



**HAL**  
open science

# Convergence rate for the $\lambda$ -Medial-Axis estimation under regularity conditions

Catherine Aaron

► **To cite this version:**

Catherine Aaron. Convergence rate for the  $\lambda$ -Medial-Axis estimation under regularity conditions. 2018. hal-01558392v2

**HAL Id: hal-01558392**

**<https://hal.science/hal-01558392v2>**

Preprint submitted on 11 Jul 2018 (v2), last revised 4 Jul 2019 (v3)

**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 destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Convergence rate for the $\lambda$ -Medial-Axis estimation under regularity conditions

C. Aaron

June 26, 2018

## Abstract

Let  $\mathcal{X} = \{X_1, \dots, X_n\} \subset \mathbb{R}^d$  be a random sample of observations drawn with a probability distribution supported on  $S$  satisfying that both  $S$  and  $\bar{S}^c$  are  $r_0$ -convex ( $r_0 > 0$ ). In this paper we propose an estimator of the medial axis of  $S$  based on the  $\lambda$ -medial axis and the  $r$ -convex hull. Its convergence rate is derived. An heuristic to tune the parameters of the estimator is given and a small simulation study is performed.

**Keywords:** Manifold Learning, Geometric Inference, Skeleton.

## 1 Introduction

The medial axis of a set  $S$  was introduced in [5] as the set of points that has at least two different projections on  $\partial S$ . It was proposed as a tool for biological shape recognition and became popular in image compression and image analysis (see [11] for instance). More recently it has also been applied to wireless networks [20].

In this paper we will focus on the inner medial axis, that is, the intersection of the medial axis and  $S$  which carries all the information for set estimation and image analysis.

More precisely we are interested in  $\mathcal{M}(S) = \{x \in S, \text{diam}(\Gamma(x)) > 0\}$  where  $\Gamma(x) = \{y \in \partial S, \|y - x\| = d(x, \partial S)\}$ .

There exists another object, the skeleton, that can be defined, following [21] or [15], by the set of the centers of the maximal balls included in  $S$ . More precisely, if  $B(x, r)$  denotes the open ball centered in  $x$  and of radius  $r$  and if  $\mathring{S}$  denotes the interior of  $S$  we consider:

$$\mathcal{M}^*(S) = \{x, \exists r(x) \text{ such that } B(x, r(x)) \subset \mathring{S} \text{ and } B(x, r(x)) \subsetneq B(x', r') \Rightarrow B(x', r') \not\subset \mathring{S}\}. \quad (1)$$

It can be proved (see [15]) that  $\mathcal{M}(S) \subset \mathcal{M}^*(S) \subset \overline{\mathcal{M}(S)}$  and an example where the last inclusion is strict can be found in [8]. Nevertheless in this work we will assume that  $\mathcal{M}(S)$  is closed which directly implies that  $\mathcal{M}(S) = \mathcal{M}^*(S)$ .

Let  $\mathcal{X}_n = \{X_1, \dots, X_n\}$  be a random sample of observations drawn with a probability distribution supported on an unknown support  $S$ . Being able to estimate  $\mathcal{M}(S)$  has various statistical applications. First of all, the knowledge of  $(x, d(x, \partial S))$  for all  $x \in \mathcal{M}(S)$ , known as the “medial axis transformation”, allows one to reconstruct  $S$  (see [21]). Being able to estimate  $\mathcal{M}(S)$  and  $d(x, \partial S)$  provides a natural plug-in **estimator of the support**.

In [12] the author proposed to apply the medial axis to **filament estimation** (one dimensional curves embedded in  $\mathbb{R}^d$ ). The same kind of idea can be applied in any dimension for **manifold estimation**. Suppose that we observe the random variable  $X$  that is drawn “close to” a  $d'$ -dimensional manifold i.e.  $X = Y + rU$  where  $Y$  is a hidden random variable whose

distribution is supported by  $M$ , a  $d'$ -dimensional manifold embedded in  $\mathbb{R}^d$ ,  $r$  is a positive real number, and  $U$  is a random variable uniformly drawn on the  $d$ -dimensional unit ball. Then the support  $S$  of the random variable  $X$  is  $M \oplus rB = \{x, \exists y \in M, \|x - y\| \leq r\}$ . Under some regularity assumptions (namely that  $M$  has a positive reach  $r_0 > r$ ) we have  $M = \mathcal{M}(M \oplus rB)$  and medial axis estimation based on the observations of  $X$  allows us to estimate the underlying unknown lower dimensional structure. Note that the minimax bounds for the estimation of  $M$  has been obtained in [13], the estimator proposed in Section 3.1. is minimax when the co-dimension is 1 (when  $d' = d - 1$ ) and not optimal when  $d' < d - 1$  (this seems reasonable, as noticed in [13], because we should use the knowledge of  $d'$  to build a minimax estimator). More generally suppose that  $X = Y + r(Y)U$  where the distribution of the unknown  $Y$  is supported by a  $d'$ -dimensional sub-manifold of  $\mathbb{R}^d$  and  $U$  is a random variable uniformly drawn on the  $d$ -dimensional unit ball. Manifold estimation consists of estimating  $M$  by observing  $X$ . Let  $S = \{x, \exists y \in M, \|x - y\| \leq r(y)\}$  be the support of  $X$ . In a more general setup, since  $M = \mathcal{M}(S)$  (which can be interpreted as an identifiability condition for the estimation of  $M$ ) manifold estimation and medial axis estimation amount to the same problem. Remark that  $M = \mathcal{M}(S)$  allows non-constant radius  $r$  if the variations are slow enough. Also remark that, in this context, the assumption that  $\mathcal{M}(S)$  is closed is naturally satisfied. In Section 3.4 some simulations are performed for  $d = 2$  and  $d = 3$ . The three dimensional case illustrates the use of the medial axis as a manifold-estimator.

Unfortunately, the medial axis is not continuous with respect to the Hausdorff distance  $d_h$  (see pages 217–238 of [15] for topological properties of the medial axis). This implies that estimating the medial axis using a finite sample of points  $\mathcal{X}_n = \{X_1, \dots, X_n\}$  is unstable and provides a challenging problem that has been investigated in various papers (see [3] for a state-of-the-art report).

Two different approaches have been investigated. The first one consists of pruning the medial axis of an estimation of  $S$  (see [18], [6], [9],[4] or [16]); the second one consists of estimating the  $\lambda$ -medial axis instead of the medial axis  $\mathcal{M}_\lambda(S) = \{x \in \mathcal{M}(S), \Gamma(x) \subset \mathcal{B}(a, r) \Rightarrow r \geq \lambda\}$ , the latter object being stable with respect to the Hausdorff distance (see [7]). More precisely the Authors prove that, if  $d_h(S^{rc}, S^c) \leq \varepsilon$  then  $d_h(\mathcal{M}_\lambda(S), \mathcal{M}_\lambda(S')) \leq C_S \sqrt{\varepsilon}$ , then they propose an algorithm to estimate  $\hat{\mathcal{M}}(S)$  given sample points located near the boundary, that is, given  $E_n$  a sequence of finite set of points such  $d_h(E_n, \partial S) \rightarrow 0$  they propose a function  $\hat{\mathcal{M}}_\lambda(E_n)$  and prove that  $d_h(\hat{\mathcal{M}}_\lambda(E_n), \mathcal{M}_\lambda(S)) \rightarrow 0$ .

Later, given a sample point  $\mathcal{X}_n$  drawn on  $S$  (instead of “near  $\partial S$ ”), it is proved in [10], under no more shape hypothesis than regularity, that given a support estimator  $\hat{S}_n$  such that  $d_h(\hat{S}_n, S) \rightarrow 0$  a.s. and  $d_h(\partial \hat{S}_n, \partial S) \rightarrow 0$  a.s. then  $d_h(\mathcal{M}_\lambda(\hat{S}_n), \mathcal{M}_\lambda(S)) \rightarrow 0$  a.s.

