Approximating Covering and Minimum Enclosing Balls in Hyperbolic Geometry

Frank Nielsen[1] and Gaëtan Hadjeres[2]

[1] École Polytechnique, France/Sony Computer Science Laboratories, Japan
Frank.Nielsen@acm.org
[2] Sony Computer Science Laboratories, France
gaetan.hadjeres@etu.upmc.fr

Abstract. We generalize the $O(\frac{dn}{\epsilon^2})$-time $(1+\epsilon)$-approximation algorithm for the smallest enclosing Euclidean ball [2, 10] to point sets in hyperbolic geometry of *arbitrary* dimension. We guarantee a $O\left(1/\epsilon^2\right)$ convergence time by using a closed-form formula to compute the geodesic α-midpoint between any two points. Those results allow us to apply the hyperbolic k-center clustering for statistical location-scale families or for multivariate spherical normal distributions by using their Fisher information matrix as the underlying Riemannian hyperbolic metric.

1 Introduction and prior work

Given a metric space $(X, d_{\mathrm{X}}(.,.))$, fitting the smallest enclosing ball of a point set $P = \{p_1, \ldots, p_N\}$ consists in finding the circumcenter $c \in X$ minimizing $\max_{p \in P} d_{\mathrm{X}}(c, p)$. In practice, this non-differentiable problem is computationally intractable as the dimension increases, and has thus to be approximated. A simple algorithm was proposed in [2] for euclidean spaces and generalized in [7] to dually flat manifolds.

In this article, we consider the case of the hyperbolic Poincaré conformal ball model $\mathbb{B}^d$ which is a model of d-dimensional geometry [8]. Even if balls in this hyperbolic model can be interpreted as euclidean balls with *shifted* centers [9], we cannot transpose directly results obtained in the euclidean case to the hyperbolic one because the euclidean enclosing balls may intersect the boundary ball $\partial\mathbb{B}^d$ (and are thus not proper hyperbolic balls, see Fig.1).

An exact solution for the hyperbolic Poincaré ball was proposed in [6] as a LP-type problem, but such an approach cannot be used in practice in high dimensions. A generic Riemannian approximation algorithm was studied by Arnaudon and Nielsen [1] but no explicit bounds were reported in hyperbolic geometry besides convergence, and moreover the heuristic assumed to be able to precisely cut geodesics and that step was approximated in [1].

We propose an intrinsic solution based on a closed-form formula making explicit the computation of geodesic α-midpoints (generalization of barycenter between two points) in hyperbolic geometry. We derive a $(1+\epsilon)$-approximation algorithm for computing and enclosing ball in hyperbolic geometry in arbitrary dimension in $O(\frac{dn}{\epsilon^2})$. This is all the more interesting from a machine learning perspective when dealing with data whose underlying geometry is hyperbolic. As an example, we illustrate our results on location-scale families or of multivariate spherical normal distributions. In the reminder, we assume the reader familiar with the basis of differential and Riemannian geometry, and recommend the textbook [8], otherwise.

The paper is organized as follows: Section 2 presents the exact computation of the α-midpoint between any arbitrary pair of points. Section 3 describes and analyzes the approximation algorithms for (i) fixed-radius covering balls and (ii) minimum enclosing balls. Section 4 presents the experimental results and discusses on k-center clustering applications.

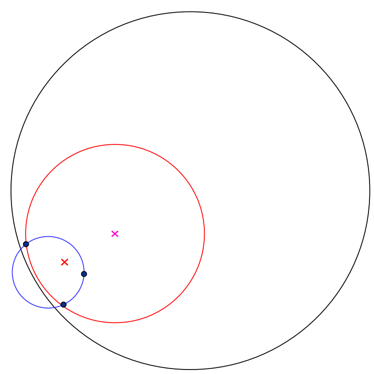

Fig. 1. Difference between euclidean MEB (in blue) and hyperbolic MEB (in red) for the set of blue points in hyperbolic Poincaré disk (in black). The red cross is the hyperbolic center of the red circle while the pink one is its euclidean center.

