



**HAL**  
open science

## A simple and optimal algorithm for strict circular seriation

Mikhael Carmona, Victor Chepoi, Guylain Naves, Pascal Pr ea

► **To cite this version:**

Mikhael Carmona, Victor Chepoi, Guylain Naves, Pascal Pr ea. A simple and optimal algorithm for strict circular seriation. *SIAM Journal on Mathematics of Data Science*, 2023, 5 (1), pp.201-221. 10.1137/22M149534 . hal-03714735v1

**HAL Id: hal-03714735**

**<https://hal.science/hal-03714735v1>**

Submitted on 26 Mar 2024 (v1), last revised 2 Apr 2024 (v2)

**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 ee au d ep ot et  a la diffusion de documents scientifiques de niveau recherche, publi es ou non,  emanant des  tablissements d'enseignement et de recherche fran ais ou  trangers, des laboratoires publics ou priv es.

# A SIMPLE AND OPTIMAL ALGORITHM FOR STRICT CIRCULAR SERIATION

MIKHAEL CARMONA\*<sup>†</sup>, VICTOR CHEPOI\*, GUYSLAIN NAVES\*, AND PASCAL PRÉA\*<sup>†</sup>

**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).

**Key words.** Circular seriation, circular Robinson matrix, circular compatible order, hypercycle.

**AMS subject classifications.** 68W05, 68R01, 68R12, 68T09

**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 (*linear*) *seriation problem* (called also the matrix reordering problem), introduced by Robinson [27], 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  $\mathbf{d}(x, z) \geq \max\{\mathbf{d}(x, y), \mathbf{d}(y, z)\}$  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*. 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*. Seriation (linear or circular) has numerous applications in data science, originating from various research areas: archeological dating, hypertext orderings, overlapping clustering, gene expression, DNA sequencing, DNA replication and 3D conformation, planar tomographic reconstruction, quadratic assignment problem, numerical ecology, sparse matrix ordering, musicology, matrix visualization methods, etc.

Due to its importance, 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 [23] consists in testing if the hypergraph of balls is an interval hypergraph and runs in  $O(n^4)$  time. Chepoi and Fichet [7] gave a simple divide-and-conquer algorithm running in  $O(n^3)$  time. Seston [29] presented another  $O(n^3)$  time algorithm, by using threshold graphs. In [28], 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 [24] presented an algorithm running in optimal  $O(n^2)$  time. The efficiency of the algorithm of [24] is due to the use of the PQ-trees of Booth and Lueker [4] as a data structure for encoding all compatible orders. Even if optimal, the algorithm of [24] is far from being simple. Subsequently, two new recognition algorithms were proposed by Laurent and Seminaroti: in [19] they

---

\*LIS, Aix-Marseille Université, CNRS, and Université de Toulon, Marseille, France ([mikhael.carmona](mailto:mikhael.carmona), [victor.chepoi](mailto:victor.chepoi), [guyslain.naves](mailto:guyslain.naves), [pascal.prea](mailto:pascal.prea))@lis-lab.fr

<sup>†</sup>École Centrale de Marseille, Marseille, France

presented an algorithm of complexity  $O(\alpha \cdot n)$  based on classical LexBFS traversal and divide-and-conquer (where  $\alpha$  is the depth of the recursion tree, which is at most the number of distinct elements of the input matrix), and in [20] 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 [6] we gave a simple and practical  $O(n^2)$  divide-and-conquer algorithm based on decompositions of dissimilarity spaces into modules (subsets of points not distinguishable from the outside). 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. The case when the eigenvector is not simple was considered in the recent paper by Concas et al. [11].

The circular seriation problem takes its origins in the papers by Hubert, Arabie, and Meulman [15, 17, 16]. More recently, circular seriation found interesting applications in planar tomographic reconstruction, gene expression, DNA replication and 3D conformation, see the papers [10, 22, 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 et al. [15, 17, 16] defined a class of circular Robinson dissimilarities via a certain 4-point condition. Brucker and Osswald [5] 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 [25], are the circular analogs of interval hypergraphs and are the hypergraphs whose hyperedges can be represented as circular intervals (arcs). In the case of circular Robinson dissimilarities, requiring that the ball, the 2-ball, or the cluster hypergraphs are hypercycles lead to three different classes of dissimilarity spaces. Their structural properties have been thoroughly studied by Brucker and Osswald [5]. The dissimilarities whose ball hypergraph is a hypercycle is the most general one and was characterized in [5] via a simple 4-point condition. We call such dissimilarities quasi-circular Robinson. To characterize the dissimilarities for which the 2-ball or the cluster hypergraphs are hypercycles, Brucker and Osswald [5] introduced the notion of pre-circular Robinson dissimilarities.

The algorithmic problem of recognizing circular Robinson dissimilarities was less studied and optimal or even subcubic time algorithms for different versions of this problem are not known. A spectral approach to circular seriation was developed in the papers [12, 13, 26]. Hsu and McConnell [14] showed how to efficiently recognize the hypercycles, by using a data structure, which they call PC-tree, and which is a generalization of well-known PQ-trees. Using this result, Brucker and Osswald [5] designed cubic time algorithms for recognizing quasi-circular Robinson dissimilarities. It comes to some surprise when recently Armstrong, Guzmán, and Sing Long [1] presented an optimal  $O(n^2)$  time algorithm for the recognition of strict quasi-circular Robinson dissimilarities. Among other tools, their algorithm uses PQ-trees.

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, 5, 17]. 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  $<$  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  $<$  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. The results of [1] and of this paper can be viewed as the first step toward designing efficient algorithms for circular seriation. Designing algorithms which solve the circular seriation problem in subcubic time (say, in  $O(n^2 \log n)$ ) or in optimal time  $O(n^2)$  is an interesting *open question*. While a random circular Robinson space is almost surely a strict circular Robinson space, circular Robinson dissimilarities are an ideal case and the dissimilarity matrices, which are measured only approximatively, fail to satisfy the circular Robinson property. Therefore, the *second open algorithmic problem* is the approximation of a dissimilarity space by a circular Robinson space. In the linear case, these fitting problems are NP-hard for  $\ell_1$ -norm [3] and  $\ell_\infty$ -norm [8] (no polynomial time algorithm is known for the  $\ell_p$ -norms with  $1 < p < \infty$ ) and a constant factor approximation algorithm for  $\ell_\infty$ -fitting problem was designed in [9].

The remaining part of the paper is organized as follows. In Section 2 we define the various notions of circular Robinson spaces. In Section 3 we prove our first main result about the characterization of pre-circular Robinson spaces. In Section 4 we present the main properties of quasi-circular and strict circular Robinson spaces. We also show how to efficiently verify if a dissimilarity space is (strictly) quasi-circular Robinson or (strictly) circular Robinson with respect to a fixed circular order. Section 4 can be viewed as the preparatory work for the recognition algorithm, which is described in Section 5. Using the algorithm, we prove that a strict quasi-circular Robinson space has one or two compatible orders and their opposites and that a strict circular Robinson space has one compatible order and its opposite. Section 6 provides a brief conclusion.

**2. Preliminaries.** In this section, first we introduce the notions related to the dissimilarity spaces and linear Robinson spaces. Then, we define the circular orders and the different types of circular Robinson spaces.

**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  $\mathbf{d}$  from  $X^2$  to the nonnegative real numbers such that  $\mathbf{d}(x, y) = \mathbf{d}(y, x) \geq 0$  and  $\mathbf{d}(x, y) = 0$  if and only if  $x = y$ . Then  $\mathbf{d}(x, y)$  is called the *distance* between  $x, y$  and  $(X, \mathbf{d})$  is called a *dissimilarity space*. The *ball* (respectively, the *sphere*) of radius  $r \geq 0$  centered at  $x \in X$  is the set  $B_r(x) := \{y \in X : \mathbf{d}(x, y) \leq r\}$  (respectively,  $S_r(x) := \{y \in X : \mathbf{d}(x, y) = r\}$ ). The *eccentricity* of a point  $x$  is  $r_x := \max\{\mathbf{d}(x, y) : y \in X\}$ . Given a point  $x \in X$ , a point  $y \in X$  is called a *farthest neighbor* of  $x$  if  $\mathbf{d}(x, y) = r_x$ . Denote by  $F_x$  the set of all farthest neighbors of  $x$  and note that  $F_x = S_{r_x}(x)$ . The *distance* between two subsets  $A, B$  of  $X$  is  $\mathbf{d}(A, B) = \min\{\mathbf{d}(a, b) : a \in A, b \in B\}$ .

