

# An algorithm for computing Fréchet means on the sphere

Gabriele Eichfelder\*    Thomas Hotz<sup>†</sup>    Johannes Wieditz<sup>‡</sup>

February 26, 2018

## Abstract

For most optimisation methods an essential assumption is the vector space structure of the feasible set. This condition is not fulfilled if we consider optimisation problems over the sphere. We present an algorithm for solving a special global problem over the sphere, namely the determination of Fréchet means, which are points minimising the mean distance to a given set of points. The Branch and Bound method derived needs no further assumptions on the input data, but is able to cope with this objective function which is neither convex nor differentiable. The algorithm's performance is tested on simulated and real data.

Keywords: Branch and Bound, Fréchet means, Global optimisation

## 1 Introduction

Data assuming values on the unit sphere  $\mathbb{S}^2$ , also known as *spherical data*, arise frequently in applications, for example as directions of remanent magnetisation in soil samples, as arrival directions of cosmic particle showers, or as wind directions. For general information on the statistical analysis of spherical data, see e.g. [7], [11].

When considering optimisation problems on the sphere with these input data, one quickly encounters problems because most approaches fail due to the missing vector space structure. Neither the sum of two points of the sphere, nor the multiplication with a real number is an element of the sphere anymore.

A well-known example for such an optimisation problem arises from non-Euclidean statistics. There, one aim is to define a *mean* on the sphere. Because of the reasons

---

\*Institute for Mathematics, Technische Universität Ilmenau, Weimarer Straße 25, 98693 Ilmenau, Germany, gabriele.eichfelder@tu-ilmenau.de

<sup>†</sup>Institute for Mathematics, Technische Universität Ilmenau, Weimarer Straße 25, 98693 Ilmenau, Germany, thomas.hotz@tu-ilmenau.de

<sup>‡</sup>Institute for Mathematical Stochastics, Georg-August-Universität Göttingen, Goldschmidtstraße 7, 37077 Göttingen, Germany, johannes.wieditz@uni-goettingen.de

named above, how to define such a *mean* is at first sight unclear. Thus, often a more intrinsic point of view is used, namely considering the sphere as a Riemannian manifold with arc length as the metric measuring distances between two points. A mean may then be defined to be a point which minimises the average distance to some power measured to a set of given data points  $x_1, \dots, x_n \in \mathbb{S}^2$ , i.e. an element of

$$\operatorname{argmin} \left\{ \frac{1}{n} \sum_{i=1}^n d^p(m, x_i) \mid m \in \mathbb{S}^2 \right\} \quad (1)$$

for a fixed non-negative number  $p$ . This is the set of so-called *Fréchet- $p$ -means*. Note that on a Euclidean space, the case  $p = 2$  leads to the usual notion of mean, whereas the case  $p = 1$  results in a (spatial) median.

While computing Fréchet-2-means on the 1-sphere, i.e. the circle, is well investigated and efficiently possible, cf. [9], [12], the case of the 2-sphere turns out to be quite challenging. Even for real numbers  $p > 1$ , the objective function is neither differentiable nor convex in the sense of manifolds [2]. Therefore, Fréchet- $p$ -means need not be unique; however, since the objective function is continuous and the sphere is compact, there does always exist a Fréchet- $p$ -mean.

In the literature, there are many different ways to tackle the problem of computing a Fréchet mean on the sphere. A probabilistic method is introduced in [5], where the authors use a Monte Carlo approach. Moreover, [3] developed a steepest descent method for the case that the given points are located in one half of the sphere which implies that the unique Fréchet-2-mean lies in that half sphere where the objective function in addition is convex.

Furthermore, there exists a variety of problems similar to (1) each making slightly different assumptions or focussing on different aspects. If  $d$  is the Euclidean distance, there are several results about lower bounds and characterising means of the function  $m \mapsto \sum_{i=1}^n d^p(x_i, m)$  where the latter needs some further assumptions for the points, cf. [15], [16], [17]. Fréchet functions and Fréchet means over more general spaces were considered in [8] and [13]. An essential assumption there is an upper bound for the so-called local injectivity radius which is violated in our case of the 2-sphere.

A deterministic approach for computing the set of Fréchet means on the 2-sphere for general configurations of points based on (1) has not been considered so far; general deterministic optimisation algorithms on manifolds appear to be local search algorithms, see e.g. [1].

