Extending results of TDA to Bregman point clouds

Masterarbeit

Máté László Telek

Juli 2020

Ruprecht-Karls-Universität Heidelberg
Fakultät für Mathematik und Informatik

Betreuer: Prof. Dr. Anna Wienhard

Abstract

Topological Data Analysis (TDA) aims to study the structure of a data set represented as a finite metric space. Edelsbrunner and Wagner [1] extended the framework of TDA to data measured with Bregman divergences, which made TDA methods accessible to new applications.

In this thesis, we consider how the usual constructions and algorithms in TDA can be generalized to the setting with Bregman divergences. Furthermore, we present a stability result of persistence diagrams and an approximative Künneth formula.

Zusammenfassung

Topologische Datenanalyse (TDA) verfolgt das Ziel, die Struktur einer Datenmenge zu untersuchen, welche als ein endlicher metrischer Raum repräsentiert ist. Edelsbrunner und Wagner [1] erweiterten die Theorie für Daten, deren Distanzen durch Bregman-Divergenzen gegeben sind. Damit wurden Methoden der TDA für neue Anwendungen zugänglich gemacht.

In dieser Arbeit diskutieren wir, wie die üblichen Konstruktionen und Algorithmen der TDA im Falle von Bregman-Divergenzen verallgemeinert werden können. Ferner präsentieren wir ein Stabilitätsresultat für Persistenz-Diagramme und eine approximative Künneth-Formel.

Acknowledgements

I would first like to thank my supervisor, Prof. Dr. Anna Wienhard for the opportunity to work on this fascinating topic. I wish to thank all the people at Heidelberg University, who I met during my study. Thanks should also go to my friends, who encouraged me and never let me down. Finally, I must express my very profound gratitude to my parents for providing me with unfailing support throughout my years of study.

Selbstständigkeitserklärung

Ich versichere, die Masterarbeit selbstständig und lediglich unter Benutzung der angegebenen Quellen und Hilfsmittel verfasst zu haben.
Ich erkläre weiterhin, dass die vorliegende Arbeit noch nicht im Rahmen eines anderen Prüfungsverfahrens eingereicht wurde.

Ort, Datum:

Unterschrift

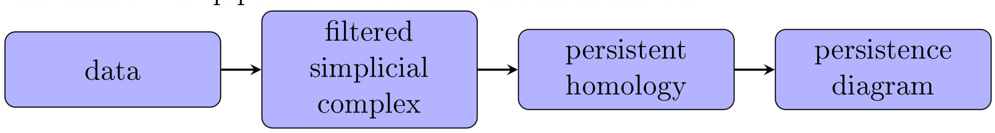
Contents

1. Introduction

Understanding large and complex data sets can be very demanding. The goal of data analysis is to find descriptors of the data that are easy to understand, can be computed fastly, but still reflect relevant information. Topological Data Analysis (TDA) provides such a descriptor, the so-called *persistence diagram*. It describes the shape of the data in the form of easily understandable diagrams.

In 1999, V. Robins introduced in her paper [2] *persistent homology* that became a central tool in TDA. The first papers about TDA were written by Edelsbrunner et al. [3] in 2002 and by Carlsson et al. [4] in 2005. Since then, there are many available papers, surveys, and books about topological data analysis. It has become an important research topic in modern mathematics. TDA also has a wide range of applications. It is used, e.g., to analyse viral evolution [5], sensor networks [6], neuronal network dynamics [7].

The usual TDA pipeline can be summarized as follows:



Usually, the data is represented as a finite metric space, also called a *point cloud*. From this, we build a *filtered simplicial complex* that stores the topological features of the data set at different *scale parameters*. After applying the homology functor, we get persistent homology groups that we can characterize using persistence diagrams.

The entire process strongly depends on the choice of the metric. In practice, it could be challenging to find the "right" metric for the data set. For some applications, one wishes to measure the data with a dissimilarity measure that does not satisfy the axioms of a metric function. The *Kullback-Leibler divergence* is such a dissimilarity measure. It is commonly used to analyse text documents [8] and images [9],[10]. The *Itakura-Saito divergence* is well suited for sound data [11]. Both divergences are members of the class of *Bregman divergences*.

