



HAL
open science

Recognition Algorithms for Triadic and Multiway Distances, with an Application to Approximation into a Distance, an Ultrametric or a Robinson Dissimilarity

Pascal Pr ea, Fran ois Brucker

► **To cite this version:**

Pascal Pr ea, Fran ois Brucker. Recognition Algorithms for Triadic and Multiway Distances, with an Application to Approximation into a Distance, an Ultrametric or a Robinson Dissimilarity. Metric Graph Theory and Related Topics, Dec 2020, Marseille, France. hal-03513229

HAL Id: hal-03513229

<https://hal.science/hal-03513229>

Submitted on 5 Jan 2022

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.

Recognition Algorithms for Triadic and Multiway Distances, with an Application to Approximation into a Distance, an Ultrametric or a Robinson Dissimilarity*

Pascal Pr ea Fran ois Brucker

Aix-Marseille Universit , CNRS, Universit  de Toulon, LIS

 cole Centrale Marseille

Marseille, France

{pascal.prea, francois.brucker}@lis-lab.fr

January 5, 2022

Abstract

Given a triadic distance [6, 7] D on a n -set S , we consider the following two recognition problems: (i) determining if D is a *perimeter*, i.e. if there exists a (ordinary) dissimilarity d such that $\forall x, y, z \in S$, $D(x, y, z) = d(x, y) + d(y, z) + d(x, z)$ and (ii) determining if D is a *diameter*, i.e. if there exists a dissimilarity d such that $\forall x, y, z \in S$, $D(x, y, z) = \max\{d(x, y), d(y, z), d(x, z)\}$.

In this note, we give optimal $O(n^3)$ algorithms for these problems. Algorithms for Problem (i) can be extended, with the same complexity, to problems like determining if there exist a dissimilarity d such that (iii) $\forall x, y, z \in S$, $D(x, y, z) = \sqrt[p]{d(x, y)^p + d(y, z)^p + d(x, z)^p}$ or (iv) $\forall x, y, z \in S$, $D(x, y, z) = \sqrt{d(x, y) \cdot d(y, z) \cdot d(x, z)}$. In addition, all these algorithms can be extended to k -way dissimilarities, with complexity $O(n^k)$.

From the algorithm for Problem (ii), we derive algorithms to approximate a dissimilarity into a distance, an ultrametric or a Robinson dissimilarity. The general frame of these algorithms is the following: *if the dissimilarity is closed to a distance (resp. an ultrametric, resp. a Robinson dissimilarity), return a solution to the approximation problem, otherwise return None.*

keywords : triadic distances, k -way dissimilarities, dissimilarities, diameter, perimeter, approximation, ultrametries, Robinson dissimilarities.

1 Introduction

A *dissimilarity* on a finite set S is a symmetric function $d : S \times S \rightarrow \mathbb{R}^+$ such that $d(x, y) = 0$ if $x = y$. Dissimilarities can be seen as a generalization of distances; they are equivalent to symmetric matrices with null diagonal.

Triadic distances or cubes were introduced in [6, 7] as a generalization of dissimilarities. A *triadic distance*, *three-way distance* or (*dissimilarity*) *cube* on a n -set $S = \{x_1, \dots, x_n\}$ is function $D : S \times S \times S \rightarrow \mathbb{R}^+$ such that, if we denote $D(x_i, x_j, x_k)$ by $b_{i,j,k}$:

$$\begin{aligned} b_{ijk} &= b_{ikj} = b_{jik} = b_{jki} = b_{kij} = b_{kji} \\ b_{jji} &= b_{ijj} \end{aligned}$$

*This work was supported in part by ANR project DISTANCIA (ANR-17-CE40-0015).

$$b_{iii} = 0$$

The two first conditions are a generalization to a 3-dimensional array of the symmetry condition for dissimilarities ($d(x, y) = d(y, z)$). The equivalent of rows and columns in a matrix are, for a cube, *rows*, *columns* and *tubes*: for $i, j \in \{1, \dots, n\}$, the (i, j) -th row (resp. column, resp. tube) is the set $\{b_{ijk}, k \in \{1, \dots, n\}\}$ (resp. $\{b_{ikj}, k \in \{1, \dots, n\}\}$, resp. $\{b_{kij}, k \in \{1, \dots, n\}\}$).