We are going to introduce a new medial axis estimator  $\hat{\mathcal{M}}_\lambda(\mathcal{X}_n)$  that is morally very close to the one introduced in [7]: we just restrict our medial axis to be the inner part of the medial axis and we consider the diameter of  $\Gamma(x)$  instead of the radius of the smallest ball containing  $\Gamma(x)$  (which is a bit more quick to compute).

Namely suppose that one can extract from  $\mathcal{X}_n$  a subset  $\mathcal{Y} = \{Y_1, \dots, Y_N\} \subset \mathcal{X}_n$  of observations “close” to  $\partial S$  and that we can estimate  $S$  by  $\hat{S}_n$ . We propose the following medial axis estimator:

$$\hat{\mathcal{M}}_{\lambda, S_n}(\mathcal{Y}) = \left\{ x \in \text{Vor}_{\mathcal{Y}}(y) \cap \text{Vor}_{\mathcal{Y}}(z) \cap \hat{S}_n, (y, z) \in \mathcal{Y}^2, \|y - z\| \geq \lambda \right\}, \quad (2)$$

where  $\text{Vor}_{\mathcal{Y}}(y) = \{z, \|z - y\| \leq d(z, \mathcal{Y})\}$  is the Voronoi cell of  $y$  with respect to the set  $\mathcal{Y}$ .

In this paper we aim to prove that, under some shape hypothesis on  $S$ , there exists suitable values of  $\lambda$  such that, if  $d_h(\mathcal{Y}_n, \partial S) \xrightarrow{a.s.} 0$ ,  $d_h(\hat{S}_n, S) \xrightarrow{a.s.} 0$  and  $d_h(\partial \hat{S}_n, \partial S) \xrightarrow{a.s.} 0$  then  $d_h(\hat{\mathcal{M}}_\lambda(\mathcal{X}_n), \mathcal{M}(S)) \xrightarrow{a.s.} 0$  almost surely (convergence rates are also given). This means that, under shape hypotheses, the  $\lambda$  can be seen as a pruning parameter for the estimation of  $\mathcal{M}(S)$ .

This improves the results given in [10] in which the authors proved that we can estimate  $\mathcal{M}_\lambda(S)$  without giving convergence rates.

Let  $\mathcal{C}_r(\mathcal{X}_n)$  be the  $r$ -convex hull of the observations (see [17]). For some  $r < r_0$ , we will prove the choice of  $\mathcal{Y} = \partial\mathcal{C}_r(\mathcal{X}_n) \cap \mathcal{X}_n$  and  $S_n = C_r(\mathcal{X}_n)$  allows us to obtain  $d_h(\hat{\mathcal{M}}_{\lambda, \hat{S}_n}(\mathcal{Y}), \mathcal{M}(S)) \leq (\ln n/n)^{2/(d+1)}$  for suitable values of  $\lambda$ .

Finally we propose an empirical way to choose the tuning parameters, and we perform a small simulation study.

## 2 Shape Hypothesis and main result

Throughout the paper we consider subsets of  $\mathbb{R}^d$ . Let us start with some definitions and notation. Let  $A$  be an open set and  $C$  a closed set such that  $A \cap C = \emptyset$  and introduce  $\mathcal{B}(A, C) = \{B(x, r), x \in A, B(x, r) \cap C = \emptyset\}$ ,  $\mathcal{C}^{\max}(A, C)$  the set of the maximal balls (for the inclusion) of  $\mathcal{B}(A, C)$  and  $\overline{\mathcal{M}}(A, C)$  the set of the centers of the balls of  $\mathcal{B}(A, C)$ . In the following,  $\mathcal{S}(x, r)$  denotes the sphere of radius  $r$  centered at  $x$ .

Let  $S$  be a compact set. As defined in (1) we have  $\mathcal{M}^*(S) = \overline{\mathcal{M}}(\overset{\circ}{S}, \partial S)$ . For any point  $x \in \mathcal{M}^*(S)$  let us introduce  $r(x) = d(x, \partial S)$ , i.e.  $B(x, r(x))$  is a maximal ball of  $\mathcal{B}(\overset{\circ}{S}, \partial S)$ . As noted in [15], the medial axis may have various ‘‘bad’’ properties, and we are going to restrict our study to regular enough sets, namely we will require that balls of radius  $r_0$  roll freely inside and outside  $S$ , that  $\mathcal{M}(S)$  is closed and that the radius  $r(x) = d(x, \partial S)$  for  $x \in \mathcal{M}(S)$  is a  $K$ -Lipschitz continuous function for some  $K < 1$ .

**Definition 1.** *Balls of radius  $r_0$  roll freely outside and inside  $S$  if, for each  $x \in \partial S$  there exists  $O_x^{\text{out}}$  and  $O_x^{\text{in}}$  such that  $B(O_x^{\text{out}}, r_0) \subset S^c$  and  $B(O_x^{\text{in}}, r_0) \subset S$ . In this case, we introduce  $u_x = \frac{O_x^{\text{out}} - x}{r_0}$  the unit vector, normal to  $\partial S$  and pointing outward from  $S$ .*

According to Walther, it is equivalent to have balls of radius  $r_0$  rolling freely inside and outside  $S$  and the  $r_0$ -convexity of  $S$ ,  $\overline{S^c}$  and  $\overset{\circ}{S} \neq \emptyset$  as cited in the introduction. Nevertheless the ball vision of such a notion is the most helpful for the proofs. If  $S$  is a compact set such that balls roll freely inside and outside  $S$  then it is regular enough to have a medial axis satisfying some good properties described in the following lemma.

**Proposition 1.** *If  $S$  is a compact set such that balls of radius  $r_0$  roll freely inside and outside  $S$  then  $S = \overline{(\overset{\circ}{S})}$  (one then says  $S$  is regular which is a common condition when considering the medial axis);*

*Proof.* The inclusion  $\overline{(\overset{\circ}{S})} \subset S$  comes from the closeness of  $S$ . Considering the second inclusion  $S \subset \overline{(\overset{\circ}{S})}$ , for any  $x \in S$ , on one hand if  $x \in \overset{\circ}{S}$  then  $x \in \overline{(\overset{\circ}{S})}$ ; on the other hand if  $x \in \partial S$ , introduce  $x_n = x - \frac{r_0}{2n} u_x$ . The rolling ball property implies that  $x_n \in \overset{\circ}{S}$  and as  $x_n \rightarrow x$  we have  $x \in \overline{(\overset{\circ}{S})}$ .  $\square$

In addition to the regularity of the support we need additional assumptions, summarized in Definition 2

**Definition 2.** *Let  $r_0 > 0$  and  $K < 1$  be two numbers,  $S$  be a compact set in  $\mathbb{R}^d$ . We say  $S$  is  $(K, r_0)$ -regular if:*

1. balls of radius  $r_0$  roll freely inside and outside  $S$ ;
2.  $\mathcal{M}(S)$  is closed;

3. for all  $(x, y) \in \mathcal{M}(S)^2$ ,  $|d(x, \partial S) - d(y, \partial S)| / \|x - y\| \leq K$

The second assumption, that  $\mathcal{M}(S)$  is closed, ensures that medial axis and skeleton are the same object. The third assumption is a bit more technical and is required in the proof. Notice that it is not so restrictive since the maximality of the balls  $B(x, d(x, \partial S))$  and  $B(y, d(y, \partial S))$  and the triangular inequality imply that  $|d(x, \partial S) - d(y, \partial S)| \leq \|x - y\|$ .

Now recall that we aim to estimate the medial axis of a set  $S$  via an estimation based on a finite number of points. It will be seen later that, regarding the medial axis, the two parts of the Hausdorff distance don't have the same importance. This leads us to define  $(\varepsilon, h)$ -estimations as follows.

