yielding $d(x, v) \leq d(x, y) < d(x, u)$ and $d(y, u) \leq d(x, y) < d(y, v)$. By unimodality of distances, we conclude that $F_x \subseteq X_{uy}^\beta$ and $F_y \subseteq X_{vx}^\beta$. Consequently, if $x' \in F_x$ and $y' \in F_y$, then we get $x \prec x' \prec y \prec y'$. Moreover, $y \notin F_x$ and $x \notin F_y$. Since $x \prec x' \prec v \prec y'$ and $d(x, v) < d(x, x')$, by $\text{qcR}(x, x', v, y')$ we conclude that $d(x, v) \geq d(x, y')$, yielding $y' \notin F_x$. Analogously, one can show that $x' \notin F_y$. \square

4.3. Verification of compatibility. In this subsection, given circular order β , we describe how to verify in $O(n^2)$ if a dissimilarity space (X, d) on n points is (strictly) quasi-circular Robinson or (strictly) circular Robinson with respect to β . This verification task can also be done in $O(n^2)$ for strict circular Robinson spaces, as defined in [10].

To test whether (X, d) is (strictly) quasi-circular Robinson with respect to β , by Proposition 1 we have to test whether all balls of (X, d) are arcs of β . This can be done in the following way. Let D be the distance matrix of (X, d) ordered according to the circular order β . The matrix D is called *unimodal* if for each row i , when moving circularly from the element d_{ii} on the main diagonal of D to the right until the last element d_{in} and then from the first element d_{i1} until d_{ii} , the elements first increase monotonically, stay at the maximal values, and then decrease monotonically. Since D is symmetric, the same monotonicity property holds also for each column i : moving down from d_{ii} until d_{ni} and then from d_{1i} until d_{ii} , the elements first increase monotonically, stay at the maximal value, and then decrease monotonically. We say that D is *strictly unimodal* if the values strictly increase, have one or two maximal elements, and then strictly decrease. It was shown in [1, Proposition 3.7] that β is a compatible circular order for a quasi-circular Robinson space (respectively, strictly quasi-circular Robinson space) if and only if D is unimodal (respectively, strictly unimodal). From the definition, testing if D is (strictly) unimodal can be easily done in $O(n^2)$ time. In case of strict unimodality we also have to check that each row has at most two maximal elements (this correspond to computing for each $x \in X$ the set F_x and checking if $|F_x| \leq 2$). Notice that for strictly circular Robinson spaces defined in [10], this testing task can be also done in $O(n^2)$ time.

Next, we consider the task of testing whether (X, d) is (strictly) circular Robinson with respect to a circular order β . As (strictly) circular Robinson spaces are particular cases of (strictly) quasi-Robinson spaces, the (strict) unimodality of the distance matrix D is a necessary condition for compatibility. Under this condition, we can use Proposition 3. Namely, we compute the arc F_x for each $x \in X$, and store the indices of its extremities. This can be done by dichotomy (using β) in

$O(n \log n)$ total time. Then for each pair $x, y \in X$, we can check in constant time whether there are $x' \in F_x, y' \in F_y$ as given in Proposition 3. If such elements exist, then by Proposition 2, (X, d) is not (strictly) circular Robinson. Otherwise, (X, d) is (strictly) circular Robinson with respect to β . This testing task can be done in $O(n^2)$ time. As a consequence, we have the following result:

Proposition 4. *For a dissimilarity space (X, d) on n points and a circular order β on X , one can check in $O(n^2)$ time whether, with respect to β , (X, d) is (1) (strictly) quasi-circular Robinson, (2) (strictly) circular Robinson.*

5. THE RECOGNITION ALGORITHM

In this section, we describe a simple but optimal algorithm to recognize strictly circular Robinson spaces and strictly quasi-circular Robinson spaces. Our algorithm consists in partitioning X into four sets with respect to any point $x \in X$ and any $x' \in F_x$. We prove that those four sets are arcs in any compatible circular order β and that the restriction of β to each of these four sets is obtained by sorting its points by distances to x . Concatenating these four arcs, we obtain two circular orders. Finally, it suffices to verify that one of these circular orders is compatible. This also shows that any strict circular Robinson space (in each of the three versions) has one or two compatible circular orders and their opposites.

5.1. How to define arcs X_{xy}^β metrically. Given a dissimilarity space (X, d) and two distinct points $x, y \in X$, we set $J^\circ(x, y) = \{u \in X : d(x, y) > \max\{d(x, u), d(u, y)\}\}$ and $J(x, y) = J^\circ(x, y) \cup \{x, y\}$. In all results of this subsection, we assume that (X, d) is a strict quasi-circular Robinson space and β is an arbitrary compatible circular order on X .