In this note, we will only consider values $D(x, y, z)$ with $x \neq y \neq z \neq x$, so the first condition is the only one which will interest us. Similarly, for $k > 3$, we define a k -way dissimilarity as a function $D : S^k \rightarrow \mathbb{R}^+$ such that, for every $\{x_1, x_2, \dots, x_k\} \subset S$ and every permutation σ on $\{1, \dots, k\}$, $D(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(k)}) = D(x_1, x_2, \dots, x_k)$.

Our first goal in this note is to determinate if a triadic distance D can be “linked” with an (ordinary) dissimilarity d . The two most natural links are:

- D is the *diameter* of d , i.e. $\forall x, y, z \in S, D(x, y, z) = \max\{d(x, y), d(y, z), d(x, z)\}$
- D is the *perimeter* of d , i.e. $\forall x, y, z \in S, D(x, y, z) = d(x, y) + d(y, z) + d(x, z)$

In Section 2, we give two $O(n^3)$ algorithms for the Perimeter problem and, in Section 3, an $O(n^3)$ algorithm to solve the Diameter problem. Since the size of the data (a triadic distance on a n -set) is $\approx n^3$, these three algorithms are optimal. These algorithms can be extended, with complexity $O(n^k)$, to k -way dissimilarities. In Section 4, we derive, from the algorithm of Section 3, algorithms to approximate a dissimilarity into a distance, an ultrametric or a Robinson dissimilarity [8].

2 Optimal Algorithms for the Perimeter Problem

2.1 An analytic algorithm

In this section, we will see that if a triadic distance D is the perimeter of a dissimilarity d , then d can be analytically expressed from D .

Given a triadic distance D on a n -set S , perimeter of a dissimilarity d . We define:

$$D(\cdot, \cdot, \cdot) := \sum_{\{x, y, z\} \subset S} D(x, y, z) = \sum_{\{x, y, z\} \subset S} d(x, y) + d(y, z) + d(x, z) \quad (1)$$

$\forall x \in S$:

$$D(x, \cdot, \cdot) := \sum_{\{y, z\} \subset S \setminus \{x\}} D(x, y, z) = \sum_{\{y, z\} \subset S \setminus \{x\}} d(x, y) + d(y, z) + d(x, z) \quad (2)$$

$\forall x, y \in S$:

$$D(x, y, \cdot) := \sum_{z \in S \setminus \{x, y\}} D(x, y, z) = \sum_{z \in S \setminus \{x, y\}} d(x, y) + d(y, z) + d(x, z) \quad (3)$$

Similarly:

$$d(\cdot, \cdot) := \sum_{\{x, y\} \subset S} d(x, y) \quad (4)$$

$\forall x \in S$:

$$d(x, \cdot) := \sum_{y \in S \setminus \{x\}} d(x, y) \quad (5)$$

In equation (1), each $d(x, y)$ appears $n - 2$ times (one for each $z \neq x, y$), so:

$$D(\cdot, \cdot, \cdot) = (n - 2) \cdot d(\cdot, \cdot) \quad (6)$$

In Equation 2, each $d(x, y)$ appears $n - 2$ times. So:

$$D(x, \cdot, \cdot) = (n - 2) \cdot d(x, \cdot) + \sum_{y, z \neq x} d(y, z) \quad (7)$$

As $\sum_{y, z \neq x} d(y, z) = d(\cdot, \cdot) - d(x, \cdot)$, we have:

$$d(x, \cdot) = \frac{D(x, \cdot, \cdot) - d(\cdot, \cdot)}{n - 3} \quad (8)$$

In Equation 3, $d(x, y)$ appears $n - 2$ times. In addition, $\sum_{z \neq x, y} d(x, z) = d(x, \cdot) - d(x, y)$. So, we have:

$$d(x, y) = \frac{D(x, y, \cdot) - d(x, \cdot) - d(y, \cdot)}{n - 4} \quad (9)$$

If D is the perimeter of a dissimilarity d , Equations (6-8-9) yield an analytic expression of d as a function of D (Equation 10) which yields to an $O(n^3)$ Algorithm.