**Definition 3.** Let  $S$  and  $\tilde{S}$  be two sets in  $\mathbb{R}^d$ . Then  $\tilde{S}$  is an  $(\varepsilon, h)$ -estimation of  $S$  if:

$$\max_{y \in \tilde{S}} d(y, S) \leq \varepsilon \text{ and } \max_{x \in S} d(x, \tilde{S}) \leq h.$$

The behavior of the estimated medial axis is made explicit in the following theorem. Note that the smoothness conditions on the boundary are close to the one used in [2] where a similar theorem is obtained for a sufficiently dense sample of the boundary. Nevertheless, here we can just observe points close to the boundary and we avoid the heavy computation of a weighted Delaunay tessellation.

**Theorem 1.** Let  $S$  be a  $(K, r_0)$ -regular compact set. Suppose that there exists  $\varepsilon_n \rightarrow 0$ ,  $\mathcal{Y}(\mathcal{X}_n) \subset \mathcal{X}_n$  and  $\hat{S}_n \subset S$  such that

1.  $\mathcal{Y}(\mathcal{X}_n) \subset \mathcal{X}_n$  is a  $(\varepsilon_n^2, \varepsilon_n)$ -estimation of  $\partial S$
2.  $d_h(\hat{S}_n, S) \leq \varepsilon_n^2$  and  $d_h(\partial \hat{S}_n, \partial S) \leq \varepsilon_n^2$

There exists an explicit constant  $\lambda_0$  and a constant  $C$  such that for all  $\lambda < \lambda_0$ , for  $n$  large enough:

$$d_h(\hat{\mathcal{M}}_{\lambda, \hat{S}_n}(\mathcal{Y}), \mathcal{M}) \leq C\varepsilon_n^2$$

*Proof.* First note that the inner rolling ball condition implies that for all  $x \in \mathcal{M}$ ,  $r(x) = d(x, \partial S) \geq r_0$ .

For any  $x \in \mathcal{M}(S)$  we have that  $x \in \hat{S}_n$  for  $n$  large enough. Indeed suppose the reverse. As there exists  $x' \in \hat{S}_n$  with  $\|x' - x\| \leq \varepsilon_n^2$ , there exists  $x'' \in [x, x'] \cap \partial \hat{S}_n$ ,  $\|x - x''\| \leq \varepsilon_n^2$ . As  $B(x, r(x)) \subset S$  we also have  $B(x'', r(x) - \varepsilon_n^2) \subset S$  and  $d(x'', \partial S) \geq r(x) - \varepsilon_n^2 \geq r_0 - \varepsilon_n^2$ . That is impossible when  $n$  is large enough to have  $\varepsilon_n^2 < \frac{r_0}{2}$ . We also have that  $B(x, r(x) - 2\varepsilon_n^2) \cap \mathcal{Y}(\mathcal{X}_n) = \emptyset$ . Thus  $B(x, r(x) - 2\varepsilon_n^2) \in \mathcal{B}(\hat{S}_n, \mathcal{Y})$  and there exist  $x' \in \hat{S}_n$  and  $r'$  such that:

$$B(x, r(x) - 2\varepsilon_n^2) \subset B(x', r') \text{ with } B(x', r') \in \mathcal{C}^{\max}(\hat{S}_n, \mathcal{Y}). \quad (3)$$

We are now going to prove that for all  $(x', r')$  such that  $B(x, r(x) - 2\varepsilon_n^2) \subset B(x', r')$  and  $B(x', r') \in \mathcal{C}^{\max}(\hat{S}_n, \mathcal{Y})$  we have :

$$\|x - x'\| \leq \frac{1+K}{1-K}(2+3/r_0)\varepsilon_n^2. \quad (4)$$

Introduce  $x'^*$  a point of  $\partial S$  such that  $d(x', \partial S) = \|x' - x'^*\|$  and  $\gamma = r' - \|x' - x'^*\|$ . Notice that

1. Because  $x'^* \in \partial S$ , there exists  $y_i \in \mathcal{Y}$  such that  $\|y_i - x'^*\| \leq \varepsilon_n$ . Thus  $B(x', \|x' - x'^*\| + \varepsilon_n) \notin \mathcal{B}(\hat{S}_n, \mathcal{Y})$  and we obtain :  $r' \leq \|x' - x'^*\| + \varepsilon_n$ .

2. Conversely, because  $B(x', r')$  is in  $\mathcal{C}^{\max}(\hat{S}_n, \mathcal{Y})$  there exists  $y \in \mathcal{Y}$  such that  $\|x', y\| = r'$  and there exists  $z \in \partial S$  such that  $\|y - z\| \leq \varepsilon_n^2$  so we obtain :  $\|x' - x'^*\| \leq r' + \varepsilon_n^2$ .

Finally, we have that

$$-\varepsilon_n^2 \leq \gamma \leq \varepsilon_n \quad (5)$$

Remark now that there exists  $y_i \in \mathcal{Y}$  such that:

1.  $\|y_i - x'^*\| \leq \varepsilon_n$ , because  $\mathcal{Y}$  is a  $(\varepsilon, \varepsilon_n^2)$ -estimation of  $\partial S$ ;
2.  $y_i \notin \mathcal{B}(O_{x'^*}^{\text{out}}, r_0)$ , because  $y_i \in S$ ;
3.  $\|x' - y_i\| \geq r'$ , because  $B(x', r')$  is a ball of  $\mathcal{B}(\hat{S}_n, \mathcal{Y})$ .

Let us write  $y_i = au + bw$  where  $u = \frac{x'^* - x'}{\|x'^* - x'\|}$  and  $w$  is a unit vector of  $u^\perp$ . Notice that  $a \geq 0$  and that  $w$  can be chosen such that  $b \geq 0$  (see Figure 1).

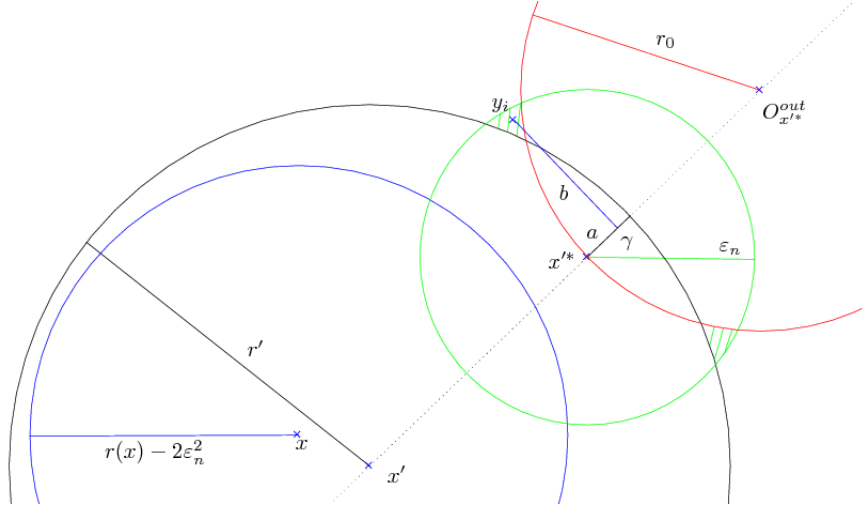


Figure 1: Let  $B(x', r') \in \mathcal{C}^{\max}(\hat{S}_n, \mathcal{Y})$  and  $x'^*$  be a point of  $\partial S$  such that  $d(x', \partial S) = \|x' - x'^*\|$ . Then there exists a  $y_i \in \mathcal{Y}$  in the green dashed area.

We have:

$$\begin{cases} a^2 + b^2 \leq \varepsilon_n^2 \\ b^2 + (r_0 - a)^2 \geq r_0^2 \\ r' \leq \|x' - y_i\|^2 = (r' - \gamma + a)^2 + b^2 \end{cases} \Rightarrow \begin{cases} a^2 + b^2 \leq \varepsilon_n^2 \\ 2ar_0 \leq \varepsilon_n^2 \\ 0 \leq \gamma^2 - 2(r' + a)\gamma + (1 + \frac{r'}{r_0})\varepsilon_n^2 \end{cases}$$

The last inequality gives  $\gamma \leq r' + a - \sqrt{(r' + a)^2 - (1 + \frac{r'}{r_0})\varepsilon_n^2}$  or  $\gamma > r' + a + \sqrt{(r' + a)^2 - (1 + \frac{r'}{r_0})\varepsilon_n^2}$ .