**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  $\mathbf{d}(x, z) \geq \max\{\mathbf{d}(x, y), \mathbf{d}(y, z)\}$ . If  $\mathbf{d}$  and  $<$  are compatible, then  $\mathbf{d}$  is also compatible with the linear order  $<^{\text{op}}$  opposite to  $<$ . If a dissimilarity space  $(X, \mathbf{d})$  admits a compatible order, then  $d$  is said to be *Robinson* and  $(X, \mathbf{d})$  is called a *Robinson space*. Equivalently,  $(X, \mathbf{d})$  is Robinson if its distance matrix  $D$  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  $(X, \mathbf{d})$  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  $\mathbf{d}(x, y) \leq \max\{\mathbf{d}(x, z), \mathbf{d}(y, z)\}$  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  $\mathbb{R}$  such that  $p_1 < \dots < p_n$ . Any line-distance has exactly two compatible orders: the order  $p_1 < \dots < p_n$  and its opposite.

A dissimilarity  $\mathbf{d}$  on a set  $X$  is *strictly Robinson* if there exists a linear order, said *compatible*, on  $X$  such that  $x < y < z$  implies  $\mathbf{d}(x, z) > \max\{\mathbf{d}(x, y), \mathbf{d}(y, z)\}$ .  $(X, \mathbf{d})$  is then called a *strict circular Robinson space*.

**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  $\beta$  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 [18]: 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  $\beta(u, v, x)$ .

From the definition it follows that only triplets of distinct points can be in the relation  $\beta$ . It also follows that the reverse relation  $\beta^{op}$ , defined by setting  $\beta^{op}(u, v, w)$  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  $\beta$  corresponds to the counterclockwise order of  $C$  and  $\beta^{op}$  to the clockwise order of  $C$ . Given a circular order  $\beta$  and three distinct points  $u, v, w$ , we will write  $u < v < w$  if  $\beta(u, v, w)$  holds.

For a sequence of points  $x_1, x_2, \dots, x_\ell$  containing at least three distinct points, we will write  $x_1 <_\beta x_2 <_\beta \dots <_\beta x_\ell$  (or simply  $x_1 < x_2 < \dots < x_\ell$ , if no ambiguity occurs) if for any  $1 \leq i < j < k \leq \ell$  with  $x_i, x_j, x_k$  distinct,  $\beta(x_i, x_j, x_k)$  holds. We will use the following properties of circular orders:

**PROPOSITION 2.1.** *Let  $\beta$  be a circular order on  $X$  and  $x_1, x_2, \dots, x_\ell \in X$  such that  $x_1 <_\beta x_2 <_\beta \dots <_\beta x_\ell$ . Then:*

- (i)  $x_2 < x_3 < \dots < x_\ell < x_1$ ,
- (ii) if  $1 \leq i < j < k \leq \ell$  and  $x_i = x_j \neq x_k$  hold, then for any  $m \in \{i, \dots, j\}$  we have  $x_m = x_i$ .

*Proof.* (i): We must prove that for any  $1 < i < j$  with  $x_i, x_j, x_1$  distinct,  $\beta(x_i, x_j, x_1)$  holds. This follows from (CO3) and  $x_1 <_\beta x_2 <_\beta \dots <_\beta x_\ell$ .

(ii): Assume that  $x_m \neq x_i = x_j$ . If  $x_m \neq x_k$ , then  $x_i = x_j$ ,  $x_m$  and  $x_k$  are distinct, with  $\beta(x_i, x_m, x_k)$  and  $\beta(x_m, x_j, x_k)$  (because  $x_1 <_\beta x_2 <_\beta \dots <_\beta x_k$ ), by (CO3)  $\beta(x_m, x_k, x_i)$  and  $\beta(x_j, x_k, x_m)$ , contradicting (CO2) as  $x_i = x_j$ . If  $x_m = x_k$ , by assumption there must be a point  $x_p$  distinct from  $x_i = x_j$  and  $x_m = x_k$ . By (i), we may assume that  $p = 1$ . Then  $\beta(x_1, x_i, x_k)$  and  $\beta(x_1, x_m, x_j)$  hold. By (CO3), we get  $\beta(x_i, x_k, x_1)$ , that is  $\beta(x_j, x_m, x_1)$  holds, contradicting (CO2).  $\square$

We say that a nonempty proper subset  $A$  of  $X$  is an *arc* of a circular order  $\beta$  on  $X$  if there are no four distinct points  $u, v \in A$  and  $x, y \in X \setminus A$  such that  $u < x < v < y$ . From the definition, 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  $x < y < z$ , then  $X_{xy}^\beta \subset X_{xz}^\beta$  and  $X_{yz}^\beta \subset X_{xz}^\beta$ . This implies that if Notice also that if  $x_1 <_\beta x_2 <_\beta \dots <_\beta x_m$ , then  $X_{x_1 x_m}^\beta$  is the union of the arcs  $X_{x_1 x_2}^\beta, \dots, X_{x_{m-1} x_m}^\beta$  and thus  $X$  is the union of the arcs  $X_{x_1 x_2}^\beta, \dots, X_{x_{m-1} x_m}^\beta, X_{x_m x_1}^\beta$ . For  $x, y \in X$  and  $Z \subset X$ , we write  $x < y < Z$  if for all  $z \in Z$  we have  $x < y < z$ .

Arcs are for circular orders what intervals are for linear orders. Thus the arcs can be viewed as arcs of a circle ordered counterclockwise: the arc  $X_{xy}^\beta$  is obtained by traversing the cycle counterclockwise from  $x$  to  $y$ . The intersection of two arcs is not necessarily an arc. However, we can use this geometric interpretation of arcs to prove the following elementary observation:

**LEMMA 2.2.** *Let  $A$  and  $B$  be two arcs of a circular ordered set  $(X, \beta)$ . If there exists  $x \in$*

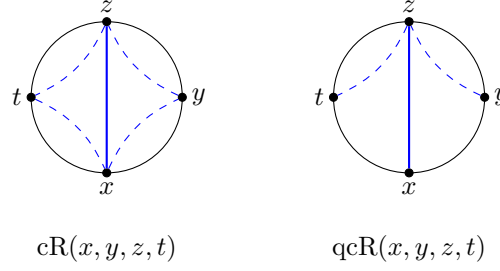


FIG. 2.1. Illustration of the constraints defined by one-side Robinson and quasi one-side Robinson.

$X \setminus (A \cup B)$ , then  $A \cap B$  is an arc or is empty. If there exists  $x \in B \setminus A$ , then  $A \setminus B$  is an arc or is empty.

*Proof.* Let  $A = X_{a', a''}^\beta$  and  $B = X_{b', b''}^\beta$ . Both assertions trivially holds when  $a' = a''$  or  $b' = b''$ . So, let  $a' \neq a''$  and  $b' \neq b''$ . First let  $x \notin A \cup B$ . Since  $x \notin A \cup B$ , we cannot have  $\beta(a', x, a'')$  or  $\beta(b', x, b'')$ . Thus, by (CO1), we can suppose that  $\beta(x, a', a'')$  and  $\beta(x, b', b'')$ . Suppose also, without loss of generality, that  $\beta(x, a', b')$ . If  $\beta(a', a'', b')$  holds, then  $A$  and  $B$  are disjoint. Thus, let  $\beta(a', b', a'')$  holds. Then  $B \subset A$  if  $\beta(b', b'', a'')$  and  $A \cap B$  is the arc  $X_{b', a''}^\beta$  if  $\beta(b', a'', b'')$ .

Now suppose that  $x \in B \setminus A$ . Since  $x \notin A$ , we can suppose that  $\beta(x, a', a'')$  holds. Furthermore, we can suppose that (1) either  $\beta(x, b'', a')$  or  $\beta(x, a', b'')$  holds and (2) either  $\beta(a'', b', x)$  or  $\beta(b', a'', x)$  holds. Combining the subcases, we conclude that  $A \setminus B$  is (a) the arc  $A = X_{a', a''}^\beta$  if  $\beta(x, b'', a')$  and  $\beta(a'', b', x)$ , (b) the arc  $X_{b'', a''}^\beta$  if  $\beta(x, a', b'')$  and  $\beta(a'', b', x)$ , (c) the arc  $X_{a', b'}^\beta$  if  $\beta(x, b'', a')$  and  $\beta(b', a'', x)$ , and (d) the arc  $X_{b'', b'}^\beta$  if  $\beta(x, a', b'')$  and  $\beta(b', a'', x)$ . This concludes the proof.  $\square$

We continue with several metric relations on four points, which will be used to define various types of circular Robinson spaces. Let  $(X, \mathbf{d})$  be a dissimilarity space,  $\beta$  be a circular order on  $X$  and  $x, y, z, t \in X$  such that  $x < y < z < t$ .