To fill this gap, we will in the following introduce a Branch and Bound method which is quite universal in the sense that it is able to cope with the non-differentiability and non-convexity of the objective function, requiring no assumptions on the given points  $x_1, \dots, x_n \in \mathbb{S}^2$ . Our approach will be outlined in Section 2 while lower bounds for the objective function will be established in Section 3. Using these bounds, we can then introduce the algorithm in Section 4; there we also consider numerical and efficiency aspects. We illustrate the performance of the algorithm in Section 5 using simulations as well as an application to real data, discussing the results obtained in Section 6.

## 2 Fréchet means on the sphere and a branching scheme

The optimisation problem we study here is that of finding so-called Fréchet- $p$ -means on the unit sphere. Therefore we consider the sphere as a metric space  $(\mathbb{S}^2, d)$ , endowed with arc length as the distance, a finite number of points  $x_1, \dots, x_n \in \mathbb{S}^2$  and a real number  $p \geq 0$ . Then we are looking for the set of all minimisers of the program

$$\min_{m \in \mathbb{S}^2} \frac{1}{n} \sum_{i=1}^n d^p(m, x_i) \quad (2)$$

where  $d(x, y) = \arccos \langle x, y \rangle$  for  $x, y \in \mathbb{S}^2$ ,  $\langle \cdot, \cdot \rangle$  denoting the standard dot product in  $\mathbb{R}^3$ . The objective function  $\hat{F}_n : \mathbb{S}^2 \rightarrow \mathbb{R}, m \mapsto \frac{1}{n} \sum_{i=1}^n d^p(m, x_i)$  of (2) is also referred to as *Fréchet- $p$ -function* (or Fréchet function in short).

One can consider (2) more generally on metric spaces. If these are fulfilling the Heine-Borel property (i.e. every closed bounded subset is compact), such minimisers always exist, although uniqueness is not guaranteed in general. Hereafter, we will focus on the spherical case in which at least existence is ensured.

The main idea of Branch and Bound in continuous optimisation is to divide (*branch*) the feasible set step-by-step into smaller subsets. One then tries to eliminate, under usage of suitable lower bounds, subsets which cannot contain a minimum (*bound*). Using this method we obtain an approximation of the set of all minimisers; here, we will use the definition of an  $(\varepsilon, \delta)$ -*approximation*. Recall that for  $\varepsilon, \delta > 0$  a set  $A \subseteq \mathbb{S}^2$  is an  $(\varepsilon, \delta)$ -approximation of the set of minimisers  $X$  of a function  $f : \mathbb{S}^2 \rightarrow \mathbb{R}$  if for all  $a \in A$  it holds that  $f(a) - \min_{m \in \mathbb{S}^2} f(m) \leq \varepsilon$  and for all  $x \in X$  there exists an  $a \in A$  with  $d(a, x) < \delta$ .

To apply a Branch and Bound algorithm to our problem, we have to first specify a rule how to subdivide the sphere as well as a discarding rule. For our purposes it is most appropriate to divide the sphere in spherical triangles. These are generated by a triple of non-coplanar vectors which are called *vertices* of the triangle. More precisely, a spherical triangle is the intersection of the (closed) convex cone spanned by its vertices with the sphere. Here, we start with the triangles induced by the vertices of a regular octahedron inscribed in the sphere as an initial triangulation. A triangle is then divided in one branch step at its midpoint of the longest side; in case of non-uniqueness we choose a side according to a deterministic rule. The so generated sequence of triangles fulfils the criterion of exhaustiveness, i.e. their diameters converge to 0, which is needed to show convergence of the algorithm, cf. [10, p. 204 ff.]. Furthermore, it is easily possible to determine whether a given point lies within a triangle using the following elementary geometric observation.

**Lemma 1.** *Let  $x \in \mathbb{S}^2$  and  $\Delta \subseteq \mathbb{S}^2$  be a spherical triangle with non-coplanar vertices  $d_1, d_2, d_3 \in \mathbb{S}^2$ .*

- (a) *Then  $x \in \Delta$  if and only if the solution  $\lambda = (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$  of the linear equations  $\sum_{i=1}^3 \lambda_i d_i = x$  is component-wise non-negative.*

(b) Moreover, let  $n_{ij} \in \mathbb{S}^2$  be orthogonal to  $\text{span}\{d_i, d_j\}$ ,  $i \neq j$ , with the property that for the third vector  $d_k$ ,  $k \neq i, j$  we have  $\langle d_k, n_{ij} \rangle \geq 0$ . Then  $x \in \Delta$  if and only if  $\langle x, n_{ij} \rangle \geq 0$  for all  $i, j \in \{1, 2, 3\}$ ,  $i \neq j$ .