The second case is impossible for  $n$  large enough because (5) is incompatible with  $\gamma > r' + a > r_0 - 2\varepsilon_n^2 - \varepsilon_n^2/(2r_0)$ , and we must have:

$$\gamma \leq r' + a - \sqrt{(r' + a)^2 - (1 + \frac{r'}{r_0})\varepsilon_n^2} \leq \frac{(1 + \frac{r'}{r_0})\varepsilon_n^2}{r' + a} \leq \left(\frac{1}{r'} + \frac{1}{r_0}\right) \varepsilon_n^2.$$

Now as  $r' \geq r(x) - 2\varepsilon$ , when  $n$  is large enough we have  $r' \geq r_0/2$  and thus  $\gamma \leq 3\varepsilon_n^2/r_0$ . This, in turn guarantees that  $B(x', r' - 3\varepsilon_n^2/r_0) \subset \mathcal{B}(\hat{S}, \partial S)$ , so there exists  $B(x'', r(x'')) \in \mathcal{C}^{\max}(\hat{S}, \partial S)$  with  $B(x', r' - 3\varepsilon_n^2/r_0) \subset B(x'', r(x''))$ , that is:

$$\exists x'' \in \mathcal{M} \text{ such that } B(x', r' - 3\varepsilon_n^2/r_0) \subset B(x'', r(x'')). \quad (6)$$

Now, by (3) and (6) it follows that  $B(x, r(x) - (2+3/r_0)\varepsilon_n^2) \subset B(x'', r(x''))$ . As a consequence, by the  $K$ -regularity of the support, we obtain  $r(x'') \geq \|x'' - x\| + r(x) - (2+3/r_0)\varepsilon_n^2$ . Thus  $(2+3/r_0)\varepsilon_n^2 \geq \|x'' - x\| + r(x) - r(x'') \geq (1-K)\|x - x''\|$  and we finally obtain:

$$\|x - x''\| \leq \frac{(2+3/r_0)\varepsilon_n^2}{1-K}. \quad (7)$$

By (3) we also have  $r' \geq \|x - x'\| + r(x) - 2\varepsilon_n^2$  and by (6) we have  $r(x'') \geq \|x'' - x'\| + r' - 3\varepsilon_n^2/r_0$ . Summing these two inequalities gives:  $r(x'') \geq r(x) + \|x - x'\| + \|x' - x''\| - (2+3/r_0)\varepsilon_n^2$  so

$$\|x - x'\| \leq r(x'') - r(x) + (2+3/r_0)\varepsilon_n^2$$

So again using the  $K$ -regularity of the support and (7) we obtain  $\|x - x'\| \leq \frac{1+K}{1-K}(2+3/r_0)\varepsilon_n^2$ . This concludes the proof of (4).

Introduce  $C_0 = \min\{\sqrt{\frac{(1-K)r_0^3}{(1+K)(2r_0+3)}}, \frac{r_0}{2}\}$ . We now aim to prove that, for all  $x \in \mathcal{M}(S)$  there exists  $x'$  and  $r'_x$  such that

$$x' \in \mathcal{M}_{C_0, \hat{S}_n}(\mathcal{Y}), B(x', r'_x) \in \mathcal{C}^{\max}(\hat{S}_n, \mathcal{Y}), \text{ and } B(x, r(x) - 2\varepsilon_n^2) \subset B(x', r'_x). \quad (8)$$

More precisely we will show that this is realized for

$$r'_x = \max\{r', B(x, r(x) - 2\varepsilon_n^2) \subset B(x', r'), B(x', r') \in \mathcal{C}^{\max}(\hat{S}_n, \mathcal{Y})\}$$

and  $x'$  a point such that  $B(x, r(x) - 2\varepsilon_n^2) \subset B(x', r'_x)$  and  $B(x', r'_x) \in \mathcal{C}^{\max}(\hat{S}_n, \mathcal{Y})$ . Clearly there exists  $y \in \mathcal{S}(x', r'_x) \cap \mathcal{Y}$ . Suppose that  $\max\{\|y - z\|, z \in \mathcal{S}(x', r'_x) \cap \mathcal{Y}\} = l \leq r_0/2$ . Introduce a point  $z_0 \in \mathcal{S}(x', r'_x) \cap B(y, l)$ ,  $u = \frac{x' - y}{\|x' - y\|}$ ,  $x''_t = x' + tu$  and  $r''_t = \|z_0 - x''_t\|$  (See Figure 2).

Note first that  $\langle x' - z_0, u \rangle = \frac{2(r'_x)^2 - l^2}{2r'_x} \geq \frac{3}{4}$  (because  $l^2 \leq r_0^2/4 \leq (r'_x)^2$ ). Thus:

1. For  $t > 0$  :  $r''_t = \|x' - z_0 + tu\| = \sqrt{(r'_x)^2 + t^2 + 2t\langle x' - z_0, u \rangle} > r'_x$ .
2. For  $t > 0$  :  $B(x''_t, r''_t) \cap B^c(x', r'_x) \cap B(y, l) = \emptyset$ , so  $d(x''_t, \mathcal{Y} \cap B(y, l)) \geq r''_t$ . Consider now the function  $\rho(t) = d(x''_t, \mathcal{Y} \cap B^c(y, l)) - r''_t$ . For  $t = 0$  we have  $\rho(0) > 0$  so that, using continuity arguments, there exists  $t_0$  such that, for all  $t \in [0, t_0]$  we have  $\rho(t) > 0$  and thus (according to the preliminary remark)  $B(x''_t, r''_t) \cap \mathcal{Y} = \emptyset$ . Observe that, because  $x'$  is far enough from the boundary of  $S$ , we can choose  $t_0$  such that for all  $t < t_0$  we have  $x''_t \in \hat{S}_n$  and  $B(x''_t, r''_t) \cap \mathcal{Y} = \emptyset$ .

As  $B(x', r'_x)$  is a ball containing  $B(x, r(x) - 2\varepsilon_n^2)$ , in  $\mathcal{C}^{\max}(\hat{S}_n, \mathcal{Y})$  and has a maximal radius we must have that, for all  $0 < t < t_0$  there exists  $y_t \in B^c(x''_t, r''_t) \cap B(x, r(x) - 2\varepsilon_n^2)$ . Observe that  $y_t \in (B^c(x''_t, r''_t) \cap B(x', r'_x)) \cap B(x, r(x) - 2\varepsilon_n^2)$ .

So, taking  $t \rightarrow 0$  and using compactness arguments we obtain that there exists  $y' \in \mathcal{S}(x', r'_x) \cap B(y, l) \cap B(x, r(x) - 2\varepsilon_n^2)$  and finally using again that  $B(x, r(x) - 2\varepsilon_n^2) \subset B(x', r'_x)$  we have

$$y' \in \mathcal{S}(x', r'_x) \cap B(y, l) \cap \mathcal{S}(x, r(x) - 2\varepsilon_n^2).$$

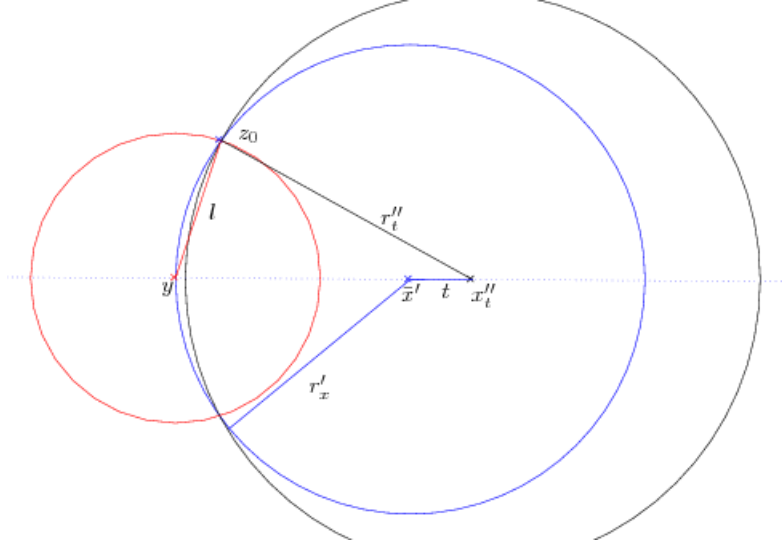


Figure 2: Construction of  $x_t''$ .