- The points  $x, y, z, t$  are *one-side Robinson*, and we denote it by  $\text{cR}(x, y, z, t)$ , if  $\mathbf{d}(x, z) \geq \min\{\max\{\mathbf{d}(x, y), \mathbf{d}(y, z)\}, \max\{\mathbf{d}(x, t), \mathbf{d}(t, z)\}\}$ . See Figure 2.1, left.
- The points  $x, y, z, t$  are *strictly one-side Robinson*, and we denote it by  $\text{scR}(x, y, z, t)$ , if  $\mathbf{d}(x, z) > \min\{\max\{\mathbf{d}(x, y), \mathbf{d}(y, z)\}, \max\{\mathbf{d}(x, t), \mathbf{d}(t, z)\}\}$ .
- The points  $x, y, z, t$  are *quasi one-side Robinson*, and we denote it by  $\text{qcR}(x, y, z, t)$ , if  $\mathbf{d}(x, z) \geq \min\{\mathbf{d}(y, z), \mathbf{d}(t, z)\}$ . See Figure 2.1, right.
- The points  $x, y, z, t$  are *strictly quasi one-side Robinson*, and we denote it by  $\text{sqcR}(x, y, z, t)$ , if  $\mathbf{d}(x, z) > \min\{\mathbf{d}(y, z), \mathbf{d}(t, z)\}$ .

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:

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

Now, we define the three types of circular dissimilarities investigated in this paper and their strict versions. A dissimilarity space  $(X, \mathbf{d})$  is called *pre-circular Robinson* if there exists a circular order  $\beta$ , 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  $(X, \mathbf{d})$  is called *circular Robinson* (respectively *strictly circular Robinson*) if there exists a circular order  $\beta$ , called a *compatible order* such that for all  $x, y \in X$ , either  $(X_{xy}^\beta, \mathbf{d})$  or  $(X_{yx}^\beta, \mathbf{d})$  is a Robinson space (respectively a strict Robinson space) and the restriction of  $<$  to the arc  $X_{xy}^\beta$  or respectively  $X_{yx}^\beta$  is a (linear) compatible order. Notice also that for all definitions, if a circular order  $\beta$  is compatible, then  $\beta^{op}$  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.

Hubert et al. [17] were the first to define circular Robinson spaces. We do not provide their definition here because they are particular pre-circular Robinson spaces (for their definition, see [5, 17]). 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  $(X, \mathbf{d})$ . Then for all  $x, y \in X$ , if  $x < y$  in the compatible order, then  $d$  is Robinson on the arc  $X_{xy}^\beta$ . As noted above, (strictly) circular Robinson spaces are (strictly) quasi-circular Robinson spaces. However not any circular order  $\beta$  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 2.2](#).



FIG. 2.2. A strict quasi-circular Robinson space  $(X = \{x, y, z, t\}, \mathbf{d})$  with a compatible circular order  $\beta$ . 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  $\text{scR}$  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  $\mathbf{d}(u, w)$ .

LEMMA 3.1. *Let  $(X, \mathbf{d})$  be a pre-circular Robinson space with a compatible circular order  $\beta$  and points  $u < y < y' < w < z < z'$ , where  $u$  and  $w$  are distinct. Then  $\mathbf{d}(u, w) \geq \min\{\mathbf{d}(y, y'), \mathbf{d}(z, z')\}$ . Moreover, if  $(X, \mathbf{d})$  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 an illustration, see [Figure 3.1](#)). For sake of contradiction, assume  $u < y < y' < w < z < z'$  is a counterexample with a minimum number of distinct points, implying that  $\mathbf{d}(u, w) < \mathbf{d}(y, y')$  and  $\mathbf{d}(u, w) < \mathbf{d}(z, z')$ . Then for  $u < y < y' < w$  we obtain the following inequalities:

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

The first inequality follows from  $\text{cR}(u, y, y', w)$ , the second inequality follows from the (easily verifiable) fact that for any four reals  $a_1, a_2, b_1, b_2$  and for any  $i, j \in \{1, 2\}$  we have

$$\min\{\max\{a_1, a_2\}, \max\{b_1, b_2\}\} \geq \min\{a_i, b_j\}$$

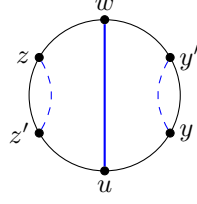


FIG. 3.1. Illustration of [Lemma 3.1](#). The solid blue line is longer than at least one of the dashed lines.

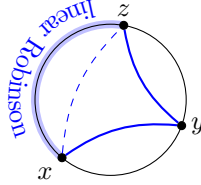


FIG. 3.2. Illustration of [Lemma 3.2](#). When  $\mathbf{d}(x, z)$  is not the maximum of the three dissimilarities, then the arc  $X_{zx}^\beta$  is a linear Robinson space.

and, finally, the third inequality is implied by the initial condition  $\mathbf{d}(u, w) \geq \min\{\mathbf{d}(y, y'), \mathbf{d}(z, z')\}$ . Consequently,  $\mathbf{d}(u, y') \geq \mathbf{d}(w, u)$ . If  $\mathbf{d}(u, y') = \mathbf{d}(w, u)$ , then  $u \triangleleft y \triangleleft w \triangleleft z \triangleleft z'$  is a counterexample. If  $\mathbf{d}(u, y') > \mathbf{d}(w, u)$ , then  $u \triangleleft u \triangleleft y' \triangleleft w \triangleleft z \triangleleft z'$  is a counterexample. Consequently, by the minimality of the counterexample  $u \triangleleft y \triangleleft y' \triangleleft w \triangleleft z \triangleleft z'$ , we conclude that either  $u = y$  or  $w = y'$  holds. Symmetrically, applying for  $w \triangleleft z \triangleleft z' \triangleleft u$  the same reasoning as for  $u \triangleleft y \triangleleft y' \triangleleft w$  with condition  $\text{cR}(w, z, z', u)$  instead of  $\text{cR}(u, y, y', w)$ , we deduce that either  $z = w$  or  $z' = u$  holds. 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  $\text{cR}(u, y'', w, z'')$ ,

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

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 assumption that  $u \triangleleft y \triangleleft y' \triangleleft w \triangleleft z \triangleleft z'$  is a counterexample.  $\square$

As a consequence we have:

**LEMMA 3.2.** *Let  $(X, \mathbf{d})$  be a pre-circular Robinson space with a compatible circular order  $\beta$  and  $x, y, z$  be three arbitrary points of  $X$  such that  $x \triangleleft y \triangleleft z$ . If  $\mathbf{d}(x, z) < \max\{\mathbf{d}(x, y), \mathbf{d}(y, z)\}$ , then  $(X_{zx}^\beta, \mathbf{d})$  is a Robinson space.*

*Proof.* Let  $y', y'' \in \{x, y, z\}$  be such that  $\mathbf{d}(x, z) < \mathbf{d}(y', y'')$  and  $x \triangleleft y' \triangleleft y'' \triangleleft z$  (for an illustration, see [Figure 3.2](#)). 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  $\mathbf{d}(u, w) < \max\{\mathbf{d}(u, v), \mathbf{d}(v, w)\}$ , let  $v', v'' \in \{u, v, w\}$  be such that  $u \triangleleft v' \triangleleft v'' \triangleleft w$  and  $\mathbf{d}(u, w) < \mathbf{d}(v', v'')$ . If  $\mathbf{d}(u, w) \leq \mathbf{d}(x, z)$ , then  $\mathbf{d}(u, w) < \mathbf{d}(y', y'')$ , and by [Lemma 3.1](#) on  $w \triangleleft y' \triangleleft y'' \triangleleft u \triangleleft v' \triangleleft v''$ , this is a contradiction. If  $\mathbf{d}(u, w) > \mathbf{d}(x, z)$ , then  $\mathbf{d}(x, z) < \mathbf{d}(v', v'')$ , and by [Lemma 3.1](#) 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 3.3.** *A dissimilarity space  $(X, \mathbf{d})$  is pre-circular Robinson if and only if  $(X, \mathbf{d})$  is circular Robinson.*