*Proof.* (a) is the algebraic formulation of  $x$  being in  $\Delta$  if and only if  $x$  is in the convex cone spanned by  $d_1, d_2, d_3$ . In fact, the latter cone is the intersection of the three (closed) half-spaces determined by the conditions  $\langle x, n_{ij} \rangle \geq 0$  in (b).  $\square$

### 3 Lower bounds for the Fréchet function

The second ingredient necessary for designing a Branch and Bound method is an appropriate discarding rule. For this, we first note the following simple result.

**Lemma 2.** *Let  $g$  be a lower bound of our objective function  $f : \mathbb{S}^2 \rightarrow \mathbb{R}$  on a subset  $\Delta \subsetneq \mathbb{S}^2$  and let  $\bar{m} \in \mathbb{S}^2 \setminus \Delta$ . If  $f(\bar{m}) < g$  holds, then  $\Delta$  cannot contain a minimiser of  $f$ .*

Hence, our aim is to construct a lower bound for our objective function, the Fréchet function, on a spherical triangle.

Since the Fréchet function is Lipschitz continuous the construction of a lower bound using the Lipschitz constant is possible. However, this approach leads to lower bounds which are too weak for our purposes. We thus will now construct sharper estimates.

Having non-negative lower bounds  $g_i$  on  $\Delta$  for the functions  $m \mapsto d(m, x_i)$  for each  $i \in \{1, \dots, n\}$  we obtain a lower bound for the Fréchet function via

$$\frac{1}{n} \sum_{i=1}^n g_i^p. \quad (3)$$

In fact, we can calculate the minimum  $\min_{m \in \Delta} d(m, x_i)$  of the distance function on a spherical triangle analytically. This is of course the best lower bound for the distance function. The following proposition shows how to calculate this minimum.

**Theorem 3.** *Let  $x \in \mathbb{S}^2$  and consider a spherical triangle  $\Delta \subseteq \mathbb{S}^2$  with non-coplanar vertices  $d_1, d_2, d_3 \in \mathbb{S}^2$ . Furthermore, for  $i, j \in \{1, 2, 3\}$ ,  $i \neq j$  let*

$$g_{ij} = \begin{cases} |\arcsin(\langle x, d_i \times d_j \rangle)|, & \text{if } x = \lambda_1 d_i + \lambda_2 d_j + \lambda_3 (d_i \times d_j) \\ & \text{for some } \lambda_1, \lambda_2 \geq 0, \lambda_3 \in \mathbb{R}, \\ \min\{d(x, d_i), d(x, d_j)\}, & \text{otherwise} \end{cases} \quad (4)$$

which is the distance of  $x$  to the great circle arc connecting  $d_i$  and  $d_j$ .

Then for the distance of  $x$  to  $\Delta$  we have

$$d(x, \Delta) = \begin{cases} 0, & \text{if } x \in \Delta, \\ \min\{g_{12}, g_{23}, g_{31}\}, & \text{otherwise.} \end{cases} \quad (5)$$

In particular, the average  $\frac{1}{n} \sum_{i=1}^n d(x_i, \Delta)$  is a lower bound for  $\hat{F}_n$  on  $\Delta$ .

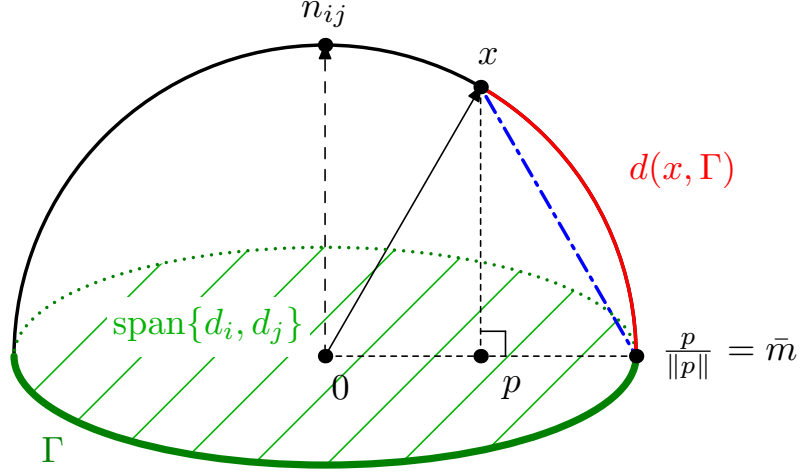


Figure 1: Upper half of the sphere with marked distance (red) of  $x \in \mathbb{S}^2$  from the great circle  $\Gamma$  induced by  $\text{span}\{d_i, d_j\}$  (green). The minimiser  $\bar{m}$  is given by  $\text{argmin}\{\|x - m\| : m \in \Gamma\} = \frac{p}{\|p\|}$ . The Euclidean distance  $\|x - \bar{m}\|$  is marked in blue (—)

*Proof.* Equation (5) holds for sure in case of  $x \in \Delta$ . Otherwise, because of monotonicity arguments the minimum will be attained at the boundary of  $\Delta$  which are three great circle arcs. Therefore, it is sufficient to show that for the distance of  $x$  to such an arc Equation (4) holds. We consider w.l.o.g. the great circle arc  $\kappa$  between  $d_1$  and  $d_2$ .