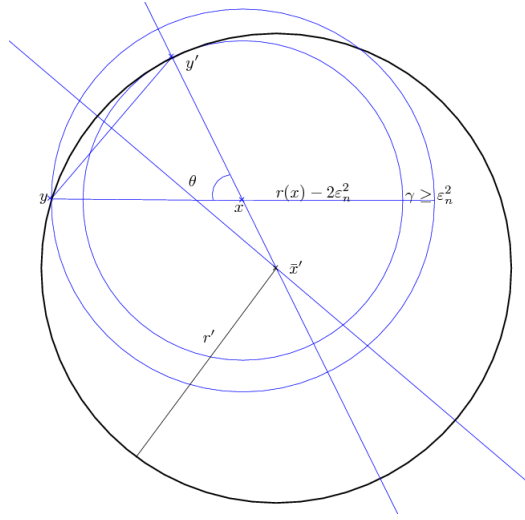


Figure 3:

Introduce  $\gamma = \|x - y\| - (r(x) - 2\varepsilon_n^2)$  and  $\theta = \angle yx'y'$ . Notice that  $\gamma > \varepsilon_n^2$  because  $d(y, \partial S) \leq \varepsilon_n^2$  (see Figure 3).

By  $\|y - y'\|^2 = \|y - x\|^2 + \|x - y'\|^2 + 2\langle y - x, x - y' \rangle$  it follows that  $\|y - y'\|^2 = (r(x) - 2\varepsilon_n^2)^2 + (r(x) - 2\varepsilon_n^2 + \gamma)^2 - 2(r(x) - 2\varepsilon_n^2)(r(x) - 2\varepsilon_n^2 + \gamma) \cos(\theta)$ . Thus we have:

$$\cos(\theta) = \frac{(r(x) - 2\varepsilon_n^2)^2 + (r(x) - 2\varepsilon_n^2 + \gamma)^2 - \|y - y'\|^2}{2(r(x) - 2\varepsilon_n^2)(r(x) - 2\varepsilon_n^2 + \gamma)}. \quad (9)$$



We also have  $x' - x = \|x' - x\|(\cos(\theta), -\sin(\theta))$ . Thus

$$\|x' - y\|^2 = (r(x) - 2\varepsilon_n^2 + \gamma + \|x' - x\| \cos(\theta))^2 + (\|x' - x\| \sin(\theta))^2. \quad (10)$$

But we also have  $\|x' - y\|^2 = \|x' - y'\|^2 = (r')^2 = (r(x) - 2\varepsilon_n^2 + \|x' - x\|)^2$ . Thus, from (10) we finally obtain:

$$(r(x) - 2\varepsilon_n^2 + \gamma)^2 + 2\|x' - x\|(r(x) - 2\varepsilon_n^2 + \gamma) \cos(\theta) = (r(x) - 2\varepsilon_n^2)^2 + 2\|x' - x\|(r(x) - 2\varepsilon_n^2),$$

and thus :

$$\|x' - x\| = \frac{1}{2} \frac{2\gamma(r(x) - 2\varepsilon_n^2) + \gamma^2}{(r(x) - 2\varepsilon_n^2) - (r(x) - 2\varepsilon_n^2 + \gamma) \cos(\theta)}.$$

This and (9) gives

$$\|x' - x\| = \frac{2\gamma(r(x) - 2\varepsilon_n^2)^2 + \gamma^2(r(x) - 2\varepsilon_n^2)}{\|y - y'\|^2 - 2\gamma(r(x) - 2\varepsilon_n^2) - \gamma^2}. \quad (11)$$

Finally, recall that  $r(x) - 2\varepsilon_n^2 \geq r_0/2$  and  $\gamma \geq \varepsilon_n^2$ , and observe that because of (4) we have  $\|x' - x\| \leq \frac{K+1}{1-K}(2+3/r_0)\varepsilon_n^2$ . Therefore, from (11) we obtain:

$$\frac{r_0^2 \varepsilon_n^2 (1 + \varepsilon_n^2 / r_0)}{2\|y - y'\|^2} \leq \|x - x'\| \leq \frac{K+1}{1-K}(2+3/r_0)\varepsilon_n^2$$

and thus, for  $n$  large enough that  $\varepsilon_n \leq \sqrt{r_0}$  we have:  $\|y - y'\|^2 \geq \frac{(1-K)r_0^3}{(1+K)(2r_0+3)}$ . Thus, if  $l \leq r_0/2$  we have  $l \geq \sqrt{\frac{(1-K)r_0^3}{(1+K)(2r_0+3)}}$ . This concludes the proof of (8).

Now from (4) and (8) we have that, for all  $\lambda < C_0$  :

$$\text{for all } x \in \mathcal{M}(S) \text{ there exists } x' \in \hat{\mathcal{M}}_{\lambda, \hat{S}_n}(y) \text{ such that } \|x - x'\| \leq \frac{K+1}{1-K}(2+3/r_0)\varepsilon_n^2. \quad (12)$$

Consider points  $x \in S$  such that  $d(x, \mathcal{M}) \geq e$ . In this last part of the proof we put  $l = d(x, \partial S)$ . Consider a point  $x'^* \in \partial S$  such that  $\|x - x'^*\| = l$ . As  $B(x, l) \subset \dot{S}$  one can introduce  $B(x', r(x'))$ , a ball of  $\mathcal{C}^{\max}(\dot{S}, \partial S)$  containing  $B(x, l)$ . Recall that the regularity condition on  $S$  allows the existence of  $O_{x'^*}^{\text{out}}$  such that  $B(O_{x'^*}^{\text{out}}, r_0) \subset S^c$  and observe that  $x, x', x'^*$  and  $O_{x'^*}^{\text{out}}$  are on the same line, and that  $r(x') = \|x' - x\| + l$  with  $e' = \|x' - x\| \geq e$ .

Because  $\mathcal{Y}$  is a  $(\varepsilon_n^2, \varepsilon_n)$ -estimation of  $\partial S$  there exists  $y \in \mathcal{Y}$  such that  $\|x'^* - y\| \leq \varepsilon_n$ . Obviously, as  $d(y, \partial S) \leq \varepsilon_n^2$ , we have  $\|y - x'\| \geq r(x') - \varepsilon_n^2$ , and since  $y \in S$  we have  $\|y - O_{x'^*}^{\text{out}}\| \geq r_0$ , that is  $y \in B(x'^*, \varepsilon_n) \cap B^c(x', r(x') - \varepsilon_n^2) \cap B^c(O_{x'^*}^{\text{out}}, r_0)$ . See Figure 4 for the construction.

Now, since  $\|x'^* - y\| \leq \varepsilon_n$  and  $\|y - O_{x'^*}^{\text{out}}\| \geq r_0$ , a short calculation shows that

$$\|x - y\|^2 \leq (r(x') - e')^2 + 2r_0 \left( 1 - \sqrt{1 - \frac{\varepsilon_n^2}{r_0^2}} \right) (r_0 + r(x') - e),$$

Thus there exists  $C_1$  such that, for  $n$  large enough, for all  $y_i \in \mathcal{Y}$  such that  $x \in \text{Vor}(y_i)$ , we have  $y_i \in B(x, \sqrt{(r(x') - e')^2 + C_1 \varepsilon_n^2}) \cap B^c(x', r(x') - \varepsilon_n^2) = E_x$ . See Figure 4 again.