Lemma 6. *Let $x < y < z$ be three points of X such that $d(x, y) \leq \min\{d(x, z), d(y, z)\}$. Then $X_{xy}^\beta = J(x, y)$.*

Proof. First, let $v \in X_{xy}^\beta \setminus \{x, y\}$, i.e., $x < v < y$. By $\text{sqcR}(x, v, y, z)$ and since $d(x, y) \leq d(x, z)$, we have $d(x, v) < d(x, y)$. By $\text{sqcR}(y, z, v, x)$ and since $d(y, x) \leq d(y, z)$, we have $d(y, v) < d(y, x)$. Hence $X_{xy}^\beta \subseteq J(x, y)$. To prove the converse inclusion, let $u \in X_{yz}^\beta \setminus \{y, z\}$, that is $y < u < z$. By $\text{sqcR}(x, y, u, z)$, $d(x, u) > \min\{d(x, y), d(x, z)\} \geq d(x, y)$, hence $u \notin J^\circ(x, y)$. Similarly if $u \in X_{zx}^\beta \setminus \{z, x\}$, applying $\text{sqcR}(y, z, u, x)$ we also get $u \notin J^\circ(x, y)$. Consequently, $(X_{yz}^\beta \cup X_{zx}^\beta) \cap J^\circ(x, y) = \{x, y\}$, establishing the inclusion $J(x, y) \subseteq X_{xy}^\beta$. Thus $X_{xy}^\beta = J(x, y)$. \square

Now, let x be an arbitrary point of X and $x' \in F_x$. Let $N = \{u \in X : d(u, x) \leq d(u, x')\}$ and $F = \{u \in X : d(u, x) \geq d(u, x')\}$. Note that $N \cup F = X$ and $x \in N \setminus F, x' \in F \setminus N$.

Lemma 7. *N and F are arcs of β .*

Proof. It suffices to prove that N is an arc, as $F = X \setminus N$ and $x' \in F \neq \emptyset$. Let $y, z \in X \setminus \{x, x'\}$ be distinct points with $x < y < z < x'$ and z in N . We assert that $y \in N$. By $\text{sqcR}(x, y, z, x')$, we have $d(x, y) < d(x, z)$. By $\text{sqcR}(x', x', y, z)$,

$$d(x', y) > \min\{d(x', z), d(x', x)\} \geq \min\{d(x, z), d(x, x)\} = d(x, z) > d(x, y),$$

where the second inequality follows from the facts that $z \in N$ and $x' \in F_x$. This implies that $y \in N$. Symmetrically, if $z < y < x$ with $z \in N$, then $y \in N$. Hence, N is an arc of β . \square

Lemma 8. *If there exists $y \in X$ with $d(x, y) = d(y, x')$, then $J(x, y) \cup J(y, x')$ either coincides with $X_{xx'}^\beta$ when $x < y < x'$ or with $X_{x'x}^\beta$ when $x' < y < x$.*

Proof. Suppose without loss of generality that $x < y < x'$ (see Figure 2 (a)). Since $d(x, y) = d(y, x')$ and $d(x, y) \leq d(x, x')$, by Lemma 6 we conclude that $X_{xy}^\beta = J(x, y)$. By (CO3), we have $y < x' < x$. From the choice of the points $x' \in F_x$ and y we have $d(y, x') \leq \min\{d(x', x), d(y, x)\}$. By Lemma 6

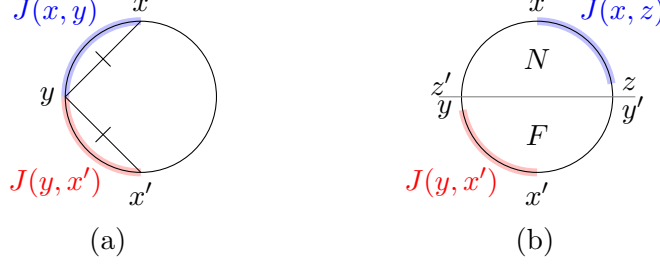


FIGURE 2. Configurations occurring in (a) Lemma 8 and (b) Lemma 9. In (b), the positions of z and z' may be exchanged, as well as those of y and y' .

we conclude that $X_{yx'}^\beta = J(y, x')$. Finally, since $x \triangleleft y \triangleleft x'$, we have $X_{xx'}^\beta = X_{xy}^\beta \cup X_{yx'}^\beta$, yielding $X_{xx'}^\beta = J(x, y) \cup J(y, x')$. \square

Consequently, if a point y with $d(x, y) = d(y, x')$ exists, then according to Lemma 8, the circular order β such that $x \triangleleft y \triangleleft x'$ can be computed in $O(n \log n)$ time. This is done by computing $X_{xx'}^\beta = J(x, y) \cup J(y, x')$, then ordering the points of $X_{xx'}^\beta$ and of its complement $X \setminus X_{xx'}^\beta$ by distances to x , by Lemma 4. Notice that in this case the compatible circular order β is unique up to reversal. Thus, we may next assume that $N \cap F = \emptyset$.