$$d(x, y) = \frac{D(x, y, \cdot) - \frac{D(x, \cdot, \cdot) + D(y, \cdot, \cdot) - 2\frac{D(\cdot, \cdot, \cdot)}{n-2}}{n-3}}{n - 4} \quad (10)$$

2.2 A recursive algorithm

We now present another algorithm for this problem, based on the fact that the perimeter problem can be settled as a system of $n(n - 1)(n - 2)/6$ linear equations (each equation corresponds to a value $D(x, y, z)$) on $n(n - 1)/2$ unknowns (each unknown corresponds with a value $d(x, y)$), namely all the equations

$$d(x, y) + d(y, z) + d(x, z) = D(x, y, z) \quad (11)$$

Solving this system would not yield to an efficient algorithm, but it can help to design one.

For $S' \subset S$, we define $\mathfrak{E}(S')$ as the system of all equations (11) built on the elements of S' . If $|S'| = 5$, $\mathfrak{E}(S')$ is a system of 10 equations on 10 unknowns; if $|S'| < 5$, $\mathfrak{E}(S')$ has more unknowns than equations and if $|S'| > 5$, $\mathfrak{E}(S')$ has more equations than unknowns.

The following Algorithm PERIMETER_INVERSION first partition S into $n/5$ sets of 5 points and solves \mathfrak{E} on each of these subsets. Then it merges these subsets. The values of D are not all used to compute d , so there is a final step of verification (Function IS_PERIMETER).

Algorithm 1: PERIMETER_INVERSION

Input: A 3-way dissimilarity D on a n -set $S = \{x_1, x_2, \dots, x_n\}$

Output: A dissimilarity d on S such that D is the perimeter of d , or **None** if such a dissimilarity does not exist

begin

forall $x_i \in S$ **do**

$d(x_i, x_i) \leftarrow 0$;

$\chi_i \leftarrow \{x_i\}$;

$\Theta \leftarrow \{\chi_1, \chi_2, \dots, \chi_n\}$;

while $|\Theta| > 1$ **do**

$(\Theta, d) \leftarrow \text{ZONE_FUSION}(S, D, \Theta, d)$;

if IS_PERIMETER(d, D) **then**

return d

It is possible, in Algorithm ZONE_FUSION, that $|I_{\lceil k/5 \rceil}| = a < 5$; In this case, in order to have a system $\mathfrak{E}(\{p_i : i \in I_j\})$ with as many unknowns that equations, we complete $\{p_i\}$ with $5 - a$ other points.

Algorithm 2: ZONE_FUSION

Input: A 3-way dissimilarity D on a n -set S
A partition $\Theta = \{\theta_1, \theta_2, \dots, \theta_k\}$ of S
A dissimilarity d on S such that $\forall \theta_i \in \Theta, \forall x, y \in \theta_i, d(x, y)$ is known
(the other values are yet undetermined)

Output: A partition $\Omega = \{\omega_1, \dots, \omega_{\lceil k/5 \rceil}\}$ of S // The dissimilarity d is now known $\forall x, y \in \omega_i,$
 $\forall \omega_i \in \Omega$

begin

- forall** $\theta_i \in \Theta$ **do**
 - └ Chose $p_i \in \theta_i$;
 - forall** $j \in \{1, \dots, \lceil k/5 \rceil\}$ **do**
 - └ $I_j \leftarrow \{j, j+1, j+2, j+3, j+4\}$;
 - └ $\omega_j \leftarrow \cup_{i \in I_j} \theta_i$;
 - └ Solve $\mathfrak{S}(\{p_i : i \in I_j\})$;
 - └ **forall** $\{l, l'\} \subset I_j$ **do**
 - └ **forall** $x \in \theta_l$ **do**
 - └ **if** $x \neq p_l$ **then**
 - └ $d(x, p_{l'}) \leftarrow D(x, p_l, p_{l'}) - d(x, p_l) - d(p_l, p_{l'})$;
 - └ **forall** $y \in \theta_{l'} \setminus \{p_{l'}\}$ **do**
 - └ $d(x, y) \leftarrow D(x, y, p_{l'}) - d(x, p_{l'}) - d(y, p_{l'})$;
- └ **return** $(\Omega := \{\omega_1, \omega_2, \dots, \omega_{\lceil k/5 \rceil}\}, d)$;