For all  $y = x + au + bw$ , where  $u = \frac{x - x'}{\|x - x'\|}$  and  $w \in u^\perp$ ,  $y \in E_x$  we have

$$\begin{cases} b^2 + a^2 \leq (r(x') - e')^2 + C_1 \varepsilon_n^2 \\ b^2 + (e' + a)^2 \geq (r(x') - \varepsilon_n^2)^2. \end{cases}$$

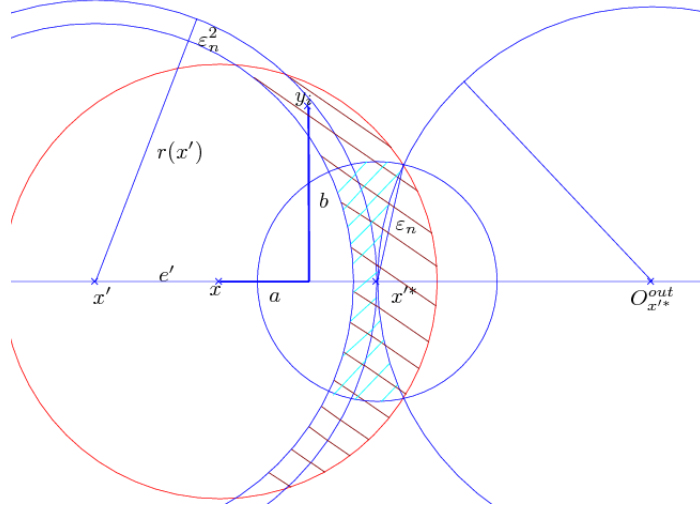


Figure 4: As there exists an observation in the blue crosshatched domain, all the  $y_i$  such that  $x \in \text{Vor}_y(y_i)$  are in  $E_x$ , the brown crosshatched domain.

After calculation we obtain that

$$\begin{cases} a_{\min}(e') = \min\{a, y = x + au + bw, u = \frac{x-x'}{\|x-x'\|}, w \in u^\perp, y \in E_x\} \geq r(x') - e' - \frac{C_1 + 2r(x')}{2} \frac{\epsilon_n^2}{e'} \\ b_{\max}(e') = \max\{b, y = x + au + bw, u = \frac{x-x'}{\|x-x'\|}, w \in u^\perp, y \in E_x\} \leq \sqrt{(r(x') - e')^2 + C_1 \epsilon_n^2} - a_{\min}(e'). \end{cases}$$

Considering  $\text{diam}(E_x)$ , we have:

$$\begin{cases} \text{diam}(E_x) \text{ decrease when } e' \text{ increases} \\ \text{diam}(E_x) \leq 2\sqrt{(r(x') - e')^2 + C_1 \epsilon_n^2} \\ \text{diam}(E_x) \leq 2b_{\max} \text{ when } a_{\min} \geq 0 \end{cases}$$

Choose  $e_n = \frac{(C_1 + 2r_0)}{r_0} \epsilon_n^2$ , and note that, for  $n$  large enough,  $a_{\min}(e_n) \geq 0$  therefore for  $n$  large enough, when  $e' > e_n$ ,

$$\text{diam}(E_x) \leq 2b_{\max}(e_n) \leq \sqrt{(r(x') - e_n)^2 + C_1 \epsilon_n^2} - r(x') + e_n + \frac{C_1 + 2r(x')}{2} \frac{\epsilon_n^2}{e_n}.$$

With the given value of the sequence  $e_n$  we can conclude that there exist  $C_2 > 0$  such that, for  $n$  large enough, if  $d(x, \mathcal{M}) \geq e_n$  then  $\text{diam}(E_x) \leq C_2$  and as a direct consequence  $x \notin \hat{\mathcal{M}}_{C_2, \hat{s}_n}(y)$ .

Finally, choose  $\lambda_0 = \min(C_0, C_2)$ . Then (12) and previous result allows to conclude that, for all  $\lambda < \lambda_0$  exists  $D = \max(\frac{(k+1)(2r_0+3)}{(1-K)r_0}, \frac{C_1+2r_0}{r_0})$  such that, for all  $\lambda < \lambda_0$  there for  $n$  large enough

$$d_h(\mathcal{M}(S), \hat{\mathcal{M}}_{\lambda_0, \hat{s}_n}(y)) \leq D\epsilon_n^2.$$

This concludes the proof.  $\square$

### 3 Practical aspects

#### 3.1 Proposition of an explicit estimator

In the case where the points are randomly drawn on  $S$  satisfying the regularity conditions of Theorem 1 and assuming additional hypothesis on the probability distribution, the following corollary proposes a ways to practically estimate the inner medial axis. For this we use  $C_r(\mathcal{X}_n)$ , the  $r$ -convex hull estimator of  $S$  (see [17]), to estimate  $S$  and to identify the subset  $\mathcal{Y}$  of sample points located close to the boundary.

**Corollary 1.** *Let  $\mathcal{X}_n = \{X_1 \dots X_n\}$  be an iid sample of points, drawn on  $S$  a  $(K, r_0)$ -regular compact set. Assume that the density  $f$  of the sample satisfies  $f(x) \geq f_0 > 0$  for all  $x \in S$ . For all  $r < r_0$  denote by  $\hat{C}_r(\mathcal{X}_n)$  the  $r$ -convex hull of  $\mathcal{X}_n$  and put  $\mathcal{Y}_r(\mathcal{X}_n) = \partial\hat{C}_r(\mathcal{X}_n) \cap \mathcal{X}_n$ . There exists  $\lambda_0$  such that, for all  $\lambda < \lambda_0$  there exists  $B_\rho$  such that*

$$d_h(\mathcal{M}, \hat{\mathcal{M}}_{\lambda, \hat{C}_r(\mathcal{X}_n)}(\mathcal{Y}_r(\mathcal{X}_n))) \leq B_\rho \left( \frac{\ln n}{n} \right)^{\frac{2}{d+1}} \quad e.a.s.$$

*Proof.* We refer the reader to [17] to be convinced that there exist constants  $a_r, b_r$  and  $c_r$  such that, eventually almost surely:

1.  $\hat{C}_r(\mathcal{X}_n)$  is (eventually almost surely) a  $(0, a_r(\ln n/n)^{\frac{2}{d+1}})$ -estimation of  $S$ ,
2.  $\mathcal{Y}_r(\mathcal{X}_n)$  is (eventually almost surely) a  $(a_r(\ln n/n)^{\frac{2}{d+1}}, b_r(\ln n/n)^{\frac{1}{d+1}})$ -estimation of  $S$ .

□

#### 3.2 Algorithm

**Step 1: Detection of boundary points.** First we have to identify the set  $\mathcal{Y}_r(\mathcal{X}_n) = \partial\hat{C}_r(\mathcal{X}_n) \cap \mathcal{X}_n$ . This is an easy task since:

**Proposition.**  $X_i \in \mathcal{Y}_r(\mathcal{X}_n)$  if and only if  $\max\{\|y - X_i\|, y \in (\text{Vor}_{\mathcal{X}_n}(X_i))\} \geq r$ .