Lemma 9. *If $N \cap F = \emptyset$, then there exist $z, z' \in N$ and $y, y' \in F$ such that $N = X_{zx}^\beta \cup X_{xz'}^\beta$ and $F = X_{yx'}^\beta \cup X_{x'y'}^\beta$. The pairs $\{z, z'\}$, $\{y, y'\}$ and the partition $X_{zx}^\beta \cup X_{xz'}^\beta \cup X_{yx'}^\beta \cup X_{x'y'}^\beta$ can be computed in $O(n)$ time.*

Proof. Assume that $N \neq \{x\}$, and let $z \in N$ with $d(x, z)$ maximal (see Figure 2 (b)). Then applying Lemma 6 to x, z, x' , we have that $J(x, z)$ is either X_{xz}^β or X_{zx}^β (depending of whether $x \triangleleft z \triangleleft x'$ or $x' \triangleleft z \triangleleft x$). We denote $N' = J(x, z)$. If $N'' = N \setminus N' \neq \emptyset$, let $z' \in N''$ with $d(x, z')$ maximal. By Lemma 6, $J(x, z')$ is either $X_{xz'}^\beta$ or $X_{z'x}^\beta$. By Lemma 4, z and z' are x -separated, that is:

- either $J(x, z') = X_{z'x}^\beta$ and $J(x, z) = X_{xz}^\beta$,
- or $J(x, z) = X_{zx}^\beta$ and $J(x, z') = X_{xz'}^\beta$.

By the choice of z and z' , we conclude that $N = J(x, z) \cup J(x, z')$. If z or z' are not defined (because $N = \{x\}$ or $N' = \emptyset$), we may suppose them equal to x .

Pick any $y \in F$. Then $d(x', y) < d(y, x) \leq d(x', x)$. Thus we can also use Lemma 6 and get similarly that there exist points $y, y' \in F$ (where $y \in F$ with $d(x', y)$ maximal, $F' = J(x', y)$, and $y' \in F'' = F \setminus F'$ with $d(x', y')$ maximal) such that

- either $J(x', y') = X_{y'x'}^\beta$ and $J(x', y) = X_{x'y}^\beta$,
- or $J(x', y) = X_{yx'}^\beta$ and $J(x', y') = X_{x'y'}^\beta$,

and $F = J(x', y) \cup J(x', y')$. From their definition, it immediately follows that the pairs $\{z, z'\}$, $\{y, y'\}$ and the partition $X_{zx}^\beta \cup X_{xz'}^\beta \cup X_{yx'}^\beta \cup X_{x'y'}^\beta$ can be computed in $O(n)$ time. \square

Sorting by distances to x and to x' , we can find in $O(n \log n)$ time the linear orders of N and F such that each of these orders is either as in β or in β^{op} . It remains to determine whether the two chosen linear orders are correctly merged in a circular order. This is done by the following result:

Lemma 10. *Let $N = \{x_1, x_2, \dots, x_k\}$ and $F = \{y_1, y_2, \dots, y_\ell\}$ be such that*

- (a) *either $x_1 \triangleleft_1 x_2 \triangleleft_1 \dots \triangleleft_1 x_k \triangleleft_1 y_1 \triangleleft_1 y_2 \triangleleft_1 \dots \triangleleft_1 y_\ell$,*
- (b) *or $x_k \triangleleft_2 x_{k-1} \triangleleft_2 \dots \triangleleft_2 x_1 \triangleleft_2 y_1 \triangleleft_2 y_2 \triangleleft_2 \dots \triangleleft_2 y_\ell$.*

One can decide which of these two circular orders (possibly both) is compatible in $O(n)$ time.