If  $x \perp \text{span}\{d_1, d_2\}$ , then  $x = \pm \frac{1}{\|d_1 \times d_2\|} (d_1 \times d_2)$  and every arc connecting  $x$  and  $\kappa$  has the same length  $\frac{\pi}{2}$ . Thus,  $d(x, \kappa) = \frac{\pi}{2} = g_{12}$ .

In the case  $x \not\perp \text{span}\{d_1, d_2\}$ , we compute at first the distance of  $x$  to the whole great circle  $\Gamma$  (containing  $\kappa$ ), see Figure 1. The point  $\bar{m}$  minimising the distance in arc length from  $x$  to  $\Gamma$  is the same as the one minimising the Euclidean distance from  $x$  to  $\Gamma$  since both result from each other by a monotonous transformation. For the latter one, application of Pythagoras' theorem leads to  $\|x - \bar{m}\|^2 = \|x - p\|^2 + \|p - \bar{m}\|^2$  where  $p$  is the orthogonal projection of  $x$  onto  $\text{span}\{d_1, d_2\}$ . Because the first term is constant, only the second term has to be minimised and the point  $\bar{m} \in \Gamma$  with the smallest Euclidean distance to  $p$  is  $\bar{m} = \frac{1}{\|p\|}p$ . This exists and is unique since  $x \not\perp \text{span}\{d_1, d_2\}$ . The distance from  $x$  to  $\Gamma$  is therefore given by  $d(x, \Gamma) = d(x, \bar{m})$ .

Assume now, that  $n_{12} = \frac{d_1 \times d_2}{\|d_1 \times d_2\|}$  (which – possibly up to sign – equals the corresponding vector in Lemma 1) and  $x$  lie in the same half space induced by  $\text{span}\{d_1, d_2\}$ . Then  $n_{12}$ ,  $x$  and  $\bar{m}$  lie on the same great circle arc and we have  $d(\bar{m}, x) + d(x, n_{12}) = d(\bar{m}, n_{12}) = \frac{\pi}{2}$ . Using the identity  $\arccos(\cdot) + \arcsin(\cdot) = \frac{\pi}{2}$  we obtain  $d(x, \Gamma) = \arcsin\langle x, d_1 \times d_2 \rangle$ . The case  $-(d_1 \times d_2)$  and  $x$  lying in the same half space can be treated analogously using the property  $\arcsin(-t) = -\arcsin(t)$  for all  $t \in [-1, 1]$ .

We consider now the computation of the distance of  $x$  to the great circle arc  $\kappa$ . If the minimiser  $\bar{m} = \text{argmin}_{m \in \Gamma} d(x, m) \in \kappa$ , then we can compute the distance

as shown above. Otherwise it can be shown, under usage of some monotonicity arguments, that one of the endpoints  $d_1, d_2$  of the arc has the smallest distance to  $x$ .

It remains to show that  $\bar{m} \in \kappa$  is equivalent to the first condition in Equation (4) holds. For that, we consider the great circle arc  $\kappa$  connecting  $d_1, d_2$  and assume  $x \notin \text{span}\{d_1, d_2\}$ . From our reasoning above we already know that the minimiser of  $\min_{m \in \Gamma} d(x, m)$  is obtained via normalising the orthogonal projection of  $x$  to  $\text{span}\{d_1, d_2\}$ . The points obtained from normalising points of  $\kappa$  can be written as  $\lambda_1 d_1 + \lambda_2 d_2$  with  $\lambda_1, \lambda_2 \geq 0$ . Since the orthogonal part to  $\text{span}\{d_1, d_2\}$  vanishes after projection, any point  $x$  having a minimiser in  $\kappa$  can be written as  $\lambda_1 d_1 + \lambda_2 d_2 + \lambda_3 (d_1 \times d_2)$  with  $\lambda_1, \lambda_2 \geq 0$  and  $\lambda_3 \in \mathbb{R}$  which is the first condition in Equation (4).  $\square$

**Remark 4.** One can easily show that this lower bound fulfils the exactness in the limit property, i.e. if the diameter of the triangle converges to zero then the bound converges to the function value at the limit point, cf. [10]. Thus the convergence of our method follows with the aid of elementary estimates and the triangle inequality. Furthermore, it is efficiently possible to check the conditions in Equation (4) and (5) using the results of Lemma 1.