2 Geodesic α-midpoints in the hyperbolic Poincaré ball model

Let $\langle \cdot, \cdot \rangle$ and $\|x\| = \sqrt{\langle x, x \rangle}$ denote the usual scalar product and norm on the euclidean space $\mathbb{R}^d$. The Poincaré conformal ball model of dimension d is defined as the d-dimensional open unit ball $\mathbb{B}^d = \{x \in \mathbb{R}^d : \|x\| < 1\}$ together with the hyperbolic metric distance $\rho\,(.,.)$ given by:

$$\rho\,(p, q) = \operatorname{arcosh}\left(1 + \frac{2\|p-q\|^2}{(1-\|p\|^2)\,(1-\|q\|^2)}\right), \quad \forall p, q \in \mathbb{B}^d.$$

This distance induces on $\mathbb{B}^d$ a Riemannian structure.

Definition 1. *Let $p, q \in \mathbb{B}^d$ and $\gamma_{p,q}$ the unique geodesic joining p to q in the hyperbolic Poincaré model. For $\alpha \in [0,1]$, we define the α-midpoint, $p\#_\alpha q$, between p and q as the point $m_\alpha \in \gamma_{p,q}([0,1]) \subset \mathbb{B}^d$ on the geodesic $\gamma_{p,q}$ such that*

$$\rho(p, m_\alpha) = \alpha \rho(p,q).$$

Lemma 1. *For all $\alpha \in [0,1]$, we can compute the α-midpoint $p\#_\alpha q$ between two points p, q in the d-dimensional hyperbolic Poincaré ball model in constant time.*

Proof. We first consider the case where one of the point, say p, is equal to the origin $(0, \dots, 0)$ of the unit ball. In this case, the only geodesic running through p and q is the straight euclidean line. As the distance ρ on the hyperbolic ball is invariant under rotation around the origin, we can assume without loss of generality that $q = (x_q, 0, 0, \dots, 0)$, $x_q \geq 0$. In this case, we have:

$$\rho(p,q) = \operatorname{arcosh}\left(1 + \frac{2\|q\|^2}{1-\|q\|^2}\right) = \log\left(\frac{1+\|q\|}{1-\|q\|}\right) = \log\left(\frac{1+x_q}{1-x_q}\right), \tag{1}$$

using $\operatorname{arcosh}(x) = \log\left(x + \sqrt{x^2-1}\right)$. The α-midpoint $p\#_\alpha q$ has coordinates $(x_\alpha, 0, \dots, 0)$, $x_\alpha \geq 0$, which satisfies $\rho(p, p\#_\alpha q) = \alpha\rho(p,q)$. By (1), this is equivalent to solving, after exponentiating, $\frac{1+x_\alpha}{1-x_\alpha} = \left(\frac{1+x_q}{1-x_q}\right)^\alpha$. It follows that:

$$x_\alpha = \frac{c_{\alpha,q}-1}{c_{\alpha,q}+1}, \quad \text{where} \quad c_{\alpha,q} := e^{\alpha\rho(p,q)} \left(= \left(\frac{1+x_q}{1-x_q}\right)^\alpha \right). \tag{2}$$

For $p = (0, \dots, 0)$ and $q \neq p$ arbitrary, we have $p\#_\alpha q = \frac{\|x_\alpha\|}{\|q\|} q$, taking (2) as a definition for x_α with $c_{\alpha,q} = e^{\alpha\rho(p,q)}$.

Now, for arbitrary p and q, we first perform a *hyperbolic translation* T_{-p} of vector $-p$ to both p and q in order to resort to the preceding case, then compute the α-midpoint and translate it using the inverse hyperbolic translation T_p. The translation of $x \in \mathbb{B}^d$ by a vector $p \in \mathbb{B}^d$ of the hyperbolic Poincaré conformal ball model is given by (see [8], page 124 formula (4.5.5)):

$$T_p(x) = \frac{\left(1-\|p\|^2\right)x + \left(\|x\|^2 + 2\langle x, p\rangle + 1\right)p}{\|p\|^2\|x\|^2 + 2\langle x, p\rangle + 1}, \tag{3}$$

Since hyperbolic translations preserve the hyperbolic distance, using Def.1, we have indeed

$$T_p\left(T_{-p}(p) \#_\alpha T_{-p}(q)\right) = p\#_\alpha q.$$

Note that those computations can be made *exactly* without numerical loss since they involve only rationals and square root operations, see [3].