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

Conversely, let  $(X, \mathbf{d})$  be a pre-circular Robinson space and  $\beta$  be a compatible circular order. Pick any pair of points  $a, b$  of  $X$ . If  $(X_{ab}^\beta, \mathbf{d})$  is not a Robinson space, then there exists three points  $x, y, z \in X_{ab}^\beta$  such that  $x < y < z$  and  $\mathbf{d}(x, z) < \max\{\mathbf{d}(x, y), \mathbf{d}(y, z)\}$ . By [Lemma 3.2](#),  $(X_{zx}^\beta, \mathbf{d})$  is a Robinson space. Since  $X_{ba}^\beta \subset X_{zx}^\beta$ , this proves that  $(X_{ba}^\beta, \mathbf{d})$  is Robinson, establishing that  $(X, \mathbf{d})$  is a circular Robinson space and  $\beta$  is a compatible circular order.  $\square$

As a consequence,  $(X, \mathbf{d})$  is a *strictly circular Robinson space* if and only 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. Notice that these properties are also true for (strictly) circular Robinson spaces. We start with the following characterization of quasi-circular Robinson spaces of [\[5\]](#):

**PROPOSITION 4.1.** *[5] A dissimilarity space  $(X, \mathbf{d})$  is quasi-circular Robinson if and only if there exists a circular order  $\beta$  such that for any  $x \in X$  and  $r \in \mathbb{R}^+$ , the ball  $B_r(x)$  and its complement  $X \setminus B_r(x) = \{t \in X : \mathbf{d}(x, t) > r\}$  are arcs of  $\beta$ .*

*Proof.* First suppose that  $(X, \mathbf{d})$  is a quasi-circular Robinson space and  $\beta$  is a compatible circular order. Let  $B_r(x)$  be any ball of  $(X, \mathbf{d})$ . We will show that  $X \setminus B_r(x)$  is an arc; since the complement of an arc is an arc, this will also show that  $B_r(x)$  is an arc. Let  $y, y' \in X \setminus B_r(x)$  and suppose, with no loss of generality, that  $x < y < y'$ . Let  $z \in X_{yy'}^\beta$ . If  $z \notin X \setminus B_r(x)$ , then  $\mathbf{d}(z, x) \leq r < \mathbf{d}(y, x), \mathbf{d}(y', x)$ , contradicting the condition  $\text{qCR}(z, y', x, y)$ . Consequently,  $X \setminus B_r(x)$  and  $B_r(x)$  are arcs. Conversely, suppose that there exists a circular order  $\beta$  such that each ball  $B_r(x)$  is an arc of  $\beta$ . Pick arbitrary points  $x, y, z, t \in X$  such that  $x < y < z < t$ . Let  $r = \mathbf{d}(x, z)$ . Since  $B_r(z)$  is an arc of  $\beta$  and  $\beta(x, y, z), \beta(z, t, x)$  hold, either  $y$  or  $t$  must belong to the ball  $B_r(z)$ . Consequently,  $\mathbf{d}(x, z) \geq \min\{\mathbf{d}(y, z), \mathbf{d}(t, z)\}$ , establishing  $\text{qCR}(x, y, z, t)$ .  $\square$

Let  $(X, \mathbf{d})$  be a quasi-circular Robinson space and  $\beta$  be a compatible circular order. For any point  $x \in X$ , recall that  $F_x$  consists of all farthest neighbors of  $x$  and  $r_x$  is the eccentricity of  $x$ . Let  $M_x := X \setminus (F_x \cup \{x\})$ . Note that  $M_x \cup \{x\}$  is a ball  $B_r(x)$  for some  $r$  that is strictly smaller but sufficiently close to  $r_x$ . Thus, by [Proposition 4.1](#),  $M_x \cup \{x\}$  and  $F_x$  are complementary arcs of  $\beta$ . Consequently, the set  $M_x$  is partitioned into two arcs  $L_x := \{t \in M_x : x < t < F_x\}$  and  $R_x := \{t \in M_x : F_x < t < x\}$  (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  $R_x$ .

The algorithmic importance of the sets  $L_x$  and  $R_x$  is due to the fact that, as the following lemma shows, the circular order of each of these sets is given by the order of the distances to  $x$ .

**LEMMA 4.2.** *Let  $(X, \mathbf{d})$  be a quasi-circular Robinson space,  $\beta$  be a compatible circular order, and  $x$  any point of  $X$ . If  $y, z \in L_x$  and  $x < y < z$  or  $y, z \in R_x$  and  $z < y < x$ , then  $\mathbf{d}(x, y) \leq \mathbf{d}(x, z)$ . Moreover, if  $(X, \mathbf{d})$  is strict quasi-circular, then  $\mathbf{d}(x, y) < \mathbf{d}(x, z)$ .*

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

LEMMA 4.3. *Let  $(X, \mathbf{d})$  be a strict quasi-circular Robinson space,  $\beta$  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  $y \triangleleft y' \triangleleft y''$ . Since  $X$  is covered by the arcs  $X_{yy'}^\beta, X_{y'y''}^\beta$ , and  $X_{y''y}^\beta$ , we can suppose that  $x$  belongs to  $X_{y''y}^\beta$ , i.e., that  $x \triangleleft y \triangleleft y' \triangleleft y''$ . By condition  $\text{sqcR}(y', y'', x, y)$ , we must have  $d(x, y') > \min\{d(x, y''), d(x, y)\}$ . Since  $d(x, y) = d(x, y') = d(x, y'') = r$ , this is impossible. Thus  $|S_r(x)| \leq 2$ . Since  $(X, \mathbf{d})$  is strict quasi-circular, by Lemma 4.2,  $|S_r(x) \cap L_x| \leq 1$  and  $|S_r(x) \cap R_x| \leq 1$ , proving the second assertion.  $\square$

**4.2. Differences between (strictly) circular and quasi-circular Robinson spaces.** In this subsection, we present two properties which distinguish (strictly) circular Robinson spaces and (strictly) quasi-circular Robinson spaces. Roughly speaking, in a circular space, for  $x, y \in X$ ,  $x' \in F_x$  and  $y' \in F_y$ ,  $xx'$  and  $yy'$  have to cross each other, but this is not the case for quasi-circular spaces (see Figure 4.1).

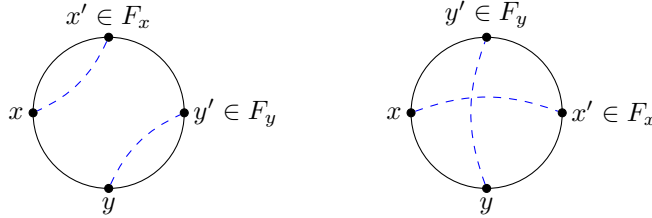


FIG. 4.1. *Illustration of Propositions 4.4. and 4.5. On the left, since  $xx'$  and  $yy'$  do not cross, the space can be quasi-circular but not circular. On the right, the space can be circular.*

PROPOSITION 4.4. *Let  $(X, \mathbf{d})$  be a circular Robinson space with a compatible order  $\beta$ . 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)  $x \triangleleft y \triangleleft x' \triangleleft y'$ ,
- (b)  $x \triangleleft y' \triangleleft x' \triangleleft y$ ,
- (c)  $\{y, y'\} \cap F_x \neq \emptyset$  or  $\{x, x'\} \cap F_y \neq \emptyset$ .

*Moreover if  $(X, \mathbf{d})$  is strictly circular Robinson, then either (a) or (b) holds.*

*Proof.* Suppose first that  $(X, \mathbf{d})$  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  $x < y < y' < x'$ , 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:

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

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

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

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

a contradiction.

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

Now, let  $\beta$  be a circular order on  $X$  that is compatible with a quasi-circular Robinson space  $(X, \mathbf{d})$ . We determine under which conditions  $\beta$  is not compatible with respect to the circular Robinson property of  $(X, \mathbf{d})$ .