## 4 The Spherical Branch and Bound algorithm (SBB algorithm)

The algorithm derived here is based on ideas from [6], where a similar algorithm was presented which aimed at finding globally optimal minimisers, but over a box in a linear space. Adapting this algorithm to our optimisation problem over the sphere leads to the following *Spherical Branch and Bound algorithm* or short *SBB algorithm* presented below. In the following, we explain the basic steps in detail.

Initially, we start with generating a list of all elements to be visited. Every item consists of a description of the triangle  $\Delta$ , its diameter  $s$  as well as a lower and upper bound for the objective function on this set. Here, the upper bound is calculated as the function's value at a certain point  $\text{MID}(\Delta)$ . We use the centroid of a spherical triangle for this. For the lower bound LB, we use the bound obtained from Theorem 3.

The outer loop (line 3–31) contains

- a branch part (line 5), in which we subdivide the currently chosen triangle using the BRANCH rule of Section 2,
- a bounding part (line 6–15), where elements of the branch part will be added to the list  $\mathcal{L}$  and elements fulfilling the discarding criterion from Lemma 2 with the lower bound LB obtained from (3) combined with (5), will be eliminated from  $\mathcal{L}$ ,

**Input** : Data points  $x_1, x_2, \dots, x_n$ ,  $p \geq 0$ , accuracies  $\varepsilon > 0, \delta > 0$ , initial subdivision  $\Delta_1, \Delta_2, \dots, \Delta_k$  of the sphere  $(\mathbb{S}^2, d)$  where  $\bigcup_{i=1}^k \Delta_i \supseteq \mathbb{S}^2$

**Output**:  $(\varepsilon, \delta)$ -approximation of the set of Fréchet- $p$ -means

```

1 Initialise list  $\mathcal{L} = \{X_1 = \{\Delta_1, s_1, v_1, u_1\}, \dots, X_k = \{\Delta_k, s_k, v_k, u_k\}\}$  where
    $s_i = \text{diam}(\Delta_i)$  is the diameter of  $\Delta_i$ ,  $v_i = \hat{F}_n(\text{MID}(\Delta_i))$  and  $u_i = \text{LB}(\Delta_i)$ .
2 Set  $A = \emptyset$ ,  $X^* = \{\Delta_1, s_1, v_1, u_1\}$ ,  $X_{\text{act}} = X^*$ ,  $u^* = -\infty$ ,  $v_{\text{act}} = \infty$ ,  $v_{\text{glob}} = v_{\text{act}}$ .
3 while  $\mathcal{L} \neq \emptyset$  do
4    $\mathcal{L} \leftarrow \mathcal{L} \setminus \{X^*\}$ 
5    $sB(X^*) \leftarrow \text{BRANCH}(X^*)$ 
6   for all  $\bar{X} = \{\bar{\Delta}, \bar{s}, \bar{v}, \bar{u}\} \in sB(X^*)$  do
7     Compute value  $\bar{v}$  of  $\hat{F}_n$  in  $\text{MID}(\bar{X})$ , diameter  $\bar{s}$  of  $\bar{X}$  and lower bound
      $\bar{u} = \text{LB}(\bar{X})$ .
8     if  $\bar{u} \leq v_{\text{glob}}$  then
9       Add  $\{\bar{\Delta}, \bar{s}, \bar{v}, \bar{u}\}$  at the end of  $\mathcal{L}$ .
10      if  $\bar{v} \leq v_{\text{act}}$  then
11         $X_{\text{act}} \leftarrow \bar{X}$ ,  $v_{\text{act}} \leftarrow \bar{v}$ ,  $v_{\text{glob}} \leftarrow \min\{v_{\text{act}}, v_{\text{glob}}\}$ 
12        Delete all elements  $\{\Delta, s, v, u\}$  from  $\mathcal{L}$  where  $u > v_{\text{glob}}$ .
13      end if
14    end if
15  end for
16  if  $\mathcal{L} \neq \emptyset$  then
17    Define  $X^* = \{\Delta^*, s^*, v^*, u^*\}$  as the first element of  $\mathcal{L}$  with
     $u^* = \min_{\{\Delta, s, v, u\} \in \mathcal{L}} u$ .
18    while  $\mathcal{L} \neq \emptyset \wedge v_{\text{act}} - u^* \leq \frac{\varepsilon}{2}$  do
19      if  $X_{\text{act}} \in \mathcal{L} \wedge s_{\text{act}} \leq \delta$  then
20        Set  $A \leftarrow A \cup \{X_{\text{act}}\}$ .
21        Delete  $X_{\text{act}}$  from  $\mathcal{L}$ .
22      end if
23      if  $X_{\text{act}} \in \mathcal{L}$  then
24         $X^* \leftarrow X_{\text{act}}$ , go to line 3.
25      end if
26      if  $\mathcal{L} \neq \emptyset$  then
27        Define  $X^* \leftarrow \{\Delta^*, s^*, v^*, u^*\}$  as the first element of  $\mathcal{L}$  with
         $u^* = \min_{\{\Delta, s, v, u\} \in \mathcal{L}} u$ .
28        Compute  $X_{\text{act}}$  as the first element in  $\text{argmin}_{\{\Delta, s, v, u\} \in \mathcal{L}} v$  and
        update upper bound  $v_{\text{act}} \leftarrow \hat{F}_n(\text{MID}(X_{\text{act}}))$ .
29      end if
30    end while
31  end if
32 end while