3 Approximation algorithms: Enclosing Balls and Minimum Enclosing Balls

In the following, we will denote by $P = \{p_1, \dots, p_N\}$ a set of N points of the hyperbolic Poincaré ball model. For $q \in \mathbb{B}^d$, we write $\rho(q, P) := \max_{p \in P} \rho(q,p)$.

Definition 2. *Let $r > 0$. A point $c \in \mathbb{B}^d$ is called the center of a hyperbolic enclosing ball of P of radius r (abbreviated* EHB(P, r)*) if*

$$\rho(c, P) \leq r. \tag{4}$$

If c is the center of a EHB *of minimal radius among all hyperbolic enclosing balls of P, then c is called the center of a minimum hyperbolic enclosing ball of P (abbreviated* MEHB(P)*).*

As the MEHBis unique [1], let R^* denote its radius and c^* its center. In practice, computing the MEHB(P) is intractable in high dimensions, we will focus on approximation algorithms by modifying (4).

Definition 3. *The point $c \in \mathbb{B}^d$ is the center of an $(1+\epsilon)$-approximation of* EHB(P, r) *if $\rho(c, P) \leq (1+\epsilon) r$ and the center of an $(1+\epsilon)$-approximation of* MEHB(P) *if $\rho(c, P) \leq (1+\epsilon) R^*$.*

3.1 A $(1+\epsilon)$-approximation of an enclosing ball of fixed radius

We generalize the EHB(P, r) approximation introduced in [10], Alg.2 Sect.3.1, to point clouds in the hyperbolic Poincaré ball model. Given P, $r > R^*$ and ϵ, this algorithm returns the center of a $(1+\epsilon)$-approximation of EHB(P, r).

Algorithm 1 $(1+\epsilon)$-approximation of EHB(P, r)

```
1: c_0 := p_1
2: t := 0
3: while ∃p ∈ P such that p ∉ B(c_t, (1+ε)r) do
4:     let p ∈ P be such a point
5:     α := (ρ(c_t, p) − r) / ρ(c_t, p)
6:     c_{t+1} := c_t #_α p
7:     t := t+1
8: end while
9: return c_t
```

As in [1] or [7], we took into consideration the fact that this geometry is not euclidean. The update move (step 5 and 6) consists in taking a point c_{t+1} on the geodesic from c_t to p such that the ball $B(c_{t+1}, r)$ "touches" p (i.e. such that $\rho(c_{t+1}, p) = r$).

Proposition 1. *Algorithm 1 returns the center of an $(1+\epsilon)$-approximation of* EHB(P, r) *in $O\left(1/\epsilon^2\right)$ iterations (exactly less than $4/\epsilon^2$ iterations).*

Proof. Let $\rho_t := \rho(c_t, c^*)$. Figure 2 illustrates the update of c_t, straight lines represent geodesic between points. From step 5 and 6, we have $\rho(c_{t+1}, p_t) = r$ which implies $\rho(c_{t+1}, c_t) > \epsilon r$. Since $B(c^*, R^*)$ is a MEHB, $\rho(c^*, p_t) \leq R^*$. Denote the angle $\angle c^* c_{t+1} c_t$ by θ and the distance $\rho(c^*, p_t)$ by r'. The hyperbolic law of cosines gives: $\cos(\theta)\operatorname{sh}(\rho_{t+1})\operatorname{sh}(r) = \operatorname{ch}(\rho_{t+1})\operatorname{ch}(r) - \operatorname{ch}(r')$, so that $\cos(\theta) \geq 0$ since $\operatorname{ch}(r') \leq \operatorname{ch}(r)$ and $\operatorname{ch}(\rho_{t+1}) \geq 1$.