**PROPOSITION 4.5.** *Let  $(X, \mathbf{d})$  be a (strict) quasi-circular Robinson space and  $\beta$  a compatible order, such that  $(X, \mathbf{d})$  is not (strict) circular Robinson with respect to  $\beta$ . Then there exist  $x, y \in X, x' \in F_x, y' \in F_y$  such that  $x < x' < y < y'$  or  $x < y' < y < x'$ . 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  $x < u < y < v$  be such that  $\text{scR}(x, u, y, v)$  does not hold:

$$(4.1) \quad \mathbf{d}(x, y) \leq \min\{\max\{\mathbf{d}(x, u), \mathbf{d}(u, y)\}, \max\{\mathbf{d}(x, v), \mathbf{d}(v, y)\}\}.$$

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

$$(4.2) \quad \mathbf{d}(x, y) > \min\{\mathbf{d}(x, u), \mathbf{d}(x, v)\},$$

$$(4.3) \quad \mathbf{d}(x, y) > \min\{\mathbf{d}(y, u), \mathbf{d}(y, v)\}.$$

Combining these inequalities, we get:

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

and then:

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

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

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

We may assume the first. Then  $\mathbf{d}(x, v) = \min\{\mathbf{d}(x, v), \mathbf{d}(x, u)\} < \mathbf{d}(x, y) \leq \min\{\mathbf{d}(x, u), \mathbf{d}(v, y)\} \leq \mathbf{d}(x, u)$  (by [Inequalities \(4.1\)](#) and [\(4.2\)](#)), which implies by the strict unimodality of distances from  $x$  that  $F_x \subseteq X_{uy}^\beta$ . Similarly,  $\mathbf{d}(y, u) < \mathbf{d}(x, y) \leq \mathbf{d}(y, v)$  which implies that  $F_y \subseteq X_{vx}^\beta$ . Consequently, if  $x' \in F_x$  and  $y' \in F_y$ , then we get  $x < x' < y < y'$ , as expected.

In the non-strict case, [Inequality \(4.1\)](#) becomes a strict inequality, while [Inequalities \(4.2\)](#) and [\(4.3\)](#) become non-strict inequalities. Combining these inequalities, we get the same conclusion as in the strict case. For example, in the first case we get that  $\mathbf{d}(x, v) = \min\{\mathbf{d}(x, v), \mathbf{d}(x, u)\} \leq \mathbf{d}(x, y) < \min\{\mathbf{d}(x, u), \mathbf{d}(v, y)\} \leq \mathbf{d}(x, u)$ , yielding  $\mathbf{d}(x, v) \leq \mathbf{d}(x, y) < \mathbf{d}(x, u)$  and  $\mathbf{d}(y, u) \leq \mathbf{d}(x, y) < \mathbf{d}(y, v)$ . 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  $x < x' < y < y'$ . Moreover,  $y \notin F_x$  and  $x \notin F_y$ . Since  $x < x' < v < y'$  and  $\mathbf{d}(x, v) < \mathbf{d}(x, x')$ , by  $\text{qcR}(x, x', v, y')$  we conclude that  $\mathbf{d}(x, v) \geq \mathbf{d}(x, y')$ , yielding  $y' \notin F_x$ . Analogously, one can show that  $x' \notin F_y$ .  $\square$

**4.3. Verification of compatibility.** In this subsection, given a circular order  $\beta$ , we describe how to check in  $O(n^2)$  whether a dissimilarity space  $(X, \mathbf{d})$  on  $n$  points is (strictly) quasi-circular Robinson or (strictly) circular Robinson with respect to  $\beta$ . This verification task can also be done in  $O(n^2)$  for strict circular Robinson spaces, as defined in [\[17\]](#). Then, we will show how to extend this result to strict versions of the other definitions of circular dissimilarities introduced by Brucker and Osswald [\[5\]](#), namely the dissimilarities whose 2-balls or clusters are arcs.

To test whether  $(X, \mathbf{d})$  is (strictly) quasi-circular Robinson with respect to  $\beta$ , by [Proposition 4.1](#) we have to test whether all balls of  $(X, \mathbf{d})$  are arcs of  $\beta$ . This can be done in the following way. Let  $D$  be the distance matrix of  $(X, \mathbf{d})$  ordered according to the circular order  $\beta$ . 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  $\beta$  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 [\[17\]](#), this testing task can be also done in  $O(n^2)$  time.

Next, we consider the task of testing whether  $(X, \mathbf{d})$  is (strictly) circular Robinson with respect to a circular order  $\beta$ . 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 4.5](#). 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  $\beta$ ) 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 4.5](#). If such elements exist, then by [Proposition 4.4](#),  $(X, \mathbf{d})$  is not (strictly) circular Robinson. Otherwise,  $(X, \mathbf{d})$  is (strictly) circular Robinson with respect to  $\beta$ . This testing task can be done in  $O(n^2)$  time. As a consequence, we have the following result:

**PROPOSITION 4.6.** *For a dissimilarity space  $(X, \mathbf{d})$  on  $n$  points and a circular order  $\beta$  on  $X$ , one can check in  $O(n^2)$  time whether, with respect to  $\beta$ ,  $(X, \mathbf{d})$  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  $\beta$  and that the restriction of  $\beta$  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  $X_{xy}^\beta$  metrically.** Given a dissimilarity space  $(X, \mathbf{d})$  and two distinct points  $x, y \in X$ , we set  $J^\circ(x, y) = \{u \in X : \mathbf{d}(x, y) > \max\{\mathbf{d}(x, u), \mathbf{d}(u, y)\}\}$  and  $J(x, y) = J^\circ(x, y) \cup \{x, y\}$ . In all results of this subsection, we assume that  $(X, \mathbf{d})$  is a strict quasi-circular Robinson space and  $\beta$  is an arbitrary compatible circular order on  $X$ .

LEMMA 5.1. *Let  $x < y < z$  be three points of  $X$  such that  $\mathbf{d}(x, y) \leq \min\{\mathbf{d}(x, z), \mathbf{d}(y, z)\}$ . 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  $\mathbf{d}(x, y) \leq \mathbf{d}(x, z)$ , we have  $\mathbf{d}(x, v) < \mathbf{d}(x, y)$ . By  $\text{sqcR}(y, z, v, x)$  and since  $\mathbf{d}(y, x) \leq \mathbf{d}(y, z)$ , we have  $\mathbf{d}(y, v) < \mathbf{d}(y, x)$ . 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  $\text{sqcR}(u, z, x, y)$ ,  $\mathbf{d}(x, u) > \min\{\mathbf{d}(x, y), \mathbf{d}(x, z)\} = \mathbf{d}(x, y)$ , 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  $N = \{u \in X : \mathbf{d}(u, x) \leq \mathbf{d}(u, x')\}$  and  $F = \{u \in X : \mathbf{d}(u, x) \geq \mathbf{d}(u, x')\}$ . Note that  $N \cup F = X$  and  $x \in N \setminus F, x' \in F \setminus N$ .

LEMMA 5.2.  *$N$  and  $F$  are arcs of  $\beta$ .*

*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  $y \in N$ . Since  $z < x' < x < y$  by (CO3) and  $\mathbf{d}(x, x') \geq \mathbf{d}(x, y)$  because  $x' \in F_x$ , by  $\text{sqcR}(z, x', x, y)$ , we have  $\mathbf{d}(x, y) < \mathbf{d}(x, z)$ . Since we also have  $y < z < x' < x$  by (CO3), by condition  $\text{sqcR}(y, z, x', x)$ ,

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

The second inequality follows from the fact that  $\mathbf{d}(x', z) \geq \mathbf{d}(x, z)$  since  $z \in N$  and  $\mathbf{d}(x', x) \geq \mathbf{d}(x, z)$  since  $x' \in F_x$ . Consequently,  $\mathbf{d}(x', y) > \mathbf{d}(x, y)$ , implying that  $y \in N$ . Symmetrically, if  $z < y < x < x'$  with  $z \in N$ , then  $y \in N$ . Hence,  $N$  is an arc of  $\beta$ .  $\square$

LEMMA 5.3. *If  $N \cap F \neq \emptyset$  and  $y$  is a point of  $X$  with  $\mathbf{d}(x, y) = \mathbf{d}(y, x')$ , 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 5.1 (a)). Since  $\mathbf{d}(x, y) = \mathbf{d}(y, x')$  and  $\mathbf{d}(x, y) \leq \mathbf{d}(x, x')$ , by Lemma 5.1 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  $\mathbf{d}(y, x') \leq \min\{\mathbf{d}(x', x), \mathbf{d}(y, x)\}$ . By Lemma 5.1 we conclude that  $X_{yx'}^\beta = J(y, x')$ . Finally, since  $x < y < x'$ , 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  $N \cap F \neq \emptyset$ , and  $y$  is a point with  $\mathbf{d}(x, y) = \mathbf{d}(y, x')$ , then according to Lemma 5.3, the circular order  $\beta$  such that  $x < y < x'$  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

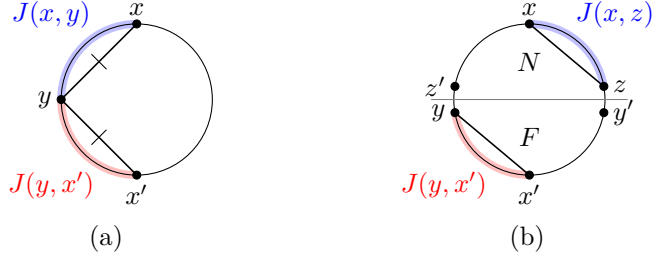


FIG. 5.1. Configurations occurring in (a) Lemma 5.3 and (b) Lemma 5.4. In (b), the positions of  $z$  and  $z'$  may be swapped, as well as those of  $y$  and  $y'$ .