```

**Algorithm 1:** SBB algorithm for computing a  $(\varepsilon, \delta)$ -approximation of all Fréchet- $p$ -means

- a selection part (line 16–32), where the next element of the list is selected. Then we decide whether this element can be added to the list  $A$  of the  $(\varepsilon, \delta)$ -approximation (line 18–21) or if it has to be revisited in a later iteration. In the first case, we delete the element from  $\mathcal{L}$  and select the element with the smallest lower bound in  $\mathcal{L}$  as the new current element (line 25–28).

For an efficient implementation it is important to accelerate the computation steps as much as possible. An important key role plays the calculation of the lower bounds on the triangles resulting from the branch steps.

To compute the lower bound, we have to make several tests. For a triangle  $\Delta$  with vertices  $d_1, d_2, d_3 \in \mathbb{S}^2$  we have to decide whether  $x_\ell \in \Delta$  and if this is not the case, we have to test whether  $x_\ell \in \{\lambda_1 d_i + \lambda_2 d_j + \lambda_3 (d_i \times d_j) \mid \lambda_1, \lambda_2 \geq 0, \lambda_3 \in \mathbb{R}\}$  for  $(i, j) \in \{(1, 2), (2, 3), (3, 1)\}$  and for all  $\ell = 1, \dots, n$ . For the second case, we have to compute the cross product of the vertices anyway, so here it is more efficient to take the criterion of Lemma 1 (b) for testing. Then we only have to compute three scalar products instead of solving a linear equation. To evaluate the last three tests, we calculate a  $QR$  decomposition of the matrix  $(d_i \ d_j \ d_i \times d_j)$  where  $Q$  is orthogonal and  $R$  is an upper triangular matrix. Using this, the linear equation can be solved quickly via back-solving  $Rx = Q^\top x_\ell$  for  $x$ . De facto, we calculate there also only three scalar products for the right hand side; the additional effort consists in computing the  $QR$  decomposition. But since the decomposition remains the same for all  $\ell$ , this seems to be worthwhile especially for large  $n$ .

Further technicalities like considering numerical inaccuracies have been taken into account in the implementation used in Section 5.

## 5 Numerical test on simulated and real data

Now, we will compare the performance of our SBB algorithm for the case  $p = 2$  given four different types of simulated datasets where the points are drawn as independent and identically distributed samples from different distributions on the sphere. In particular, we consider a uniform distribution on a half sphere as well as on the whole sphere, a sample of points forming a tetrahedron rotated around a random angle, and two random points lying diametrically opposed. For points continuously distributed on the sphere one obtains an almost surely unique Fréchet mean, cf. [4, Corollary 2.3]. In the first case, this is even the unique local minimizer of the Fréchet function. One therefore would expect a fast convergence behaviour of SBB. For uniformly distributed data on the sphere, one obtains a statement about the average performance for determining a single Fréchet mean. In the other two cases, there are non-unique Fréchet means – for the tetrahedron there are four, for the diametrically opposed points even infinitely many minimisers, namely all points on the corresponding equator.

Recall that for drawing samples from a uniform distribution on the sphere it suffices to generate three-dimensional standard normally distributed vectors and normalise them afterwards.