Edelsbrunner and Wagner [1] introduced a TDA framework where the data is measured with a Bregman divergence. This thesis aims to understand TDA in this new setting. We recall the statements from [1] and we try to extend some results in the usual TDA, such as *stability theorem* and *Künneth theorem*, to our new Bregman setting.

The thesis is organized as follows. In chapter 2, we recall some basic definitions and results from TDA. The third chapter is about Bregman geometry. We introduce Bregman divergences and show some elementary properties of them. These statements will be helpful later when we are doing TDA, but one can enjoy this chapter on its own. In chapter 4, we will consider how the usual TDA setting can be extended to the setting with Bregman divergences. At the end of the chapter, we also describe some algorithms for the computation of persistence diagrams. In chapter 5, we present a stability result that tells us that a slight noise in the input data does not have too much effect on the corresponding persistence diagram. In the last chapter, we try to find a *Künneth-type formula* for *Bregman point clouds*. One can find some Matlab code in the appendix that we will use to do some experiments with synthetic data sets during the thesis. Matlab was also used to create figures that can be found in the chapters.

To avoid confusion, we fix some notations that we will use throughout the thesis.

List of symbols

$\mathbb{N}$	positive integers
$\mathbb{N}_0$	nonnegative integers
$\mathbb{R}_+$	positive real numbers
$\mid a \mid$	absolute value of $a \in \mathbb{R}$
$\langle x, y \rangle$	Euclidean scalar product of $x, y \in \mathbb{R}^n$
$\|x\|$	Euclidean norm of $x \in \mathbb{R}^n$
$\|x\|_\infty$	maximum norm of $x \in \mathbb{R}^n$
$\mathbf{k}$	a fixed but arbitrary field
$\mathbf{k}$-$\underline{\mathbf{Mod}}$	category of $\mathbf{k}$-modules
$\bigsqcup$	disjoint union of sets
$\mathrm{card}(X)$	cardinality of a finite set X
$\mathcal{P}(X)$	power set of X
$\Delta(X) := \mathcal{P}(X) \setminus \{\emptyset\}$	full simplicial complex of X
$\mathrm{int}(U)$	interior of a subset U in a topological space
$\overline{U}$	closure of U
$\mathrm{bd}(U) := \overline{U} \setminus \mathrm{int}(U)$	boundary of U
$B_r(m)$	closed ball with radius r and center m in a metric space (M, d_M)
$\mathrm{Aff}(P)$	affine hull of a subset $P \subset \mathbb{R}^n$
$\mathrm{Conv}(P)$	convex hull of P
J_f	Jacobi matrix of a differentiable function f
$\mathrm{im}(f)$	image of a function f
$\mathrm{epi}(f)$	epigraph of a function f
$\cong$	isomorphic
$\simeq$	homotopic
$\mid K \mid$	geometric realization of a simplicial complex K
$\mathcal{C}^{\mathcal{I}}$	functor category

2. Preliminaries

In this chapter, we give a brief review of topological data analysis. The aim in doing so is to fix some basic definitions and recall the most important theorems. The chapter is based on [12] and [13].

2.1. From data to persistence diagrams

In the usual setting of TDA, we always start with a finite metric space, which is also called a point cloud. In the next step of TDA, we build a *filtered simplicial complex.*

Definition 2.1 *A finite collection K of finite nonempty sets is called a simplicial complex, if $\sigma \in K$ and $\emptyset \neq \tau \subset \sigma$ implies $\tau \in K$.*

We call each $\sigma \in K$ a simplex. A simplex $\sigma \in K$ is called a vertex if it contains exactly one element. We write $\mathrm{Vert}(K)$ for the set of vertices of K.

The dimension of a simplex $\sigma \in K$ is defined as $\dim(\sigma) := \mathrm{card}(\sigma) - 1$. The dimension of K is given by $\dim(K) := \max_{\sigma \in K} \dim(\sigma)$.