*Proof.* Observe that  $\max\{\|y - X_i\|, y \in (\text{Vor}_{\mathcal{X}_n}(X_i))\} \geq r$ . This implies that there exists  $x$  such that  $B(x, r) \cap \mathcal{X}_n = \emptyset$ . Thus, by definition of  $\hat{C}_r$ ,  $B(x, r) \cap \hat{C}_r(\mathcal{X}_n) = \emptyset$  and  $x_n = (1/n)x + (r - 1/n)X_i \rightarrow X_i \in \hat{C}_r^c(\mathcal{X}_n)$  with  $X_i \in \hat{C}_r(\mathcal{X}_n)$ , so  $X_i \in \partial\hat{C}_r(\mathcal{X}_n)$ . Conversely, if  $X_i \in \partial\hat{C}_r(\mathcal{X}_n)$  then there exist two sequences  $x_n$  and  $y_n \in \hat{C}_r^c(\mathcal{X}_n)$  such that  $x_n \rightarrow X_i$ ,  $x_n \in B(y_n, r)$  and  $B(y_n, r) \cap \mathcal{X}_n = \emptyset$ . We have  $y_n \in S \oplus rB$ , which is compact. Thus, up to an extraction we can suppose that  $y_n \rightarrow y$ . As  $r < \|y_n - X_i\| \leq r + \|x_n - X_i\|$  we have in the limit  $\|y - X_i\| \geq r$ . Moreover, since for all  $n$ ,  $B(y_n, r) \cap \mathcal{X}_n = \emptyset$ , we have  $B(y, r) \cap \mathcal{X}_n = \emptyset$  and therefore  $y \in \text{Vor}_{\mathcal{X}_n}(X_i)$ . □

**Step 2: Computation of  $\hat{\mathcal{M}}_{\lambda, C_r(\mathcal{X}_n)}(\mathcal{Y}_r(\mathcal{X}_n))$ .** First recall that

$$\hat{\mathcal{M}}_{\lambda, C_r(\mathcal{X}_n)}(\mathcal{Y}_r(\mathcal{X}_n)) = \{z \in \text{Vor}_{\mathcal{Y}_r}(Y_i) \cap \text{Vor}_{\mathcal{Y}_r}(Y_j), z \in C_r(\mathcal{X}_n), \|Y_i - Y_j\| \geq \lambda\}$$

We claim that, for  $n$  large enough, we also have :

$$\hat{\mathcal{M}}_{\lambda, C_r(\mathcal{X}_n)}(\mathcal{Y}_r(\mathcal{X}_n)) = \{z \in \text{Vor}_{\mathcal{Y}_r}(Y_i) \cap \text{Vor}_{\mathcal{Y}_r}(Y_j), \text{Vor}_{\mathcal{Y}_r}(Y_i) \cap \text{Vor}_{\mathcal{Y}_r}(Y_j) \cap C_r(\mathcal{X}_n) \neq \emptyset, \|Y_i - Y_j\| \geq \lambda\}$$

Indeed, our regularity conditions imply that the total medial axis (not constraint to be included in  $S$ ) is composed of two distinct parts :  $\mathcal{M}(S)$ , the inside part, which is included in  $S \ominus r_0B$ ,

and  $\mathcal{M}^{\text{out}}$  the outside part which is included in  $(S \oplus r_0 B)^c$ . Reasoning as in the proof of Theorem 1 for the global medial axis allows us to say that for well chosen  $\lambda : \text{Vor}_{y_r}(Y_i) \cap \text{Vor}_{y_r}(Y_j)$  is close to a part of the global medial axis when  $\|Y_i - Y_j\| \geq \lambda$  and by a connectivity argument it is either close to the inner part or close to the outer part. It is then sufficient to have a point in  $C_r(\mathcal{X}_n)$  to conclude.

### 3.3 Choice of the parameters

We now have to tune the two parameters  $r$  and  $\lambda$ .

**The  $r$ -convex hull parameter** The crucial point when choosing a value for  $r$  is to identify observations that are really close to the boundary of  $S$ . Referring to the algorithm (Step 1) let us introduce  $r_i = \sup_{x \in V_i} \|X_i - x\|$ . One can clearly guess that, for all  $i$  such that  $X_i$  is far enough from the boundary,  $r_i$  is small. For instance, under our last hypotheses ( $S$  compact  $r_0$ -smooth and  $f$  bounded away from 0 on  $S$ ) the maximal spacing theory ([14] for the original paper and [1] for the extension to the same hypotheses as in Corollary 1) ensures that there exists  $\mu$  such that for all  $i$  such that  $d(X_i, \partial S) \geq \mu(\ln n/n)^{1/d}$  we have  $r_i \leq \mu(\ln n/n)^{1/d}/2$ . Conversely for points located on the boundary we have, according to [17],  $r_i \geq r_0$ . Thus the radii  $r_i$  associated to points close to the boundary that should be in  $\mathcal{Y}$  can (asymptotically) be considered as outliers in the empirical distribution of the  $r_i$  (they clearly have greater values and their proportion is decreasing to 0). According to this empirical consideration, and according to classical outlier detection based on quartile, we propose the following algorithm for the choice of  $r$ :

- $\mathcal{R} = \{r_i, i \in \{1, \dots, n\}\}$ ,  $\rho = \max_i(r_i)$ ,  $r = Q_3(\mathcal{R}) + 3.\text{IQ}(\mathcal{R})$
- while  $\rho > r$  (while there exists outliers) :  
 $\mathcal{R} := \mathcal{R} \cap [0, r]$ ,  $\rho := r$ ,  $r := Q_3(\mathcal{R}) + 3.\text{IQ}(\mathcal{R})$ .

**The  $\lambda$  parameter** Once  $r$  is chosen we can obtain  $C_r(\mathcal{X}_n)$  and  $\mathcal{Y}$  then compute:

$$\Lambda = \{\text{diam}(\{Y_{i_1}, Y_{i_{d+1}}\}), i_1 < \dots < i_{d+1}, \cap \text{Vor}_{\mathcal{Y}}(Y_i) \cap C_r(\mathcal{X}_n) \neq \emptyset\}$$

the finite set of all the different values of  $\lambda$  for which the medial axis estimator changes. We can see the points of  $\Lambda$  as the results of the mixture of two laws, one containing the small values (parasite branches) and the other one containing the largest values (stable part of the medial axis). Here we have no a priori idea of the proportion in each part of the mixture and the outlier detection approach is no more convenient. We propose to estimate the density of the  $\lambda$  (in the simulation part we used a Kernel density estimator with the Sheater and Jones [19] procedure for the bandwidth selection) and to select the critical values that are local minima of the density estimation.

### 3.4 Some simulations

To conclude, we now present some results. Each of these are based on large samples of 2000 points. For each example we present two graphics that illustrates the automatic procedure for the choice of the parameter:

1. the estimated density on the  $r_i$  and the choice  $r_0$ ,
2. the estimated density on the  $\lambda_i$  and the different choices for  $\lambda$  (one for each local minima of the density).

**The 2-dimensional case.** Since the medial axis was originally proposed for biological image analysis we propose to first test our method on some “biological” images. We present some results on a leaf image and on a sub-image of da Vinci’s Vitruvian Man (see Figure 5 which presents the original image and the “working” version in black and white).



Figure 5: original images and there extracted support

In addition to the figure illustrating the parameter choices we present, for each selected parameter  $\lambda$ :

1. The data set  $\mathcal{X}_n$  (yellow points);
2. The boundary points  $\mathcal{Y}$  (black points);
3. The parasite part medial axis (thin black lines);
4. The stable part of the medial axis (bold red lines).

When considering the Vitruvian man example, we found two local minima in the distribution of the  $\lambda$ , which leads to two medial axis estimations that are represented in Figure 6. The first one seems better and identifies the man’s skeleton while the second one only identifies his chest.

When considering the leaf example, we can see that the medial axis fails to recognize the leaf vein structure. This is not so surprising since we notice that the leaf does not fulfill the required regularity hypothesis. The choice of the  $\lambda$  parameters looks harder too but the five different medial axes are not so different.