Let θ' be the angle $\angle c_t c_{t+1} c^*$, it follows that $\cos(\theta') \leq 0$. Let h be the distance distance between c_t and c_{t+1}, the hyperbolic law of cosines gives $0 \leq \cos(\theta')\operatorname{sh}(h)\operatorname{sh}(\rho_{t+1}) = \operatorname{ch}(h)\operatorname{ch}(\rho_{t+1}) - \operatorname{ch}(\rho_t)$. Thus $\operatorname{ch}(\rho_t) \geq \operatorname{ch}(h)\operatorname{ch}(\rho_{t+1})$. After T iterations, we have the following inequality:

$$\operatorname{ch}(\rho_1) \geq \frac{\operatorname{ch}(\rho_1)}{\operatorname{ch}(\rho_T)} \geq \operatorname{ch}(h)^T \geq \operatorname{ch}(\epsilon r)^T , \tag{5}$$

which proves that the algorithm converges since $\operatorname{ch}(\epsilon r) > 1$. We can rewrite (5) as:

$$T \leq \frac{\log(\operatorname{ch}(\rho_1))}{\log(\operatorname{ch}(\epsilon r))} \leq \frac{\log(\operatorname{ch}(2r))}{\log(\operatorname{ch}(\epsilon r))} \leq \frac{4}{\epsilon^2} \tag{6}$$

using the fact that $\rho_1 \leq 2r$ and that $f := r \mapsto \frac{\log(\operatorname{ch}(2r))}{\log(\operatorname{ch}(\epsilon r))}$ is a decreasing function from $[0, +\infty[$ to $]0, +\infty[$ with $\lim_{r=0^+} f(r) = 4/\epsilon^2$.

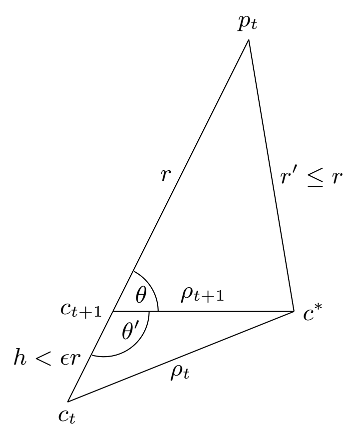

Fig. 2. Update of c_t

We now show how far from the real center c^* of the MEHB is the center of an $(1+\epsilon)$-approximation of EHB(P, r). We need the following lemma which generalizes Lemma 2 from [5]. For this, denote by $\langle ., . \rangle_p$ the scalar product given by the Riemannian metric on the tangent space $T_p\mathbb{B}^d$ in $p \in \mathbb{B}^d$ and by $\exp_p : T_p\mathbb{B}^d \to \mathbb{B}^d$ the exponential map.

Lemma 2. *For every tangent vector $v \in T_{c^*}\mathbb{B}^d$, there exists $p \in P \cap \mathcal{H}_v$ such that $\rho(c^*, p) = R^*$ where we denoted by*

$$\mathcal{H}_v = \left\{ \exp_{c^*}(u) \in \mathbb{B}^d, \quad u \in T_{c^*}\mathbb{B}^d, \quad \langle v, u \rangle_{c^*} \geq 0 \right\} \tag{7}$$

the points in $\mathbb{B}^d$ obtained by following geodesics whose tangent vector at point c^* lie in the half-space defined by v.

Proof. Assume it exists $v \in T_{c^*}\mathbb{B}^d$ such that for all $p \in P \cap \mathcal{H}_v$, $\rho(c^*, p) < R^*$. We will show that "moving" c^* in the direction $-v$ results in a new center c whose distance to P is strictly less than R^*, contradicting the fact that c^* is the center of MEHB(P).

For each point $q \in P$ not in $\mathcal{H}_v$ we have

$$\frac{d}{dt}\rho\left(\exp_{c^*}(-tv), q\right)\bigg|_{t=0} = \left\langle \frac{-\exp_{c^*}^{-1}(q)}{\rho(c^*, q)}, -v \right\rangle_{c^*} < 0 \tag{8}$$

by (7). So we can find $t > 0$ small enough to obtain $\rho\left(\exp_{c^*}(-tv), p\right) < R^*$ for all $p \in P$ since there is only a finite number of points in P.