A simplicial map $f : K \to K'$ between simplicial complexes is a function $f : \mathrm{Vert}(K) \to \mathrm{Vert}(K')$ such that for every $\{v_0, \dots, v_n\} \in K$ it holds that $\big\{f(v_0), \dots f(v_n)\big\} \in K'$.

We denote the category of simplicial complexes and simplicial maps by **Simp**.

One can turn every finite collection of sets into a simplicial complex.

Definition 2.2 *Let F be a finite collection of sets. The nerve of F is defined as*

$$\mathrm{Nrv}(F) := \{T \subset F \mid \bigcap_{\tau \in T} \tau \neq \emptyset\}.$$

There are several ways to construct a simplicial complex from a point cloud. Each of them has its advantages and disadvantages. The *Čech complex* is commonly used, because it is suitable to prove theoretical statements.

Definition 2.3 *Let (M, d_M) be a point cloud and $r \geq 0$ a fixed nonnegative real number. The simplicial complex*

$$\check{\mathrm{C}}\mathrm{ech}(M, r) := \mathrm{Nrv}\left(\big\{B_r(m)\big\}_{m \in M}\right) = \{\sigma \subset M \mid \bigcap_{m \in \sigma} B_r(m) \neq \emptyset\}$$

is called the Čech complex of (M, d_M) at scale r.

If the point cloud (M, d_M) lives in some Euclidean space $\mathbb{R}^n$ i.e. $M \subset \mathbb{R}^n$ and the metric d_M is the restricted Euclidean distance, then the *Nerve Theorem* implies that $\mid \check{\mathrm{C}}\mathrm{ech}(M,r) \mid$ and $\bigcup_{m\in M} B_r(m)$ have the same homotopy type, and therefore they have isomorphic homology groups.

Theorem 2.4 ***(Nerve Theorem) [12]*** *Let F be a finite collection of closed sets in an Euclidean space such that the intersection of each subcollection is either empty or contractible. In this case, the union of the sets in F and $\mid \mathrm{Nrv}(F) \mid$ have the same homotopy type.*

The Čech complex can be extremely high dimensional. That could cause a lot of problems if we actually wanted to compute persistence diagrams. The *Vietoris–Rips complex* is easier to compute. There are well-developed algorithms [14] and software packages [15] that can be used in practice.

Definition 2.5 *Let (M, d_M) be a point cloud and $r \geq 0$ a fixed nonnegative real number. The simplicial complex*

$$\mathrm{Rips}(M,r) := \{\sigma \subset M \mid \forall m,n \in \sigma : d_M(m,n) \leq r\}$$

is called the <u>*Vietoris-Rips complex*</u> *of (M, d_M) at scale r.*

Due to a small *interleaving* between the Vietoris-Rips and Čech complexes (see 6.25 for further details), the Vietoris-Rips complex is similar enough to the Čech complex, and therefore to the union of the balls, to reflect relevant information about the structure of the point cloud.

In the above definitions, we always fixed a real number r. Since there is no natural choice of a scale parameter r, we would like to study all parameters at once.

Definition 2.6 *A* <u>*parametrised simplicial complex*</u> *is a functor $K_\bullet : [0,\infty) \to$* <u>**Simp**</u>*, where $[0,\infty)$ denotes the poset category associated to the set of nonnegative real numbers with its usual order.*

If all the maps $K_r \to K_t$, $r \leq t$ are inclusions, we call $K_\bullet$ a <u>*filtered simplicial complex*</u>*.*

Example 2.7 *Let (M, d_M) be a point cloud.* $\mathrm{Rips}(M,\bullet)$ *and* $\check{\mathrm{C}}\mathrm{ech}(M,\bullet)$ *are filtered simplicial complexes.*

In algebraic topology, one uses homology groups to characterize topological features of spaces, for more details, we refer to [16, Chapter 2]. Applying the homology functor $\mathrm{H}_n(\bullet)$ with coefficients in a fixed field $\mathbf{k}$, we can turn each parametrised simplicial complex $K_\bullet$ into a *persistence* **k**-*module* that is also called the <u>persistent homology</u> of $K_\bullet$.