Proof. Denote the circular order from (a) by β_1 and the circular order from (b) by β_2 . Suppose that β_1 is not compatible, that is β_2 is compatible. Then there is a quadruplet $u \prec_1 v \prec_1 w \prec_1 z$ with $d(u, w) \leq \min\{d(u, v), d(u, z)\}$ or $d(v, z) \leq \min\{d(v, w), d(v, u)\}$ or $d(w, u) \leq \min\{d(w, z), d(w, v)\}$ or $d(z, v) \leq \min\{d(z, u), d(z, w)\}$, say the first. Since β_2 is a compatible circular order, we must have $u = x_i$, $v = x_{i'}$, $w = y_j$, and $z = y_{j'}$ with $i < i'$ and $j < j'$. If such a quadruplet exists, then we may assume that $j = 1$ and $j' = 2$. Indeed, the distances of the points of F from x_i are strictly increasing, then maximal, then strictly decreasing, thus by the existence of j and j' the increasing sequence is non-empty and $d(x_i, y_1) < d(x_i, y_2)$. Moreover, $d(x_i, y_1) \leq d(x_i, y_j) < d(x_i, x_{i'})$.

Furthermore, we may also assume that $i' = k$. Indeed, if $d(x_i, x_k) < d(x_i, y_1)$, then $d(x_i, x_k) < \max\{d(x_i, x_{i'}), d(x_i, y_2)\}$, which implies that $x_i \prec_2 x_{i'} \prec_2 x_k \prec_2 y_2$ is a quadruplet violating compatibility of β_2 , a contradiction. This proves that $d(x_i, x_k) \geq d(x_i, y_1)$, hence x_i, x_k, y_1, y_2 is a violating quadruplet. Thus, it is enough to check $\text{sqcR}(x_i, x_k, y_1, y_2)$ and $\text{sqcR}(x_i, y_{l-1}, y_l, x_1)$ for every point x_i , and $\text{sqcR}(y_j, y_\ell, x_1, x_2)$ and $\text{sqcR}(y_j, x_{k-1}, x_k, y_1)$ for every point y_j . Obviously, all these conditions can be tested in total $O(n)$ time. \square

5.2. The algorithm. The previous discussion leads to an algorithm for finding a compatible order, presented in Algorithms 1 and 2. The function $\text{SORT}(x, S)$ sorts S by increasing values of $d(x, t)$ for $t \in S$ (we call this an x -sorting of the set S) and the function $\text{REVERSESORT}(S, x)$ sorts S by decreasing values of $d(x, t)$. The operator \uplus between two sequences represents their concatenation into a circular order. Notice that the same algorithm works for strictly circular and strictly quasi-circular Robinson dissimilarities, and that the algorithm always outputs an ordering, which may be arbitrary if the dissimilarity space is not strictly circular or strictly quasi-circular Robinson.

Algorithm 1 FINDCOMPATIBLEORDER

Input: A dissimilarity space (X, d) .

Output: A total ordering of X , compatible if (X, d) is (quasi-)circular Robinson.

```

1: let  $x \in X$ ,  $x' \in F_x$ 
2: let  $N = \{u \in X : d(u, x) \leq d(u, x')\}$ 
3: let  $F = \{u \in X : d(u, x') \leq d(u, x)\}$ 
4: if  $N \cap F \neq \emptyset$  then
5:   let  $y \in N \cup F$ 
6:   let  $X_1 = J(x, y) \cup J(y, x')$ 
7:   let  $X_2 = X \setminus X_1 \setminus \{x, x'\}$ 
8:   return  $\text{SORT}(x, X_1) \uplus \text{REVERSESORT}(x, X_2)$ 
9: else
10:  let  $z = \arg \max_{u \in N} d(x, u)$  and  $y = \arg \max_{u \in F} d(x', u)$ 
11:  let  $N' = J(x, z)$  and  $F' = J(x', y)$ 
12:  let  $X_N = \text{REVERSESORT}(x, N \setminus N') \uplus \text{SORT}(x, N')$ 
13:  let  $X_F = \text{SORT}(x', F \setminus F') \uplus \text{REVERSESORT}(x', F')$ 
14:  if  $\text{ORDERSAGREE}(X_N, X_F)$  then
15:    return  $X_N \uplus X_F$ 
16:  else
17:    return  $X_N \uplus \text{REVERSE}(X_F)$ 
```

Theorem 2. Algorithm 1 called to a strictly quasi-circular Robinson or a strictly circular Robinson dissimilarity (X, d) on n points produces a compatible circular order in $O(n \log n)$ time.