Proposition 2. *Let c be the center of an $(1+\epsilon)$-approximation of* EHB(P, r). *We have the following inequality :*

$$\rho(c, c^*) \leq \operatorname{arcosh}\left(\frac{\operatorname{ch}((1+\epsilon)r)}{\operatorname{ch}(R^*)}\right) \tag{9}$$

where c^* and R^* are respectively the center and radius of the MEHB(P).

Proof. We can assume that $c \neq c^*$, otherwise (9) is true. Let $d := \rho(c, d^*)$. Consider the geodesic $\gamma : [0, d] \to M$ from c c^*. By applying the preceding lemma with $v := \dot{\gamma}(d)$, we obtain a point $p \in P \cap \mathcal{H}_v$ such that $\rho(c^*, p) = R^*$. By definition, the angle $\theta := \angle cc^*p$ is obtuse.

We name h the distance $\rho(c, p)$ and apply the hyperbolic law of cosines to obtain $0 \geq \cos(\theta)\operatorname{sh}(d)\operatorname{sh}(R^*) = \operatorname{ch}(d)\operatorname{ch}(R^*) - \operatorname{ch}(h)$. Since c is the center of an $(1+\epsilon)$-approximation de EHB(P, r), $h \leq (1+\epsilon)r$. we deduce $\operatorname{ch}((1+\epsilon)r) \geq \operatorname{ch}(h) \geq \operatorname{ch}(d)\operatorname{ch}(R^*)$ from which we derive (9).

3.2 A $(1 + \epsilon + \epsilon^2/4)$-approximation of MEHB($P$)

We can use the previous results to derive an algorithm computing the MEHB(P) in hyperbolic geometry of arbitrary dimension. The proposed algorithm (Alg.2) consists in a dichotomic search of the radius of the MEHB(P). Indeed, we can discard a radius smaller than R^* using (6) and use inequality (9) in order to obtain a tighter bound.

Proposition 3. *Algorithm 2 returns the center of an $(1+\epsilon+\frac{\epsilon^2}{4})$-approximation of* MEHB($P$) *in $O\left(\frac{1}{\epsilon^2}\log\left(\frac{1}{\epsilon}\right)\right)$ iterations.*

Proof. In Alg.2, as $\rho(p_1, P) > R^*$, the first call to Alg.1 returns an $(1+\epsilon/2)$-approximation of EHB($\rho(p_1, P), P$). The fact that c is the center of an $(1+\epsilon/2)$ approximation of EHB($P, r_{\max}$) becomes a loop invariant. We also ensure that at each loop

$$r_{\min} \leq R^* \leq r_{\max}, \tag{10}$$

Algorithm 2 $(1+\epsilon)$-approximation of EHB(P)

$c := p_1$
$r_{\max} := \rho(c, P)$; $r_{\min} = \frac{r_{\max}}{2}$; $t_{\max} := +\infty$
$r := r_{\max}$;
repeat
  $c_{\text{temp}} := \text{Alg1}\left(P, r, \frac{\epsilon}{2}\right)$, interrupt if $t > t_{\max}$ in Alg1
  if call of Alg1 has been interrupted **then**
    $r_{\min} := r$
  else
    $r_{\max} := r$; $c := c_{\text{temp}}$
  end if
  $dr := \frac{r_{\max} - r_{\min}}{2}$; $r := r_{\min} + dr$; $t_{\max} := \frac{\log(\text{ch}(1+\epsilon/2)r) - \log(\text{ch}(r_{\min}))}{\log(\text{ch}(r\epsilon/2))}$
until $2dr < r_{\min}\frac{\epsilon}{2}$
return c

so that the maximum number T of iterations of $\text{Alg1}(P, r, \epsilon/2)$ can be bounded by