Property 1. Algorithm PERIMETER_INVERSION solves the perimeter problem.

Property 2. Algorithm PERIMETER_INVERSION runs in $O(n^3)$.

Proof. Each resolution of $\mathfrak{S}(\{p_i : i \in I_j\})$ takes a constant time. There is less than $\sum_{i \geq 1} n/5^i$ such systems. So, the resolution of all systems $\mathfrak{S}(\{p_i : i \in I_j\})$ takes $O(n)$ time.

Each value which is not determined by a system $\mathfrak{S}(\{p_i : i \in I_j\})$ takes $O(1)$ time to be computed. So the computation of all these values takes $O(n^2)$ time.

Finally, the verification runs in $O(n^3)$. □

We remark that the $O(n^3)$ complexity of Algorithm 2 is entirely due to the verification step. The construction of d takes only $O(n^2)$.

2.3 Extensions

Given two functions $f, g: \mathbb{R} \rightarrow \mathbb{R}$, such that f^{-1} exists, we say that a triadic distance D is a *fg-perimeter* of a dissimilarity d if $g(D(x, y, z)) = f(d(x, y)) + f(d(x, y)) + f(d(x, y))$. Algorithm PERIMETER_INVERSION can be adapted to determine, with the same $O(n^3)$ complexity, if D is a *fg-perimeter* of a dissimilarity d . This extension of the Perimeter problem covers many links between triadic distance and dissimilarities. For instance:

- The harmonic mean $D(x, y, z) = \sqrt{d(x, y) \cdot d(y, z) \cdot d(x, z)}$, with $f(a) = \log(a)$ and $g(a) = \log^2(a)$.
- The L_p -norm $D(x, y, z) = \sqrt[p]{d(x, y)^p + d(y, z)^p + d(x, z)^p}$ with $f(a) = g(a) = a^p$.

In addition, all these algorithms can be extended to k -way dissimilarities with complexity $O(n^k)$.

3 The Diameter Problem

In this section, we first give an optimal algorithm for the Diameter problem.

Algorithm 3: DIAMETER_INVERSION

Input: A 3-way dissimilarity D on a n -set S

Output: A dissimilarity d on S such that D is the diameter of d , or **None** if such a dissimilarity does not exist

begin

forall $x, y \in S$ **do**

$d(x, y) \leftarrow \infty$;

forall $x, y, z \in S$ **do**

$d(x, y) \leftarrow \min(d(x, y), D(x, y, z))$;

$d(x, z) \leftarrow \min(d(x, z), D(x, y, z))$;

$d(z, y) \leftarrow \min(d(z, y), D(x, y, z))$;

forall $x, y, z \in S$ **do**

if $\max(d(x, y), d(y, z), d(x, z)) \neq D(x, y, z)$ **then**

return None ;

return d ;

Property 3. Algorithm DIAMETER_INVERSION runs in $O(n^3)$.

Property 4. The output of Algorithm DIAMETER_INVERSION is **None** if and only if D is not the diameter of a dissimilarity.

Different dissimilarities may have the same diameter, even when structurally different. For instance, the dissimilarities d , defined by $\forall x, y \in S, d(x, y) = 3$, and d' , defined by $\forall x, y \in S, d'(x, y) = 3$ except $d'(x_1, x_2) = d'(x_1, x_3) = 1$, have the same diameter although d is an ultrametric ($\forall x, y, z, d(x, y) \leq \max(d(x, z), d(y, z))$) and d' is not a distance.