Proof. The correction of the algorithm follows from Lemmas 7 to 10. Namely, Lemma 8 covers the case $N \cap F \neq \emptyset$ (Lines 4 to 8), while Lemma 9 covers the case $N \cap F = \emptyset$ (Lines 10 to 17). From these lemmas and Lemma 7 it follows that the circular orders returned in Lines 8, 15 and 17 are

Algorithm 2 ORDERSAGREE

Input: A dissimilarity space (X, d) , a partition $X = N \cup F$ with $N = \{x_1, \dots, x_k\}$ and $F = \{y_1, \dots, y_\ell\}$

Output: whether the order $N \uplus F$ may be compatible based on Lemma 10

```

1: if  $k = 1$  or  $\ell = 1$  then
2:   return true
3: for all  $i \in \{1, 2, \dots, k\}$  do
4:   if not  $\text{sqcR}(x_i, x_k, y_1, y_2)$  or not  $\text{sqcR}(x_i, y_{\ell-1}, y_{\ell-2}, x_1)$  then
5:     return false
6: for all  $i \in \{1, 2, \dots, \ell\}$  do
7:   if not  $\text{sqcR}(y_i, y_\ell, x_1, x_2)$  or not  $\text{sqcR}(y_i, x_{k-1}, x_{k-2}, y_1)$  then
8:     return false
9: return true

```

the only possible compatible circular orders for (X, d) . Since (X, d) is strictly circular Robinson or strictly quasi-circular Robinson, we can apply Lemma 10 to deduce that one of these circular orders is indeed compatible. The complexity of the algorithm is dominated by the time to sort the lists, as every other operation can easily be implemented in either constant or linear time. \square

From Proposition 4 and Theorem 2 we immediately obtain the following result:

Corollary 1. *For a dissimilarity space (X, d) on n points, one can decide in optimal $O(n^2)$ time if (X, d) is strictly circular Robinson or strictly quasi-circular Robinson.*

The complexity in Theorem 2 is dominated by the time to sort the points by their distances to x or x' , and is actually tightly related to the complexity of sorting:

Proposition 5. *The problem of sorting a set Y of n distinct integers reduces linearly to the problem of finding a compatible circular order for a strictly quasi-circular Robinson dissimilarity.*

Proof. Given a set $Y \subseteq \mathbb{N}$, let $X = Y \cup \{z\}$ and let d be a dissimilarity on X defined by

$$\begin{aligned} d(y, z) &= \Delta + 1 \text{ for all } y \in Y, \\ d(y, y') &= |y - y'| \text{ for all } y, y' \in Y, \\ d(z, z) &= 0, \end{aligned}$$

where $\Delta = \max Y - \min Y$. Then it can be readily checked that (X, d) is a strictly quasi-circular Robinson dissimilarity, whose only two compatible orders induce an increasing or decreasing ordering of Y . This reduction is linear, as long as we encode the distance function d as an oracle, to avoid the computation of $\Theta(n^2)$ values. \square

5.3. On the number of compatible circular orders. From Algorithm 1, we can derive the following result about the number of compatible orders:

Proposition 6. *A strict quasi-circular Robinson space (X, d) has one or two compatible orders and their opposites. A strict circular Robinson space has one compatible order and its opposite.*

Proof. The first assertion is a direct consequence of Algorithm 1 and the proof of Theorem 2. Now, let (X, d) be a strict circular Robinson space with two compatible circular orders β and β' . Then $N \cap F = \emptyset$ and the arcs N and F are partitioned into N', N'' and F', F'' , respectively (see the proof of Lemma 9). Then the second compatible order β' is built from β by reversing N' and N'' . If the set N is empty, then this reversal does not change the order, thus $\beta' = \beta$. If F is empty, then this reversal builds the opposite order of the original one, thus $\beta' = \beta^{op}$. So, we can suppose with no loss of generality that there exist $y \in N'$ and $z \in F'$ and that the points y and z are on the same

arc $X_{xx'}^\beta$ of β . The arcs X_{xz}^β and $X_{yx'}^\beta$ are strictly Robinson, so $d(y, z) < \min\{d(x, z), d(y, x')\}$. By $\text{scR}(z, x, y, x')$ applied to β' , we must have $d(y, z) > d(x, z)$, which is in contradiction with $d(y, z) < \min\{d(x, z), d(y, x')\}$, whence β and β' cannot be both compatible. \square

If a strict quasi-circular Robinson space has two compatible orders and their opposites, then Algorithm 1 yields a bipartition of X into $N \cup F$. Next we prove that this happens exactly when there is a threshold value that clusters the dissimilarity space into two cliques:

Proposition 7. *Let (X, d) be a strict quasi-circular Robinson space. Then (X, d) admits two compatible orders and their opposites if and only if there exists a partition $X = N \cup F$ with $|N|, |F| > 1$ and $\delta \in \mathbb{R}^+$ such that for all $u, v \in X$, we have $d(u, v) > \delta$ if and only if $|\{u, v\} \cap N| = 1$.*