$X \setminus X_{xx'}^\beta$ , by distances to  $x$ , by Lemma 4.2. Notice that in this case the compatible circular order  $\beta$  is unique up to reversal.

Thus, we may next assume that  $N \cap F = \emptyset$ . The points  $w \in N$  such that  $x \prec w \prec x'$  form an arc whose ordering is given by increasing distances from  $x$ . Analogously, the points  $w \in N$  such that  $x' \prec w \prec x$  form an arc whose ordering is given by decreasing distances from  $x$ . The points of  $F$  are similarly distributed into two arcs with respect to the distances from  $x'$ . Therefore, it is sufficient to partition the sets  $N \setminus \{x\}$  and  $F \setminus \{x'\}$  into such pairs  $N', N''$  and  $F', F''$ , respectively. This is done by the next lemma.

LEMMA 5.4. *If  $N \cap F = \emptyset$ , then there exist  $z, z' \in N$  and  $y, y' \in F$  and two bipartitions  $N \setminus \{x\} = N' \cup N''$ ,  $F \setminus \{x'\} = F' \cup F''$  such that for any compatible order  $\beta$  on  $X$  we have  $\{N' \cup \{x\}, N'' \cup \{x\}\} = \{X_{zx}^\beta, X_{xz'}^\beta\}$  and  $\{F' \cup \{x'\}, F'' \cup \{x'\}\} = \{X_{yx'}^\beta, X_{x'y'}^\beta\}$ . The sets  $N', N'', F', F''$  and the points  $z, z', y, y'$  can be computed in  $O(n)$  time.*

*Proof.* Assume that  $N \neq \{x\}$ , and let  $z \in N$  with  $\mathbf{d}(x, z)$  maximal (see Figure 5.1 (b)). Then applying Lemma 5.1 to  $x, z, x'$ , we have that  $J(x, z)$  is either  $X_{xz}^\beta$  or  $X_{zx}^\beta$  (depending of whether  $x \prec z \prec x'$  or  $x' \prec z \prec x$ ). We denote  $N' = J(x, z)$ . If  $N'' = N \setminus N' \neq \emptyset$ , let  $z' \in N''$  with  $\mathbf{d}(x, z')$  maximal. By Lemma 5.1,  $J(x, z')$  is either  $X_{z'x}^\beta$  or  $X_{xz'}^\beta$ . By Lemma 4.2,  $z$  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  $\mathbf{d}(x', y) < \mathbf{d}(y, x) \leq \mathbf{d}(x', x)$ . Thus we can also use Lemma 5.1 and get similarly that there exist points  $y, y' \in F$  (where  $y \in F$  with  $\mathbf{d}(x', y)$  maximal,  $F' = J(x', y)$ , and  $y' \in F'' = F \setminus F'$  with  $\mathbf{d}(x', y')$  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 definitions, 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 the points of  $N', N''$  and  $F', F''$  by their distances to  $x$  and to  $x'$  takes  $O(n \log n)$  time. Then Lemma 5.4 allows to partition the compatible circular order into two ordered sequences. In the first sequence,  $N$  is ordered into  $N = \{x_1, x_2, \dots, x_k\}$  by taking  $N'$  in decreasing order of distances from  $x$ , followed by  $x$  and then  $N''$  in increasing order of distances from  $x$ ; the second order is the reverse of the first order. Similarly,  $F$  is ordered into  $F = \{y_1, y_2, \dots, y_l\}$  using  $F', F''$  and the distances from  $x'$  and its reverse order. This leads to four possibilities to compose the two ordered



(up to reversal) sequences  $N = \{x_1, x_2, \dots, x_k\}$  and  $F = \{y_1, y_2, \dots, y_l\}$  into a compatible circular order. Notice that up to symmetry, this reduces to only two possibilities. The next lemma gives a criterion to decide which one of the two options is valid.

LEMMA 5.5. *Let  $N = \{x_1, x_2, \dots, x_k\}$  and  $F = \{y_1, y_2, \dots, y_l\}$  be the ordered sequences defined as above. Let  $\beta_1$  and  $\beta_2$  be the two circular orders on  $X$  defined by setting*

$$(a) \quad x_1 \prec_{\beta_1} x_2 \prec_1 \dots \prec_{\beta_1} x_k \prec_{\beta_1} y_1 \prec_{\beta_1} y_2 \prec_{\beta_1} \dots \prec_{\beta_1} y_l,$$

$$(b) \quad x_k \prec_{\beta_2} x_{k-1} \prec_{\beta_2} \dots \prec_{\beta_2} x_1 \prec_{\beta_2} y_1 \prec_{\beta_2} y_2 \prec_{\beta_2} \dots \prec_{\beta_2} y_l.$$

*One can decide which of these two circular orders  $\beta_1, \beta_2$  (possibly both) is compatible in  $O(n)$  time.*

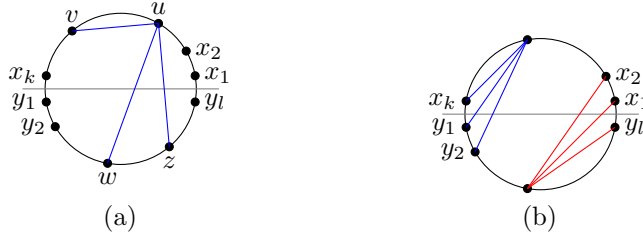


FIG. 5.2. In (a), a configuration occurring in the proof of Lemma 5.5. If  $\beta_1$  is not compatible for the quadruplet  $(u, v, w, z)$ , then it is not compatible for the quadruplet  $(u, x_k, y_1, y_2)$ . In (b), an illustration of the two families of quadruplets which are enough to check the compatibility on.

*Proof.* Suppose that  $\beta_1$  is not compatible and that  $\beta_2$  is compatible. Then there is a quadruplet  $u \prec_{\beta_1} v \prec_{\beta_1} w \prec_{\beta_1} z$  with  $\mathbf{d}(u, w) \leq \min\{\mathbf{d}(u, v), \mathbf{d}(u, z)\}$ . Since  $\beta_2$  is a compatible circular order, we must have that  $(u, v, w, z)$  is one of the four quadruplets  $(x_i, x_{i'}, y_j, y_{j'})$ ,  $(x_{i'}, y_j, y_{j'}, x_i)$ ,  $(y_j, y_{j'}, x_i, x_{i'})$ , or  $(y_{j'}, x_i, x_{i'}, y_j)$ , with  $i < i'$  and  $j < j'$ . Up to symmetry, we may assume the first without loss of generality. We may also assume that  $j = 1$  and  $j' = 2$ . Indeed, the distances from  $x_1$  of the points of  $F$ , from  $y_1$  to  $y_l$ , are strictly increasing, then maximal, then strictly decreasing, thus by the existence of  $j$  and  $j'$  the increasing sequence is non-empty and  $\mathbf{d}(x_i, y_1) < \mathbf{d}(x_i, y_2)$ . Moreover,  $\mathbf{d}(x_i, y_1) \leq \mathbf{d}(x_i, y_j) < \mathbf{d}(x_i, x_{i'})$ .