Given a dissimilarity d , we define the 3-smoothing of d by $d^{[3]} = \text{DIAMETER_INVERSION}(\text{DIAMETER}(d))$, i.e.:

$$\forall x, y \in S, d^{[3]}(x, y) := \min_{z \neq x, y} \max(d(x, y), d(y, z), d(x, z))$$

As for the Perimeter problem, Algorithm 3 can be generalized to k -way dissimilarities with complexity $O(n^k)$ and similarly, we define, for $k \in \{2, \dots, n\}$ the k -smoothing $d^{[k]}$ by:

$$\forall x, y \in S, d^{[k]}(x, y) := \min_{\substack{\{x, y\} \subset K \subset S \\ |K|=k}} \max\{d(a, b) : a, b \in K\}$$

Note that $d^{[2]} = d$ and that $d^{[n]}$ is a constant.

Claim 1. Let d be a dissimilarity on S and $k < k' \in \{2, \dots, n\}$, then, $\forall x, y \in S, d^{[k]}(x, y) \leq d^{[k']}(x, y)$.

Property 5. Let d be a dissimilarity on S and $k \in \{2, \dots, n-1\}$ such that $d^{[k]}$ is a distance. Then, for $k' \in \{k+1, \dots, n\}$, $d^{[k']}$ is a distance.

Proof. Let x, y, z be three elements of S , and K be a set of cardinality $k+1$, containing x and z and such that its diameter is $d^{[k+1]}(x, z)$. For all $t \in K, d(y, t) \leq d^{[k]}(y, t) \leq d^{[k]}(y, z) + d^{[k]}(z, t) \leq d^{[k+1]}(y, z) + d^{[k+1]}(x, z)$. The set $K \cup \{y\}$ contains x and y , its cardinality is $\geq k+1$ and its diameter is $\leq d^{[k+1]}(y, z) + d^{[k+1]}(x, z)$. Thus $d^{[k+1]}(x, y) \leq d^{[k+1]}(y, z) + d^{[k+1]}(x, z)$ and so $d^{[k+1]}$ is a distance. \square

A dissimilarity d on S is *Robinson* [8] if there exists a order (which is said to be *compatible*) σ on S such that, when S is sorted along σ , if $x < y < z$, then $d(x, y) \leq d(x, z)$ and $d(y, z) \leq d(x, z)$.

Property 6. Let d be a dissimilarity on S , σ a permutation of S and $k \in \{2, \dots, n-1\}$ such that $d^{[k]}$ is Robinson with σ as a compatible order. Then, for $k' \in \{k+1, \dots, n\}$, $d^{[k']}$ is Robinson and admits σ as a compatible order.

Proof. We suppose that S is sorted along σ , and let $x <_{\sigma} y <_{\sigma} z$ be three elements of S . Let K be a set of cardinality $k+1$, containing x and z and of diameter $d^{[k+1]}(x, z)$. For any $t \in S$, we have $d(y, t) \leq d^{[k]}(y, t) \leq d^{[k]}(x, t)$ (if $t \geq_{\sigma} y$) or $d(y, t) \leq d^{[k]}(z, t)$ (if $t \leq_{\sigma} y$). So the diameter of $K \cup \{y\}$ is $\leq d^{[k+1]}(x, z)$ and thus we have $d^{[k+1]}(x, y) \leq d^{[k+1]}(x, z)$ and $d^{[k+1]}(y, z) \leq d^{[k+1]}(x, z)$. \square

Property 7. Let d be a dissimilarity and $k \in \{2, \dots, n-1\}$ such that $d^{[k]}$ is an ultrametric. Then $d^{[k']}$ is an ultrametric for every $k' \in \{k+1, \dots, n\}$.