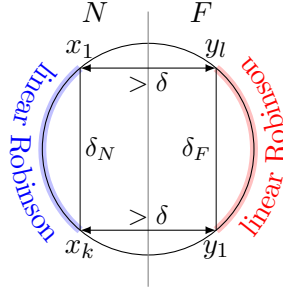


FIGURE 3. The structure of a strictly quasi-circular Robinson space with two non-opposite compatible orders, with $\delta = \max\{\delta_N, \delta_F\}$, as shown by Proposition 7. N and F have diameters δ_N and δ_F respectively, and all pairs between N and F have distance greater than δ . The proof that N (symmetrically, F) are linear Robinson follows easily from $\text{sqcR}(x_{i_1}, x_{i_2}, x_{i_3}, y_1)$ and $\text{sqcR}(x_{i_3}, y_1, x_{i_1}, x_{i_2})$

Proof. Suppose first that (X, d) admits two compatible orders and their opposites. By Lemmas 9 and 10, there is a bipartition $N \cup F$ with $N = \{x_1, x_2, \dots, x_k\}$ and $F = \{y_1, y_2, \dots, y_\ell\}$, such that the compatible orders are β (given by $N \uplus F$), β' (given by $N \uplus \text{REVERSE}(F)$), and their reverses. Notice that $k > 1$ and $\ell > 1$. Let $\delta_N = d(x_1, x_k)$ and $\delta_F = d(y_1, y_\ell)$. Then for any distinct $j, j' \in \{1, 2, \dots, \ell\}$, $\text{sqcR}(x_k, y_j, y_{j'}, x_1)$ (in β) and $\text{sqcR}(x_k, y_{j'}, y_j, x_1)$ (in β') we have:

$$\begin{aligned} d(x_k, y_{j'}) &> \min\{d(x_k, x_1), d(x_k, y_j)\}, \\ d(x_k, y_j) &> \min\{d(x_k, x_1), d(x_k, y_{j'})\}. \end{aligned}$$

Thus $\delta_N = d(x_1, x_k) < \min\{d(x_k, y_j), d(x_k, y_{j'})\}$. Analogously, $\delta_N < \min\{d(x_1, y_j), d(x_1, y_{j'})\}$. Then for any $i \in \{2, 3, \dots, k-1\}$, by $\text{sqcR}(y, x_1, x_i, x_k)$, $d(x_i, y) > \min\{d(y, x_1), d(y, x_k)\} > \delta_N$. This proves that $\min\{d(x, y) : x \in N, y \in F\} > \delta_N$.

Consequently, for any $y \in F$ and $i \in \{1, 2, \dots, k-1\}$, by $\text{sqcR}(x_k, y, x_1, x_i)$, $\delta_N = d(x_k, x_1) > \min\{d(x_k, y), d(x_k, x_i)\}$, which implies that $d(x_i, x_k) < \delta_N$. For $j \in \{i+1, i+2, \dots, k-1\}$, by $\text{sqcR}(x_i, x_j, x_k, y)$, $d(x_i, x_k) > \min\{d(x_i, x_j), d(x_i, y)\}$, which implies that $d(x_i, x_j) < \delta_N$, hence $\max\{d(u, v) : u, v \in N\} = \delta_N$. Analogously, we have $\max\{d(u, v) : u, v \in F\} = \delta_F$ and $\min\{d(x, y) : x \in N, y \in F\} > \delta_F$. Thus taking $\delta = \max\{\delta_N, \delta_F\}$ proves the assertion.

Conversely, suppose that (X, d) is a strictly quasi-circular Robinson space admitting such a bipartition $X = N \cup F$. Clearly N and F are balls of radius δ and thus, in any compatible order, by Proposition 1, N and F are arcs. Let $x_1 \leq x_2 \leq \dots \leq x_k \leq y_1 \leq y_2 \leq \dots \leq y_\ell$ be a compatible order β , with $N = \{x_1, x_2, \dots, x_k\}$ and $F = \{y_1, y_2, \dots, y_\ell\}$. Then, we can check that for any quadruplet $u < v < w < t$ of the circular order β' induced by $N \uplus \text{REVERSE}(F)$, $\text{sqcR}(u, v, w, t)$ holds. Indeed, the only nontrivial case (where the circular order is distinct for β and β' up to reversal) is when $u, v \in N$ and $w, t \in F$ (up to symmetry). In that case, we have $d(u, w) > \beta \geq d(u, v) \geq \min\{d(u, s), d(u, v)\}$,

that is $\text{sqcR}(u, v, w, t)$. This implies that β' is also compatible. Since $k, \ell > 1$, β and β' are not the reverse of each other, proving the proposition. \square