$$T \leq \frac{\log(\text{ch}(\rho(c_1, c^*)))}{\log(\text{ch}(\epsilon r/2))} \leq \frac{\log(\text{ch}(1+\epsilon/2)r) - \log(\text{ch}(r_{\min}))}{\log(\text{ch}(r\epsilon/2))} \quad (11)$$

using (9) and the left side of (6) and (10). At the end of the repeat-until loop, we know that $r_{\max} \leq R^* + dr$ and that c is the center of an $(1+\epsilon/2)$ approximation of EHB$(P, r_{\max})$. So

$$\rho(c, P) \leq \left(1 + \frac{\epsilon}{2}\right) r_{\max} \leq \left(1 + \frac{\epsilon}{2}\right)\left(R^* + r_{\min}\frac{\epsilon}{2}\right) \leq \left(1 + \epsilon + \frac{\epsilon^2}{4}\right) R^*. \quad (12)$$

This approximation is obtained in precisely $O\left(\frac{N}{\epsilon^2}\log\left(\frac{1}{\epsilon}\right)\right)$ since after T iterations of the main loop, $dr \approx \frac{R^*}{2^T}$.

4 Experimental results

4.1 Performance

To evaluate the performance of Alg. 2, we computed MEHB centers for a point cloud of $N = 200$ points for different values of the dimension d and the precision parameter ϵ. For each test, the point cloud was sampled uniformly (euclidean sampling) in the unit ball of dimension d. In order to check the relevance of our theoretical bounds, we plotted in Fig. 3 the *average number* of α-midpoints calculations and the mean execution time as a function of ϵ for different values of d. We evaluated convergence comparing the returned values of c to a value c^* computed with high precision. The algorithms have been implemented in Java using the arbitrary-precision arithmetic library Apfloat.

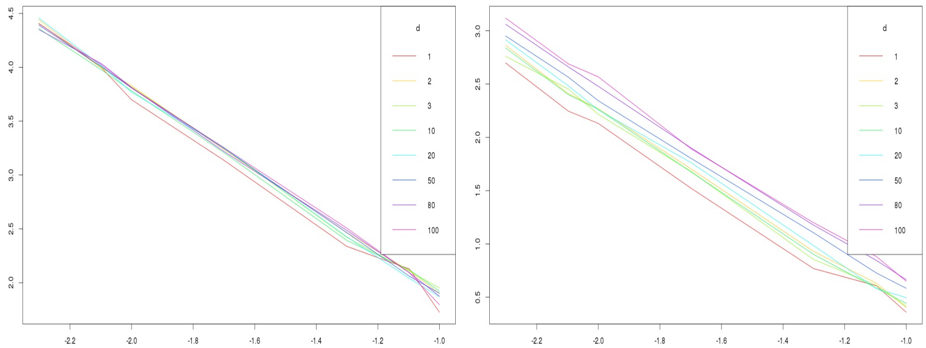

Fig. 3. Number of α-midpoint calculations (left) as a function of ϵ and execution time (right) as a function of ϵ, both in logarithmic scale for different values of d. We observe that the number of iterations does not depend on d, and that the running time is approximately $O(\frac{dn}{\epsilon^2})$ (vertical translation in logarithmic scale).

4.2 One-class clustering in some subfamilies of multivariate distributions

One-class clustering consists, given a set P of points, to sum up the information contained in P while minimizing a measure of distortion. In our case, we associate to a point set P a point c minimizing $\rho(c, P)$, i.e. the center of the MEHB(P). This is particularly relevant when applied to parameterizations of subfamilies of multivariate normal distributions. Indeed, a sufficiently smooth family of probability distributions can be seen as a statistical manifold (a Riemannian manifold whose metric is given by the Fisher information matrix, see [4]). As proved in [4],