Proof. Let $x, y, z \in S$; we suppose, with no loss of generality, that $d^{[k]}(x, y) \leq d^{[k]}(x, z) = d^{[k]}(y, z)$. The dissimilarity $d^{[k]}$ is Robinson [5]. Moreover, $d^{[k]}$ (as a Robinson dissimilarity) admits two compatible orders σ_1 and σ_2 such that $x <_{\sigma_1} y <_{\sigma_1} z$ and $y <_{\sigma_2} x <_{\sigma_2} z$. By Property 6, $d^{[k+1]}$ is Robinson and admits σ_1 and σ_2 as compatible orders. So, $\max(d^{[k+1]}(x, y), d^{[k+1]}(y, z)) \leq d^{[k+1]}(x, z)$ and $\max(d^{[k+1]}(x, y), d^{[k+1]}(x, z)) \leq d^{[k+1]}(y, z)$, i.e. $d^{[k+1]}(x, y) \leq d^{[k+1]}(x, z) = d^{[k+1]}(y, z)$. \square

Properties as 5, 6, 7 do not exist for any type of distances. For instance, there is no such properties for tree-distances¹ or quasi-ultrametrics², as shown in the example of Figure 1.

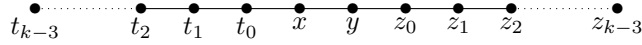


Figure 1: A dissimilarity d which is a tree-distance and a quasi-ultrametric (for all a, b , $B_d(a, b)$ is the interval $[a, b]$), although $d^{[k]}$ is neither a tree-distance nor a quasi-ultrametric: $d^{[k]}(x, y) = d^{[k]}(x, t_{k-3}) = d^{[k]}(y, t_{k-3}) = d^{[k]}(x, z_{k-3}) = d^{[k]}(y, z_{k-3}) = k-1$ and $d^{[k]}(z_{k-3}, t_{k-3}) = 2k-3$.

4 Approximation algorithms

The k -smoothing yields to a natural way to approximate a dissimilarity into a distance, an ultrametric or a Robinson dissimilarity is the following Algorithm 4, where the “desired property” can be a distance, an ultrametric, a Robinson dissimilarity, or any property which is verified by the constant dissimilarity.

Algorithm 4 may have an exponential complexity. Its variant Algorithm 5 is of complexity $O(n^3)$.

Algorithm 5 gives a solution only if the 3-smoothing of the input has the desired property. We can interpret this as “the input is closed to have the desired property”. The interest of such an algorithm is that it has a good complexity and, if the input d is “far” from having the property, it is not pertinent to approximate it.

The problem is that this notion of “far” actually means “far for the algorithm”. Does this notion correspond with a natural one? We have tested that for Robinson dissimilarities. We have generated points in a $p \times q$ rectangle and taken for d the distance between points. If $p = 0$, then d is Robinson. So, if $p \ll q$,

¹A dissimilarity d on S is a *tree-distance* [1] if there exists a valuated tree $T = (V, E)$ such that $S \subset V$ and for all $x, y \in S$, $d(x, y)$ is the length of the shortest path of T between x and y . Equivalently, for all $x, y, z, t \in S$ the two greatest of $d(x, y) + d(z, t)$, $d(x, z) + d(y, t)$ and $d(x, t) + d(y, z)$ are equal.

²Let d be a dissimilarity on S . For $x, y \in S$, the *two-ball* $B_d(x, y)$ is the set $\{z \in S : d(z, x), d(z, y) \leq d(x, y)\}$. The dissimilarity d is a *quasi-ultrametric* [3, 4] if every two-ball of d is a clique, i.e. if:

$$\forall x, y, z, t \in S, d(z, x), d(z, y), d(t, x), d(t, y) \leq d(x, y) \implies d(z, t) \leq d(x, y)$$

Algorithm 4: NAÏVE_APPROXIMATION

Input: A dissimilarity d on a n -set S **Output:** A dissimilarity d' on S which has the desired property**begin** $d' \leftarrow d$; $k \leftarrow 2$; **while** d' has not the desired property **do** $k \leftarrow k + 1$; $d' \leftarrow d^{[k]}$; **return** d' ;

Algorithm 5: PARTIAL_APPROXIMATION

Input: A dissimilarity d on a n -set S **Output:** A dissimilarity d' on S which has the desired property, or **None** if d is far from having this property**begin** **if** d has the desired property **then** **return** d ; **if** $d^{[3]}$ has the desired property **then** **return** $d^{[3]}$; **return** **None** ;

d is close to a Robinson dissimilarity and if $p = q$, d is “absolutely not” Robinson. Results of this test are shown in Figure 2.