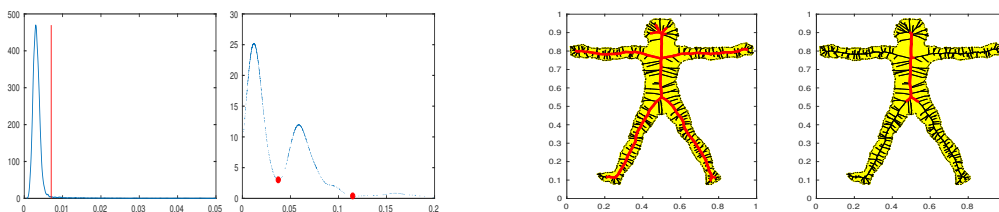


Figure 6: Vitruvian man : Estimated density of the  $r_i$  and chosen  $r_0$ : estimated density of the  $\lambda_i$  and the two different critical values; estimate medial axis for the smaller value; and estimated medial axis for the larger value.

**The 3-dimensional case.** In this section we draw points on  $M \oplus rB$  ( $r = 0.3$ ) for two cases where  $M$  is the medial axis of  $M \oplus rB$  : first a Moebus strip then a trefoil knot.

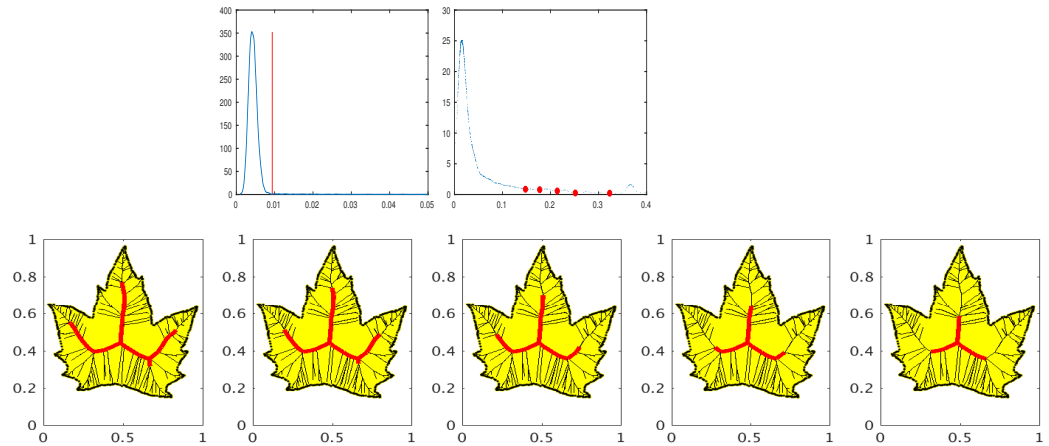


Figure 7: Leaf : Estimated density of the  $r_i$  and chosen  $r_0$ ; estimated density of the  $\lambda_i$  and the five different critical values; estimate medial axis for each value (increasing values).

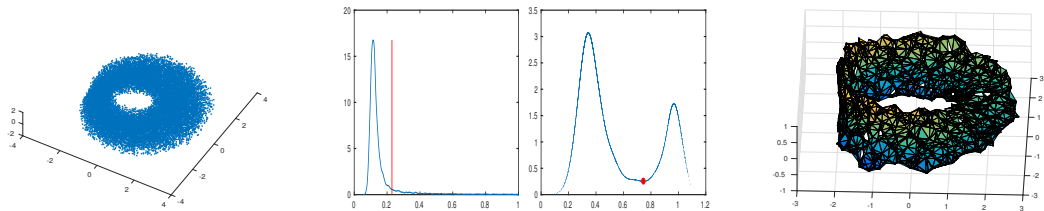


Figure 8: Moebius strip example : the data; the estimated density of  $r$  and  $\lambda$ ; the medial axis estimation

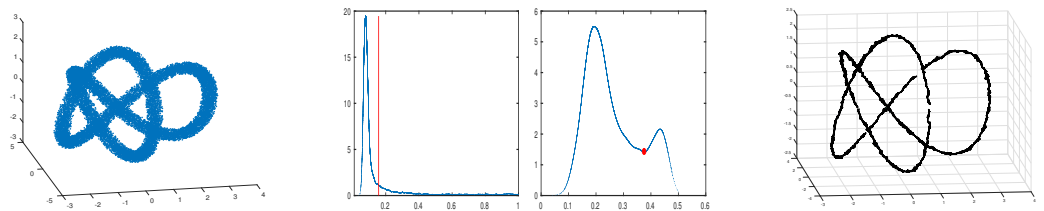


Figure 9: Trefoil knot example : the data; the estimated density of  $r$  and  $\lambda$ ; medial axis estimation

## References

- [1] C. Aaron, A. Cholaquidis, and A. Fraiman. A generalization of the maximal-spacings in several dimensions and a convexity test. *Extremes*, 10:605–634, 2017.
- [2] N. Amenta, S. Choi, and R. Kolluri. The power crust, unions of balls, and the medial axis transform. *Computational Geometry*, 19:127–153, 2001.
- [3] D. Attali, J. Boissonnat, and E. Edelsbrunner. *Mathematical Foundations of Scientific Visualization, Computer Graphics, and Massive Data Exploration*. Spr, 2009.

- [4] D. Attali and A. Montanvert. Modeling noise for a better simplification of skeletons. In *Proc. Internat. Conf. Impage Process.*, 1996.
- [5] H. Blum. A transformation for extracting new descriptors of shape. *Models for the Perception of Speech and Visual Form*, Issue 5:362–380, 1967.
- [6] J. W. Brandt and V. R. Algazi. Continuous skeleton computation by voronoi diagram. *CVGIP: Image Understanding*, 1992.
- [7] F. Chazal and A. Lieuter. The " $\lambda$ -medial axis". *Graphical Models*, 67(4):304–331, July 2005.
- [8] F. Chazal and R. Soufflet. Stability and finiteness properties of medial axis and skeleton. *Journal Of Dynamical and Control System*, 10:149–170, 2004.
- [9] S. W. Choi and H.-P. Seidel. Linear one-sided stability of mat for weakly injective 3d domain. In *Proc. 7th ACM Sympos. Solid Modeling Appl.*, pages 344–355, 2002.
- [10] A. Cuevas, P. Llop, and B. Pateiro-Lopez. On the estimation of the medial axis and inner parallel body. *Journal of Multivariate Analysis*, 129:171–185, August 2014.
- [11] D. Fritsch, S. Pizer, B. Morse, D. Eberly, and A. Liu. The multiscale medial axis and its applications in image registration. *Pattern Recognition Letter*, 15:445–452, May 1994.
- [12] C.R. Genovese, M. Perone-Pacifico, I. Verdinelli, and L. Wasserman. The geometry of non-parametric filament estimation. *Journal of the American Statistical Association*, 107(498), 2012.
- [13] C.R. Genovese, M. Perone-Pacifico, I. Verdinelli, and L. Wasserman. Minimax manifold estimation. *Journal Of Machine Learning Research*, 13:1562–1582, 2012.
- [14] S. Janson. Maximal spacings in several dimensions. *Annals of Propbability*, 15:274–280, 1987.
- [15] G. Matheron. *Image Analysis and Mathematical Morphology, Volume 2: Theoretical Advances*. Academic Press, London, 1988.
- [16] A. Montero and J. Lang. Skeleton pruning by contour approximation and the integer medial axis transform. *Computer and Graphics*, 36:477–487, 2012.
- [17] A. Rodríguez-Casal. Set estimation under convexity type assumptions. *Annales de l'Institut Henri Poincaré (B): Probability and Statistics*, 43(6):763–774, 2007.
- [18] D. Shaked and A. Bruckstein. Pruning medial axes. *Computer Vision and Image Understanding*, 69(2):156–169, February 1998.
- [19] S.J. Sheather and M.C. Jones. A reliable data-based bandwidth selection method for kernel density estimation. *Journal Of the Royal Statistical Society Ser. B.*, 53:683–690, 2012.
- [20] S. Xia, N. Ding, M. Jin, H. Wu, and Y. Yang. Medial axis construction and applications in 3d. In *Proc. IEE INFOCOM*, pages 305–309, 2013.
- [21] L. Younes. *Shapes and Diffeomorphisms*. Springer-Verlag Berlin Heidelberg, 2010.