Definition 2.8 *A persistence* **k***-module is a functor* $V_\bullet : [0, \infty) \to \mathbf{k}\text{-}\underline{\mathbf{Mod}}$.

We wish to have easily understandable invariants that describe the isomorphism classes of persistence modules. To do so, we decompose the persistence modules into simpler pieces.

Definition 2.9 *Let* $I \subset [0, \infty)$ *be an interval. The persistence* **k***-module* $\mathbb{1}_I(\bullet)$ *given by*

$$\mathbb{1}_I(i) := \begin{cases} \mathbf{k} & \text{if } i \in I \\ 0 & \text{if } i \notin I \end{cases} \quad \text{and } \mathbb{1}_I(i \leq j) := \begin{cases} \mathrm{id}_{\mathbf{k}} & \text{if } i, j \in I \\ 0 & \text{otherwise} \end{cases}$$

is called an interval module.

Theorem 2.10 ***(Crawley-Boevey Theorem) [13, Thm. 1.6]***
Every pointwise finite-dimensional persistence **k***-module* $V_\bullet$ *is a direct sum of interval modules. Moreover, the decomposition is unique up to isomorphism and permutation of the terms in the direct sum.*

Because a point cloud (M, d_M) contains only a finite number of points, we can apply the Crawley-Boevey Theorem to persistent homology $\mathrm{H}_n\left(\mathrm{Rips}(M, \bullet)\right)$ and $\mathrm{H}_n\left(\check{\mathrm{C}}\mathrm{ech}(M, \bullet)\right)$. Since the interval decomposition is unique up to isomorphism, it seems to be the invariant what we are looking for. We just need to find a way to note the lower and upper bound of the intervals.

Definition 2.11

a) A multiset is a pair (S, μ) *where* S *is a set and* $\mu : S \to \mathbb{N}_0 \cup \{\infty\}$ *is a function. For each* $s \in S$*, we call* $\mu(s)$ *the multiplicity of* s.

b) We call a bijective map

$$\eta : \bigcup_{s \in S} \coprod_{i=1}^{\mu(s)} \{s\} \to \bigcup_{s' \in S'} \coprod_{i=1}^{\mu'(s')} \{s'\}$$

a bijection between the multiset (S, μ) *and* (S', μ').
We write $\eta : (S, \mu) \xrightarrow{1:1} (S', \mu')$ *for short.*

With other words, we interpret each point of a multiset with multiplicity m as m individual points and a bijection of multisets is a bijective map between the resulting sets.

Definition 2.12 *Let $V_\bullet$ be a pointwise finite-dimensional persistence* $\mathbf{k}$*-module with interval decomposition* $V_\bullet \cong \bigoplus_{d \in D} \mathbb{1}_{I_d}(\bullet)$*. The persistence diagram of $V_\bullet$ is a multiset with underlying set* $(\mathbb{R} \cup \{\infty\})^2$ *and multiplicity function μ given by*

$$\mu(a,b) := \begin{cases} \operatorname{card}\left(\{d \in D \mid a \text{ is a lower, } b \text{ is an upper bound of } I_d\ \}\right) & \text{if } a \neq b \\ \infty & \text{if } a = b \end{cases}$$

We denote this multiset by $\operatorname{dgm}(V_\bullet)$*.*

By definition, every point with positive multiplicity lies above the diagonal. The reason to consider the diagonal with infinite multiplicity is rather technical. It allows us to define a metric for persistence diagrams, see section 5.1 for more details.

It must be noted that there is a new approach to TDA based on homotopical methods at chain complexes level. In the frame of this new approach, the points in a persistence diagram lying on the diagonal are not just noise anymore. Thus, we may lose some information if we put infinite multiplicity to the diagonal. This new approach is out of our scope, we just refer to the original paper [17] for further details.

3. Bregman geometry

The notion of *Bregman divergences* was introduced in [18]. The most basic example is the *squared Euclidean distance*. More interesting are the *Kullback-Leibler divergence* and the *Itakura-Saito divergence*, both are commonly used in applied mathematics.