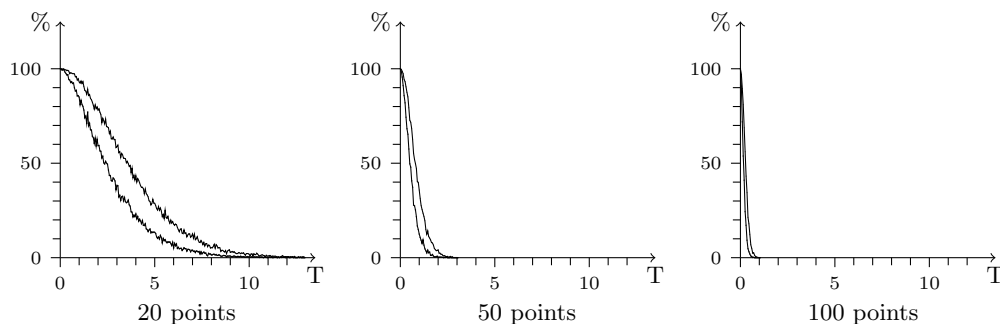


Figure 2: Results of tests for Robinson: sets of points (of size 20, 50 or 100) with integer coordinates have been generated into rectangles of length 10^8 and we considered the Euclidean distance between these points. On the abscissas, we have indicated the proportion of the rectangle (“5” indicates that the breadth of the rectangle is 5% of its length); on the ordinate, we have indicated the percent of Robinson dissimilarities. The lower curve is for dissimilarities before 3-smoothing, and the upper one for dissimilarities after 3-smoothing. Each point on the curves correspond with the average value on 1000 trials. For instance, on 1000 sets of 20 points generated in a rectangle of size $10^8 \times 5 \cdot 10^6$, the Euclidean distance between these points is Robinson for 134 sets before 3-smoothing and for 288 sets after 3-smoothing. The time taken by our algorithm on a laptop is 1.7 ms for a 20×20 matrix, 25 ms for a 50×50 matrix and 200 ms for a 100×100 matrix.

As expected, we can see the the proportion of Robinson dissimilarities after 3-smoothing decreases with the breadth of the rectangles.

References

- [1] J.P. BARTHÉLEMY & A. GUÉNOCHE (1991), *Trees and Proximity Representations*. J. Wiley & sons.
- [2] V. CHEPOI & B. FICHET (2007) A note on three-way dissimilarities and their relationship with two-way dissimilarities, in *Selected Contributions in Data Analysis and Classification*, P. Brito, P. Bertrand, G. Cucumel & F. de Carvalho Eds, pp 465-475, Springer.
- [3] J. DIATTA & B. FICHET (1994), From Apresjan Hierarcjies and Bandelt-Dress Weak Hirerachies to Quasi-Hierarchies, in *New Approaches in Classification and Data Analysis*, E. Diday, Y. Lechevallier, M. Schader, P. Bertrand & B. Burtschy Ed., pp 111-118. Springer-Verlag.
- [4] J. DIATTA & B. FICHET (1998), Quasi-ultrametrics and their 2-ball hypergraphs, *Discrete Math.* 192, 87-102.
- [5] C. DURAND & B. FICHET (1988), One-to-One Correspondences in Pyramidal Representation: an Unified Approach, in *Classification and Related Methods of Data Analysis*, H.H. Bock Ed., pp 85–90. North-Holland.
- [6] W.J. HEISER & M. BENNANI (1997), Triadic distance models: axiomatization and least squares representation, *Journal of Mathematical Psychology* 41, 189-206.
- [7] S. JOLY & G. LE CALVÉ (1995), Three way distances, *J. Classif.* 12, 191-205.
- [8] W.S. ROBINSON (1951), A Method for Chronologically Ordering Archeological Deposits, *American Antiquity* 16, 293-301.