Furthermore, we may also assume that  $i' = k$ . Indeed, if  $\mathbf{d}(x_i, x_k) < \mathbf{d}(x_i, y_1)$ , then  $\mathbf{d}(x_i, x_k) < \max\{\mathbf{d}(x_i, x_{i'}), \mathbf{d}(x_i, y_2)\}$ , which implies that  $x_i \prec_{\beta_2} x_{i'} \prec_{\beta_2} x_k \prec_{\beta_2} y_2$  is a quadruplet violating compatibility of  $\beta_2$ , a contradiction. This proves that  $\mathbf{d}(x_i, x_k) \geq \mathbf{d}(x_i, y_1)$ , hence  $x_i, x_k, y_1, y_2$  is a violating quadruplet. Thus, considering the three remaining symmetric cases, we obtain that if there is a violating quadruplet, then also there is a violating quadruplet of the form  $\{x_i, x_k, y_1, y_2\}$  or  $\{x_i, y_{l-1}, y_l, x_1\}$  for some point  $x_i$ , or  $\{y_j, y_l, x_1, x_2\}$  or  $\{y_j, x_{k-1}, x_k, y_1\}$  for some point  $y_j$ . There are at most  $2n$  such quadruplets in total, each of them may be checked in  $O(1)$  time, summing up to a complexity of  $O(n)$  time using Algorithm 5.1.  $\square$

**5.2. The algorithm.** The previous discussion leads to an algorithm for finding a compatible order, presented in Algorithms 5.1 and 5.2. The function  $\text{SORT}(x, S)$  sorts  $S$  by increasing values of  $\mathbf{d}(x, t)$  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  $\mathbf{d}(x, t)$ . The operator  $\text{++}$  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.

THEOREM 5.6. *Algorithm 5.2 called to a strictly quasi-circular Robinson or a strictly circular Robinson dissimilarity  $(X, \mathbf{d})$  on  $n$  points produces a compatible circular order in  $O(n \log n)$  time.*



**Algorithm 5.1** ORDERSAGREE

---

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

**Output:** whether the order  $N \uparrow F$  may be compatible based on [Lemma 5.5](#).

```

1: if  $k = 1$  or  $\ell = 1$  then
2:   return true
3: end if
4: for all  $i \in \{1, 2, \dots, k\}$  do
5:   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
6:     return false
7:   end if
8: end for
9: for all  $i \in \{1, 2, \dots, \ell\}$  do
10:  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
11:    return false
12:  end if
13: end for
14: return true

```

---

**Algorithm 5.2** FINDCOMPATIBLEORDER

---

**Input:** A dissimilarity space  $(X, \mathbf{d})$ .

**Output:** A total ordering of  $X$ , compatible if  $(X, \mathbf{d})$  is (quasi-)circular Robinson.

```

1: let  $x \in X, x' \in F_x$ 
2: let  $N = \{u \in X : \mathbf{d}(u, x) \leq \mathbf{d}(u, x')\}$ 
3: let  $F = \{u \in X : \mathbf{d}(u, x') \leq \mathbf{d}(u, x)\}$ 
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  $\text{SORT}(x, X_1) \uparrow \text{REVERSE SORT}(x, X_2)$ 
9: else
10:  let  $z = \arg \max_{u \in N} \mathbf{d}(x, u)$  and  $y = \arg \max_{u \in F} \mathbf{d}(x', u)$ 
11:  let  $N' = J(x, z)$  and  $F' = J(x', y)$ 
12:  let  $X_N = \text{REVERSE SORT}(x, N \setminus N') \uparrow \text{SORT}(x, N')$ 
13:  let  $X_F = \text{SORT}(x', F \setminus F') \uparrow \text{REVERSE SORT}(x', F')$ 
14:  if  $\text{ORDERSAGREE}(X_N, X_F)$  then
15:    return  $X_N \uparrow X_F$ 
16:  else
17:    return  $X_N \uparrow \text{REVERSE}(X_F)$ 
18:  end if
19: end if

```

---

*Proof.* The correctness of the algorithm follows from [Lemmas 5.2](#) to [5.5](#). Namely, [Lemma 5.3](#) covers the case  $N \cap F \neq \emptyset$  ([lines 4](#) to [8](#)), while [Lemma 5.4](#) covers the case  $N \cap F = \emptyset$  ([lines 10](#) to [17](#)). From these lemmas and [Lemma 5.2](#) it follows that the circular orders returned in [lines 8](#), [15](#) and [17](#) are the only possible compatible circular orders for  $(X, \mathbf{d})$ . Since  $(X, \mathbf{d})$  is strictly circular Robinson or strictly quasi-circular Robinson, we can apply [Lemma 5.5](#) 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 4.6](#) and [Theorem 5.6](#) we immediately obtain the following result:

**COROLLARY 5.7.** *For a dissimilarity space  $(X, \mathbf{d})$  on  $n$  points, one can decide in optimal  $O(n^2)$  time if  $(X, \mathbf{d})$  is strictly circular Robinson or strictly quasi-circular Robinson.*

The complexity in [Theorem 5.6](#) 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 5.8.** *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

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

where  $\Delta = \max Y - \min Y$ . Then it can be readily checked that  $(X, \mathbf{d})$  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 5.2](#), we can derive the following result about the number of compatible orders:

**PROPOSITION 5.9.** *A strict quasi-circular Robinson space  $(X, \mathbf{d})$  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 5.2](#) and the proof of [Theorem 5.6](#). Now, let  $(X, \mathbf{d})$  be a strict circular Robinson space with two compatible circular orders  $\beta$  and  $\beta'$ . 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 5.4](#)). Then the second compatible order  $\beta'$  is built from  $\beta$  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  $\beta$ . The arcs  $X_{xz}^\beta$  and  $X_{yx'}^\beta$  are strictly Robinson, so  $\mathbf{d}(y, z) < \min\{\mathbf{d}(x, z), \mathbf{d}(y, x')\}$ . By  $\text{scR}(z, x, y, x')$  applied to  $\beta'$ , we must have  $\mathbf{d}(y, z) > \mathbf{d}(x, z)$ , which is in contradiction with  $\mathbf{d}(y, z) < \min\{\mathbf{d}(x, z), \mathbf{d}(y, x')\}$ , whence  $\beta$  and  $\beta'$  cannot be both compatible.  $\square$

If a strict quasi-circular Robinson space has two compatible orders and their opposites, then [Algorithm 5.2](#) 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 5.10.** *Let  $(X, \mathbf{d})$  be a strict quasi-circular Robinson space. Then  $(X, \mathbf{d})$  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  $\mathbf{d}(u, v) > \delta$  if and only if  $|\{u, v\} \cap N| = 1$ .*

*Proof.* Suppose first that  $(X, \mathbf{d})$  admits two compatible orders and their opposites. By [Lemmas 5.4](#) and [5.5](#), 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  $\beta$  (given by  $N \uparrow F$ ),  $\beta'$  (given by  $N \uparrow \text{REVERSE}(F)$ ), and their reverses. Notice that  $k > 1$  and  $\ell > 1$ . Let  $\delta_N = \mathbf{d}(x_1, x_k)$  and  $\delta_F = \mathbf{d}(y_1, y_\ell)$ . Then for any

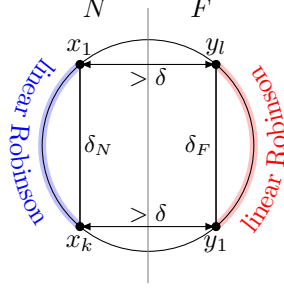


FIG. 5.3. The structure of a strictly quasi-circular Robinson space with two non-opposite compatible orders, with  $\delta = \max\{\delta_N, \delta_F\}$ , as shown by [Proposition 5.10](#).  $N$  and  $F$  have diameters  $\delta_N$  and  $\delta_F$  respectively, and all pairs between  $N$  and  $F$  have distance greater than  $\delta$ . 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  $\text{sqcR}(x_{i_3}, y_1, x_{i_1}, x_{i_2})$ .

distinct  $j, j' \in \{1, 2, \dots, \ell\}$ ,  $\text{sqcR}(x_k, y_j, y_{j'}, x_1)$  (in  $\beta$ ) and  $\text{sqcR}(x_k, y_{j'}, y_j, x_1)$  (in  $\beta'$ ) we have:

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

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

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

Conversely, suppose that  $(X, \mathbf{d})$  is a strictly quasi-circular Robinson space admitting such a bipartition  $X = N \cup F$ . Clearly  $N$  and  $F$  are balls of radius  $\delta$  and thus, in any compatible order, by [Proposition 4.1](#),  $N$  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  $\beta$ , 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  $u \leq v \leq w \leq t$  of the circular order  $\beta'$  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  $\beta$  and  $\beta'$  up to reversal) is when  $u, v \in N$  and  $w, t \in F$  (up to symmetry). In that case, we have  $\mathbf{d}(u, w) > \beta \geq \mathbf{d}(u, v) \geq \min\{\mathbf{d}(u, s), \mathbf{d}(u, v)\}$ , that is  $\text{sqcR}(u, v, w, t)$ . This implies that  $\beta'$  is also compatible. Since  $k, \ell > 1$ ,  $\beta$  and  $\beta'$  are not the reverse of each other, proving the proposition.  $\square$

**6. Conclusion.** In this paper, we presented a very simple algorithm which solves the strict quasi-circular and strict circular seriation problems 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 pre-circular Robinson space is circular Robinson, a result which can find further applications. As we already noticed in the introduction, designing an algorithm which solves the circular seriation problem in  $O(n^2)$  (or even in  $O(n^2 \log n)$  time) is an interesting *open question*. Designing approximation algorithms for fitting a dissimilarity by a circular Robinson dissimilarity is another *open problem*.