Bregman divergences can be seen as a measure of the distance between two points, even they satisfy only one of the three metric axioms, as we will see in section 3.2. Because of the missing axioms, it is not a surprise that Bregman geometry will differ from the usual Euclidean geometry.

Since the Bregman divergences are not symmetric, we will always need to define a *primal* and a *dual* object. It turns out that dual objects have better behaviour. Using *Legendre duality* we can translate primal objects to dual and dual objects back to primal objects, which allows us to show some useful properties for primal objects.

Our main goal is to introduce objects like *Bregman balls*, *Voronoi diagrams* and their properties, which we will need in the data analysis part later, but one can read this chapter on its own and enjoy the beauty of Bregman geometry. This chapter is based on [1].

3.1. Functions of Legendre type

Bregman divergences are induced from *functions of Legendre type*. In this part, we will study such functions. They are strictly convex differentiable functions with some additional properties. The authors [19], [20] and [1] require slightly different properties. We use the definition from [1] and require that the length of the gradient goes to infinity whenever we approach the boundary of the domain. This will ensure that the *convex conjugate* of a function of Legendre type is a function of Legendre type, however, this additional requirement has a disadvantage we can not restrict our domain arbitrarily.

Definition 3.1 *Given a nonempty open convex set* $\Omega \subset \mathbb{R}^n$*, a function* $F : \Omega \to \mathbb{R}$ *is of Legendre type if* F *is*

(L1) strictly convex,

(L2) differentiable,

(L3) $\|\nabla F(x_n)\| \longrightarrow \infty$ *whenever* $x_n \longrightarrow x \in \mathrm{bd}(\Omega)$ *.*

Using standard real analysis tools, one can show that the following functions are of Legendre type.

Example 3.2

function of Legendre type	$F : \Omega \to \mathbb{R}$
one dimensional half the squared Euclidean norm	$\mathbb{R} \to \mathbb{R}, x \mapsto \frac{1}{2}x^2$
one dimensional convex Shannon entropy	$\mathbb{R}_+ \to \mathbb{R}, x \mapsto x \cdot ln(x) - x$
one dimensional Burg entropy	$\mathbb{R}_+ \to \mathbb{R}, x \mapsto 1 - ln(x)$

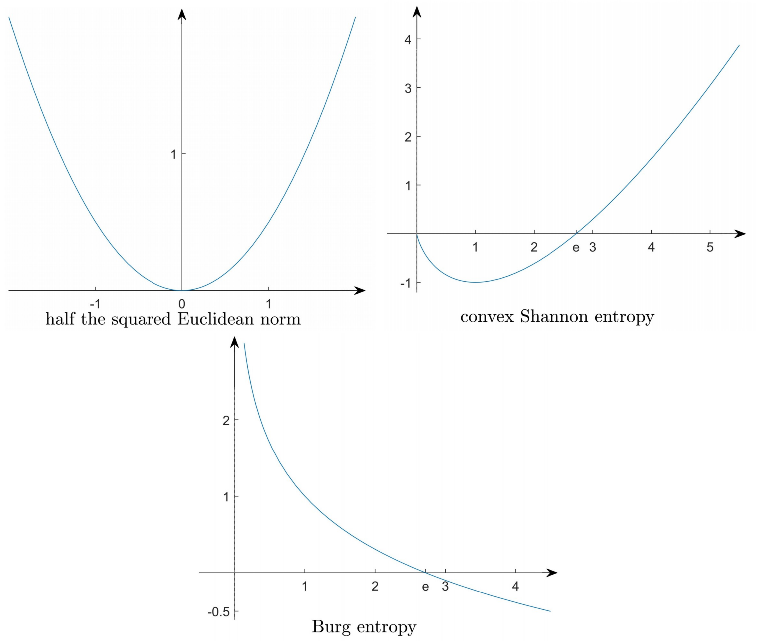


Figure 1: Graph of functions of Legendre type

The data that we want to study, lives in higher-dimensional spaces. The next proposition gives us a nice way to build higher-dimensional functions of Legendre type.