- the family $\mathcal{N}(\mu, \sigma^2 \mathrm{I}_d)$ of d-variate normal distributions with scalar covariance matrix (I_d is the $d \times d$ identity matrix)
- the family $\mathcal{N}\big(\mu, \mathrm{diag}\left(\sigma_1^2, \ldots, \sigma_n^2\right)\big)$ of n-variate normal distributions with diagonal covariance matrix
- the family $\mathcal{N}(\mu_0, \Sigma)$ of d-variate normal distributions with fixed mean μ_0 and arbitrary positive definite covariance matrix Σ

all induce a hyperbolic metric on their respective parameter spaces. We can thus apply Alg.2 to those subfamilies in order to perform one-class clustering using their natural Fisher information metric. An example showing how different from the euclidean case the results are is given in Fig.4. We used the usual Möbius transformation between the Poincaré upper half-plane and the hyperbolic Poincaré conformal ball model, see [6].

As a byproduct, we can derive a solution to the k-center problem for those specific subfamilies of multivariate normal distributions in $2^{O(k \log k \log(1/\epsilon)/\epsilon^2)} dn$.

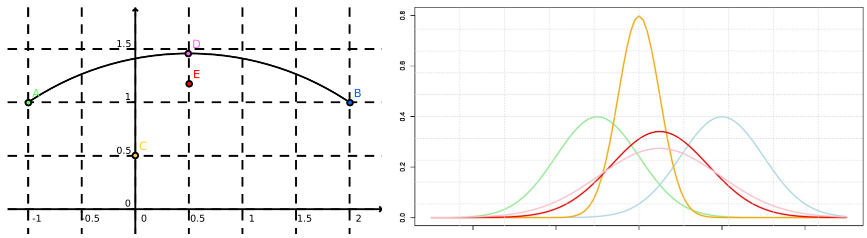

Fig. 4. (Best viewed in color). Graphical representation of the center of the MEHB, in the (μ, σ) superior half-plane (left), by showing corresponding probability density functions (right). In red (point E) is represented the center of MEHB(A,B,C). In pink (point D) is the 1/2-midpoint between A and B. The geodesic joining A to B is also displayed

5 Acknowledgements

We would like to thank François Pachet for his kind support and advice.

References

1. Arnaudon, M., Nielsen, F.: On approximating the Riemannian 1-center. Computational Geometry 46(1), 93–104 (2013)
2. Badoiu, M., Clarkson, K.L.: Optimal core-sets for balls. Comput. Geom. 40(1), 14–22 (2008)
3. Burnikel, C., Fleischer, R., Mehlhorn, K., Schirra, S.: Efficient exact geometric computation made easy. In: Proceedings of the Fifteenth Annual Symposium on Computational Geometry, Miami Beach, Florida, USA, June 13-16, 1999. pp. 341–350 (1999), http://doi.acm.org/10.1145/304893.304988
4. Costa, S.I., Santos, S.A., Strapasson, J.E.: Fisher information matrix and hyperbolic geometry. In: Information Theory Workshop, 2005 IEEE. pp. 3–pp. IEEE (2005)
5. Kumar, P., Mitchell, J.S., Yildirim, E.A.: Approximate minimum enclosing balls in high dimensions using core-sets. Journal of Experimental Algorithmics (JEA) 8, 1–1 (2003)
6. Nielsen, F., Nock, R.: Hyperbolic Voronoi diagrams made easy. In: Computational Science and Its Applications (ICCSA), 2010 International Conference on. pp. 74–80. IEEE (2010)
7. Nock, R., Nielsen, F.: Fitting the smallest enclosing Bregman ball. In: Machine Learning: ECML 2005, pp. 649–656. Springer (2005)
8. Ratcliffe, J.: Foundations of hyperbolic manifolds, vol. 149. Springer Science & Business Media (2006)
9. Tanuma, T., Imai, H., Moriyama, S.: Revisiting hyperbolic Voronoi diagrams from theoretical, applied and generalized viewpoints. In: Voronoi Diagrams in Science and Engineering (ISVD), 2010 International Symposium on. pp. 23–32. IEEE (2010)
10. Tsang, I.W., Kocsor, A., Kwok, J.T.: Simpler core vector machines with enclosing balls. In: Proceedings of the 24th international conference on Machine learning. pp. 911–918. ACM (2007)