For our simulation we used our own implementation in the programming language R, cf. [14], a computer with operating system Windows 7, Intel Core i5-5200U 2.2 GHz CPU and 4 GB RAM. The results are listed in Table 1 and are based on 100 repetitions each. Besides of the distribution the sample was drawn from, we also state the sampling size  $n$ , the theoretical number  $q$  of Fréchet means, the runtime, the number of iterations and the surface measure  $\nu(A)$  of the  $(\varepsilon, \delta)$ -approximation  $A$  for  $\varepsilon = \delta = 10^{-1}$  in relation to the whole surface of the sphere. All results are listed as means  $\pm$  empirical standard deviation.

distribution	$n$	$q$	time (in $s \pm$ s.d.)	iterations ( $\pm$ s.d.)	$\nu(A)$ ( $\pm$ s.d.)
uniform on half sphere	10	1	$3.0 \pm 1.6$	$468 \pm 145$	$1.3\% \pm 0.1\%$
	100	1	$13.8 \pm 2.2$	$539 \pm 44$	$1.3\% \pm 0.1\%$
uniform on sphere	10	1	$23.1 \pm 22.5$	$1,356 \pm 696$	$1.9\% \pm 0.5\%$
	100	1	$309.2 \pm 239.9$	$5,142 \pm 2,381$	$4.7\% \pm 1.7\%$
tetrahedron	4	4	$1,202.0 \pm 96.2$	$11,791 \pm 106$	$11.6\% \pm 0.1\%$
diametrical	2	$\infty$	$2,438.6 \pm 243.0$	$13,927 \pm 646$	$19.1\% \pm 0.3\%$

Table 1: Performance of the SBB algorithm for computing an  $(\varepsilon, \delta)$ -approximation  $A$  of the set of Fréchet means,  $\varepsilon = \delta = 10^{-1}$

Obviously, the computation of Fréchet means with the SBB algorithm took more time for samples coming from a uniform distribution on the sphere than for those originating from a uniform distribution on a half sphere. This can be explained by the great number of local minimisers of the Fréchet function in the first case which are eliminated from the list  $\mathcal{L}$  only after a long runtime. Nevertheless, in both cases SBB computes in pretty short time an approximation of the minimiser of moderate surface measure although we chose low accuracies of only  $\varepsilon = \delta = 0.1$ .

Note that for our second type of distribution the empirical standard deviation is relatively high. Since the location of the input data has a high influence on the properties of the Fréchet function, local minimisers can differ in their function value only slightly from the global minimum whereas in some cases this might not happen. In the first case mentioned, the local minimisers remain in the list  $\mathcal{L}$  for a much longer time which results in a longer running time for SBB.

For non-unique Fréchet means, the SBB algorithm takes a much higher amount of iterations for calculating an approximation of these. This is especially the case for our sample of diametrically opposed points where we have to approximate a whole great circle instead of a finite number of points. This is reflected in a much higher effort the one required for approximating a unique Fréchet mean.

We tested our algorithm also for computing the Fréchet means of a real data set called `DataB5FisherSpherical` which can be obtained from the R package `CircNNTSR`, see also [7, Example B5]. It contains  $n = 52$  measurements of magnetic remanence from specimen of sedimentary rock in Queensland. The aim is to determine the direction of the North Magnetic Pole at the time of the rock's formation. Because the sample is originated from the same region the mean direction of the magnetic field would give an approximation for this.

For illustration purposes the data is plotted in Figure 2. The depicted points are to be understood as directions of the magnetic field, more precisely as the North Magnetic Poles, of the given specimen. Points shown greyed out are on the opposite site of the spherical surface. The output of the SBB algorithm is shown in Figure 2. The  $(\varepsilon, \delta)$ -approximation covers a small area around the Geographic North Pole. Together with the fact that this is connected, one can derive that there is possibly a unique minimiser of the Fréchet-1-function. These data therefore suggest that the North Magnetic Pole at the time of the rock’s formation lay in that area.