6. CIRCULAR SPACES RECOGNITION

Lemma 11. *Let (X, d) be a dissimilarity space and \triangleleft a circular order such that for all $x \in X$, $y, z \in X \setminus (F_x \cup \{x\})$, if $x \triangleleft y \triangleleft z \triangleleft Fx$ or $Fx \triangleleft z \triangleleft y \triangleleft x$, then $d(x, y) \leq d(x, z)$. Then (X, d, \triangleleft) is a circular space.*

Proof. Suppose, by a way of contradiction, that there exists $x \triangleleft y \triangleleft z \triangleleft t$ such that $d(x, z) < \min\{\max\{d(x, y), d(y, z)\}, \max\{d(x, t), d(t, z)\}\}$.

If $z \in F_x$, then $d(x, z) \geq d(x, y), d(x, t)$, and so $d(x, z) < d(z, t), d(y, t)$. So $x \notin F_z$. As F_z is an arc, we have either $z \triangleleft F_z \triangleleft z$ or $x \triangleleft F_x \triangleleft z$. If $x \triangleleft F_z \triangleleft z$, then $d(z, t) \leq d(z, x)$; if $z \triangleleft F_z \triangleleft x$, then $d(z, y) \leq d(z, x)$; In both case, we have a contradiction with $d(x, z) \geq d(x, y), d(x, t)$.

If $z \notin F_x$, say $x \triangleleft z \triangleleft F_x$ (the case $x \triangleleft F_x \triangleleft z$ is similar), then $d(x, y) \triangleleft d(x, z)$. So $d(x, z) < d(z, y)$ and $x \notin F_z$. By Proposition 2, we have $x \triangleleft z \triangleleft F_z$, and thus $d(z, y) \leq d(z, x)$, a contradiction. \square

Corollary 2. *By repeating Algorithm 1 for each $x \in X$, one determine if (X, d) is a circular space.*

Proposition 8. *Let (X, d, β) be a circular space and A be an mmodule of (X, d) . If A is made of k arcs (for β), with $k > 2$, then X is made of $2 \cdot k$ arcs I_1, \dots, I_{2k} and for all $i \neq j$, $i' \neq j'$, $d(I_i, I_j) = d(I_{i'}, I_{j'})$.*

Proof. Suppose, by a way of contradiction, that A contains three arcs A_1, A_2 and A_3 and let $x, y, z \in X \setminus A$ be such that $x \triangleleft A_1 \triangleleft y \triangleleft A_2 \triangleleft z \triangleleft A_3$. We have $d(x, y) \geq d(x, A), d(y, A)$, $d(x, z) \geq d(x, A), d(z, A)$ and $d(x, A) \geq \min\{d(x, y), d(x, z)\}$. So $d(x, A) = d(x, y) = d(x, z)$. Similarly, $d(y, A) = d(x, y) = d(y, z)$ and $d(z, A) = d(x, z) = d(y, z)$. So $X \setminus A$ is an mmodule made of three arcs A'_x, A'_y, A'_z and $d(A'_x, A'_y) = d(A'_x, A'_z) = d(A'_y, A'_z) = d(A, A')$. By symmetry, we get $d(A_1, A_2) = d(A_1, A_3) = d(A_2, A_3) = d(A, A')$. So X is made of six mmodules, each made of one arc. \square

7. CONCLUSION

In this paper, we presented a very simple algorithm which solves two variants of the strict circular seriation problem in optimal $O(n^2)$ time. Notice that the $O(n^2)$ time is entirely due to the verification of the result, while the computation of a compatible circular order (the main part of the algorithm) is in $O(n \log n)$ time. In addition, using the algorithm we proved some structural properties of strictly quasi-circular and strictly circular Robinson spaces. We also proved that any quasi-circular Robinson space is circular Robinson, a result which can find further applications.

The starting point of this research was the attempt to use mmodules and copoint decompositions, introduced in [5] for Robinson spaces, to recognize circular Robinson spaces. While for general circular Robinson spaces it is not yet clear how to use this approach, in case of strict circular Robinson spaces it turned out that mmodules, and, more generally, spheres $S_r(x)$ have size at most 2 and if they have size two, then the two points of $S_r(x)$ must be x -separated. However, some of the used results are valid for non-strict circular seriation problem. Extending our algorithm together with the tools developed in [5] to this problem will be the topic of future research.