**Acknowledgement.** We would like to acknowledge the referees for their careful reading of the manuscript and useful suggestions and comments. 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 J. Math. Data Sci., 3 (2021), pp. 1223–1250, <https://doi.org/10.1137/21M139356X>, <https://doi.org/10.1137/21M139356X>.
- [2] J. E. ATKINS, E. G. BOMAN, AND B. HENDRICKSON, *A spectral algorithm for seriation and the consecutive ones problem*, SIAM J. Comput., 28 (1998), pp. 297–310, <https://doi.org/10.1137/S0097539795285771>, <https://doi.org/10.1137/S0097539795285771>.
- [3] J.-P. BARTHÉLEMY AND F. BRUCKER, *NP-hard approximation problems in overlapping clustering*, J. Classif., 18 (2001), pp. 159–183.
- [4] K. S. BOOTH AND G. S. LUEKER, *Testing for the consecutive ones property, interval graphs, and graph planarity using pq-tree algorithms*, J. Comput. Syst. Sci., 13 (1976), pp. 335–379, [https://doi.org/10.1016/S0022-0000\(76\)80045-1](https://doi.org/10.1016/S0022-0000(76)80045-1), [https://doi.org/10.1016/S0022-0000\(76\)80045-1](https://doi.org/10.1016/S0022-0000(76)80045-1).
- [5] F. BRUCKER AND C. OSSWALD, *Hypercycles and dissimilarities*, Journal of Classification, accepted, (2008).
- [6] M. CARMONA, V. CHEPOI, G. NAVES, AND P. PRÉA, *Modules in robinson spaces*, CoRR, abs/2203.12386 (2022), <https://doi.org/10.48550/arXiv.2203.12386>, <https://doi.org/10.48550/arXiv.2203.12386>, <https://arxiv.org/abs/2203.12386>.
- [7] V. CHEPOI AND B. FICHET, *Recognition of robinsonian dissimilarities*, Journal of Classification, 14 (1997), pp. 311–325.
- [8] V. CHEPOI, B. FICHET, AND M. SESTON, *Seriation in the presence of errors: NP-hardness of  $l_\infty$ -fitting Robinson structures to dissimilarity matrices*, J. Classif., 26 (2009), pp. 279–296.
- [9] V. CHEPOI AND M. SESTON, *Seriation in the presence of errors: A factor 16 approximation algorithm for  $l_\infty$ -fitting Robinson structures to distances*, Algorithmica, 59 (2011), pp. 521–568.
- [10] R. R. COIFMAN, Y. SHKOLNISKY, F. J. SIGWORTH, AND A. SINGER, *Graph laplacian tomography from unknown random projections*, IEEE Trans. Image Process., 17 (2008), pp. 1891–1899, <https://doi.org/10.1109/TIP.2008.2002305>, <https://doi.org/10.1109/TIP.2008.2002305>.
- [11] A. CONCAS, C. FENU, G. RODRIGUEZ, AND R. VANDEBRIL, *The seriation problem in the presence of a double fiedler value*, arXiv preprint arXiv:2204.03362, (2022).
- [12] X. EVANGELOPOULOS, A. J. BROCKMEIER, T. MU, AND J. Y. GOULERMAS, *Circular object arrangement using spherical embeddings*, Pattern Recognit., 103 (2020), p. 107192, <https://doi.org/10.1016/j.patcog.2019.107192>, <https://doi.org/10.1016/j.patcog.2019.107192>.
- [13] C. GIRAUD, Y. ISSARTEL, AND N. VERZELEN, *Localization in 1d non-parametric latent space models from pairwise affinities*, arXiv preprint arXiv:2108.03098, (2021).
- [14] W. HSU AND R. M. MCCONNELL, *PC trees and circular-ones arrangements*, Theor. Comput. Sci., 296 (2003), pp. 99–116, [https://doi.org/10.1016/S0304-3975\(02\)00435-8](https://doi.org/10.1016/S0304-3975(02)00435-8), [https://doi.org/10.1016/S0304-3975\(02\)00435-8](https://doi.org/10.1016/S0304-3975(02)00435-8).
- [15] L. HUBERT, *Some applications of graph theory and related non-metric techniques to problems of approximate seriation: The case of symmetric proximity measures*, British Journal of Mathematical and Statistical Psychology, 27 (1974), pp. 133–153.
- [16] L. HUBERT, P. ARABIE, AND J. MEULMAN, *Linear and circular unidimensional scaling for symmetric proximity matrices*, British Journal of Mathematical and Statistical Psychology, 50 (1997), pp. 253–284.
- [17] 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), pp. 341–358.
- [18] E. V. HUNTINGTON, *A set of independent postulates for cyclic order*, Proceedings of the National Academy of Sciences of the United States of America, 2 (1916), p. 630.
- [19] M. LAURENT AND M. SEMINAROTI, *A lex-bfs-based recognition algorithm for robinsonian matrices*, Discret. Appl. Math., 222 (2017), pp. 151–165, <https://doi.org/10.1016/j.dam.2017.01.027>, <https://doi.org/10.1016/j.dam.2017.01.027>.
- [20] M. LAURENT AND M. SEMINAROTI, *Similarity-first search: A new algorithm with application to robinsonian matrix recognition*, SIAM J. Discret. Math., 31 (2017), pp. 1765–1800, <https://doi.org/10.1137/16M1056791>, <https://doi.org/10.1137/16M1056791>.
- [21] J. LIU, D. LIN, G. G. YARDIMCI, AND W. S. NOBLE, *Unsupervised embedding of single-cell hi-c data*, Bioin-

- form., 34 (2018), pp. i96–i104, <https://doi.org/10.1093/bioinformatics/bty285>, <https://doi.org/10.1093/bioinformatics/bty285>.
- [22] Z. LIU, H. LOU, K. XIE, H. WANG, N. CHEN, 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), pp. 1–9.
- [23] B. G. MIRKIN AND S. N. RODIN, *Graphs and Genes*, Biomathematics, Springer, Berlin, Germany, Nov. 2011.
- [24] P. PRÉA AND D. FORTIN, *An optimal algorithm to recognize robinsonian dissimilarities*, *J. Classif.*, 31 (2014), pp. 351–385, <https://doi.org/10.1007/s00357-014-9150-2>, <https://doi.org/10.1007/s00357-014-9150-2>.
- [25] A. QUILLIOT, *Circular representation problem on hypergraphs*, *Discrete Mathematics*, 51 (1984), pp. 251–264, [https://doi.org/https://doi.org/10.1016/0012-365X\(84\)90006-2](https://doi.org/https://doi.org/10.1016/0012-365X(84)90006-2), <https://www.sciencedirect.com/science/article/pii/0012365X84900062>.
- [26] A. RECANATI, T. KERDREUX, AND A. D’ASPREMONT, *Reconstructing latent orderings by spectral clustering*, CoRR, abs/1807.07122 (2018), <http://arxiv.org/abs/1807.07122>, <https://arxiv.org/abs/1807.07122>.
- [27] W. S. ROBINSON, *A method for chronologically ordering archaeological deposits*, *American antiquity*, 16 (1951), pp. 293–301.
- [28] M. SESTON, *Dissimilarités de Robinson: algorithmes de reconnaissance et d’approximation*, PhD thesis, Aix-Marseille 2, 2008.
- [29] M. SESTON, *A simple algorithm to recognize robinsonian dissimilarities*, 2008. COMPSTAT’2008.