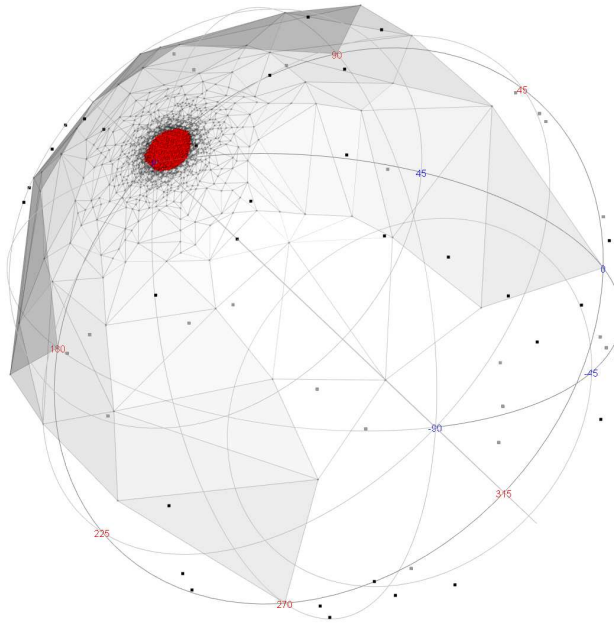


Figure 2:  $(\varepsilon, \delta)$ -approximation (red) for the Fréchet-1-mean of `DataB5FisherSpherical` computed by the SBB algorithm,  $(\varepsilon, \delta) = (10^{-2}, 10^{-1})$ , triangles obtained from branching are visualised as Euclidean triangles in grey, the data points given are shown as (■); longitudes and latitudes are sketched in red and blue, respectively.

## 6 Conclusions and discussion

We have derived a Branch and Bound method which determines all Fréchet- $p$ -means on the sphere for a given finite set of points. The main advantage of SBB is that we do not have to make any assumptions on the given points or the objective function since our lower bound is derived only via geometric considerations. Because of the very general structure of Algorithm 1, it will be applicable to other metric spaces too, as long as lower bounds and branching rules are known. In particular, extensions to higher dimensional spheres, real projective spaces, or simplices appears

straightforward. We noticed that the running time of SBB depends severely on how the points are distributed on the sphere. This is due to the fact that the location of the points influences the properties of the Fréchet function like number of local and global minimizers. Taking this information into account could improve the performance of our algorithm but requires appropriate criteria or statements for these properties, respectively.

## References

- [1] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Princeton, NJ, USA, 2007.
- [2] B. Afsari. Riemannian  $l^p$  center of mass: Existence, uniqueness, and convexity. *Proceedings of the American Mathematical Society*, 139(2):655–673, 2011.
- [3] B. Afsari, R. Tron, and R. Vidal. On the convergence of gradient descent for finding the Riemannian center of mass. *SIAM Journal on Control and Optimization*, 51(3):2230–2260, 2013.
- [4] M. Arnaudon and L. Miclo. Means in complete manifolds: uniqueness and approximation. *ESAIM: Probability and Statistics*, 18:185–206, 2014.
- [5] M. Arnaudon, L. Miclo, et al. A stochastic algorithm finding  $p$ -means on the circle. *Bernoulli*, 22(4):2237–2300, 2016.
- [6] G. Eichfelder, T. Gerlach, and S. Sumi. A modification of the abb method for box-constrained optimization and an application to inverse kinematics. *EURO Journal on Computational Optimization*, 4(1):93–121, 2016.
- [7] N. I. Fisher, T. Lewis, and B. J. J. Embleton. *Statistical Analysis of Spherical Data*. Cambridge University Press, Cambridge, 1987.
- [8] P. T. Fletcher, S. Venkatasubramanian, and S. Joshi. Robust statistics on Riemannian manifolds via the geometric median. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.
- [9] T. Hotz and S. Huckemann. Intrinsic means on the circle: uniqueness, locus and asymptotics. *Annals of the Institute of Statistical Mathematics*, 67(1):177–193, 2015.
- [10] M. Locatelli and F. Schoen. *Global optimization : theory, algorithms, and applications*. MOS-SIAM series on optimization ; 15. SIAM, Philadelphia, Pa., 2013.
- [11] K. V. Mardia and P. E. Jupp. *Directional Statistics*. Wiley, New York, 2000.

- [12] R. G. McKilliam, B. G. Quinn, and I. V. L. Clarkson. Direction estimation by minimum squared arc length. *IEEE Transactions on Signal Processing*, 60(5): 2115–2124, 2012.
- [13] M. Pálfia. Means in metric spaces and the center of mass. *Journal of Mathematical Analysis and Applications*, 381(1):383–391, 2011.
- [14] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016.
- [15] K. Stolarsky. The sum of the distances to  $n$  points on a sphere. *Pacific Journal of Mathematics*, 57(2):563–573, 1975.
- [16] G. Wagner. On means of distances on the surface of a sphere (lower bounds). *Pacific Journal of Mathematics*, 144(2):389–398, 1990.
- [17] E. Weiszfeld and F. Plastria. On the point for which the sum of the distances to  $n$  given points is minimum. *Annals of Operations Research*, 167(1):7–41, 2009.