Acknowledgement. This research was supported in part by ANR project DISTANCIA (ANR-17-CE40-0015) and has received funding from Excellence Initiative of Aix-Marseille - A*MIDEX (Archimedes Institute AMX-19-IET-009), a French "Investissements d’Avenir" Programme.

REFERENCES

- [1] S. Armstrong, C. Guzmán, and C.A. Sing Long, An optimal algorithm for strict circular seriation, *SIAM Journal on Mathematics of Data Science*, 3 (2021), 1223–1250.
- [2] J.E. Atkins, E.G. Boman, and B. Hendrickson, spectral algorithm for seriation and the consecutive ones problem, *SIAM Journal of Computing* 28 (1998), 297–310.
- [3] K.S. Booth and G.S. Lueker, Testing for the Consecutive Ones Property, interval graphs and graph planarity using PQ-tree algorithm, *Journal of Computer and System Sciences* 13 (1976), 335–379.
- [4] F. Brucker and C. Osswald, Hypercycles and dissimilarities, *Journal of Classification*, accepté, (2008).
- [5] M. Carmona, V. Chepoi, G. Naves, P. Préa, Modules in Robinson spaces, arXiv:2203.12386.
- [6] V. Chepoi and B. Fichet, Recognition of Robinsonian dissimilarities, *Journal of Classification* 14 (1997), 311–325.
- [7] R.R. Coifman, Y. Shkolnisky, F. J. Sigworth, and A. Singer, Graph laplacian tomography from unknown random projections, *IEEE Transactions on Image Processing*, 17 (2008), 1891–1899.
- [8] X. Evangelopoulos, A. J. Brockmeier, T. Mu, and J. Y. Goulermas, Circular object arrangement using spherical embeddings, *Pattern Recognition*, 103 (2020), p. 107192.
- [9] L.J. Hubert, Some applications of graph theory and related nonmetric techniques to problems of approximate seriation: The case of symmetric proximity measures, *British Journal of Mathematical Statistics and Psychology* 27 (1974), 133–153.
- [10] L. Hubert, P. Arabie, and J. Meulman, Graph-theoretic representations for proximity matrices through strongly-anti-robinson or circular strongly-anti-robinson matrices, *Psychometrika*, 63 (1998), 341–358.
- [11] L. Hubert, P. Arabie, and J. Meulman, Linear and circular unidimensional scaling for symmetric proximity matrices, *British Journal of Mathematical Statistics and Psychology*, 50 (1997), 253–284.
- [12] E.V. Huntington, A set of independent postulates for cyclic order, *Proc. Natl. Acad. Sci. U.S.A.*, 2 (1916), 630–631.
- [13] Y. Issartel, C. Giraud, and N. Verzelen, Optimal embedding on the sphere in non-parametric latent space models, 2021.
- [14] M. Laurent and M. Seminaroti, A Lex-BFS-based recognition algorithm for Robinsonian matrices, *Discrete Applied Mathematics*, 222 (2017), 151–165.
- [15] M. Laurent and M. Seminaroti, Similarity-first search: a new algorithm with application to Robinsonian matrix recognition, *SIAM Journal on Discrete Mathematics*, 31 (2017), 1765–1800.
- [16] J. Liu, D. Lin, G. G. Yardimici, and W. S. Noble, Unsupervised embedding of single-cell Hi-C data, *Bioinformatics*, 34 (2018), i96–i104.
- [17] Liu, Z., Lou, H., Xie, K., Wang, H., N. Chen, N., O.M. Aparicio, M.Q. Zhang, R. Jiang, and T. Chen, Reconstructing cell cycle pseudo time-series via single-cell transcriptome data, *Nature Communications*, 8 (2017), 22.
- [18] B. Mirkin and S. Rodin, *Graphs and Genes*, Springer-Verlag, 1984.
- [19] A. Quillot, Circular representation problem on hypergraphs, *Discrete Mathematics* 51 (3), 251–264.
- [20] P. Préa and D. Fortin, An optimal algorithm to recognize Robinsonian dissimilarities, *Journal of Classification*, 31 (2014), 351–385.
- [21] A. Recanati, T. Kerdreux, and A. d’Aspremont, Reconstructing latent orderings by spectral clustering, *CoRR* abs/1807.07122 (2018).
- [22] W.S. Robinson, A method for chronologically ordering archeological deposits, *American Antiquity* 16 (1951), 293–301.
- [23] M. Seston, A simple algorithm to recognize Robinsonian dissimilarities, *COMPSTAT’2008*, Porto.
- [24] M. Seston, *Dissimilarités de Robinson: Algorithmes de Reconnaissance et d’Approximation*, Ph.D. Thesis, Université de la Méditerranée, 2008.