Proposition 3.3 *Let $F : \Omega \to \mathbb{R}$, $G : \Omega' \to \mathbb{R}$ be functions of Legendre type. Then $F + G : \Omega \times \Omega' \to \mathbb{R}, (x, y) \mapsto F(x) + G(y)$ is a function of Legendre type.*

Proof: Since $\Omega \subset \mathbb{R}^n$ and $\Omega' \subset \mathbb{R}^m$ are nonempty open convex sets, the Cartesian product $\Omega \times \Omega' \subset \mathbb{R}^{n+m}$ is nonempty, open and convex as well.

The strict convexity of $F + G$ follows from the strict convexity of F and G.

The partial derivatives of $F + G$ in $a = (a_1, \ldots, a_{n+m}) \in \Omega \times \Omega'$ are given by

$$\frac{\partial F + G}{\partial x_i}(a) = \begin{cases} \frac{\partial F}{\partial x_i}(a_1, \ldots, a_n) & \text{for } 1 \leq i \leq n \\ \frac{\partial G}{\partial x_i}(a_{n+1}, \ldots, a_{n+m}) & \text{for } n+1 \leq i \leq n+m \end{cases}$$

They are continuous, since F and G are differentiable. This implies the differentiability of $F + G$ on $\Omega \times \Omega'$.

It remains to be shown that the length of the gradient of $F + G$ goes to infinity whenever we approach the boundary of $\Omega \times \Omega'$. Let $\big((x_n, y_n)\big)_{n\in\mathbb{N}}$ be a sequence in $\Omega \times \Omega'$ such that $(x_n, y_n) \longrightarrow (x, y) \in \operatorname{bd}(\Omega \times \Omega')$. It is well known that

$$\begin{aligned} \operatorname{bd}(\Omega \times \Omega') = \overline{\Omega \times \Omega'} \setminus \operatorname{int}(\Omega \times \Omega') &\underbrace{=}_{\text{[21, Theorem 19.5]}} \big(\overline{\Omega} \times \overline{\Omega'}\big) \setminus \big(\Omega \times \Omega'\big) = \\ &= \big(\operatorname{bd}(\Omega) \times \overline{\Omega'}\big) \bigcup \big(\overline{\Omega} \times \operatorname{bd}(\Omega')\big) \end{aligned}$$

So we can assume without loss of generality that $(x, y) \in \operatorname{bd}(\Omega) \times \overline{\Omega'}$. From this follows $x_n \longrightarrow x \in \operatorname{bd}(\Omega)$. Since F is a function of Legendre type, we have

$$\|\nabla(F + G)(x_n, y_n)\|^2 = \underbrace{\|\nabla F(x_n)\|^2}_{\longrightarrow\infty} + \underbrace{\|\nabla G(y_n)\|^2}_{\geq 0} \longrightarrow \infty.$$

This implies that $\|\nabla(F + G)(x_n, y_n)\| \longrightarrow \infty$. ■

Due to the above result, we can extend our examples 3.2 and define higher-dimensional functions of Legendre type.

Example 3.4

function of Legendre type	$F : \Omega \to \mathbb{R}$
half the squared Euclidean norm	$\mathbb{R}^n \to \mathbb{R}, x \mapsto \frac{1}{2}\sum_{i=1}^{n} x_i^2 = \frac{1}{2}\|x\|^2$
convex Shannon entropy	$\mathbb{R}_+^n \to \mathbb{R}, x \mapsto \sum_{i=1}^{n} \big(x_i \cdot ln(x_i) - x_i\big)$
Burg entropy	$\mathbb{R}_+^n \to \mathbb{R}, x \mapsto \sum_{i=1}^{n} \big(1 - ln(x_i)\big)$

Remark 3.5 *In the original paper [22], Shannon defined the entropy of a discrete probability distribution as* $-\sum_{i=1}^{n} \big(x_i \cdot ln(x_i)\big)$. *Following [1] we turn this into a function of Legendre type and use the definition from 3.4.*

In the rest of this section, we follow [23] and show that the convex conjugate of a function of Legendre type is a function of Legendre type, and the gradient map is a homeomorphism onto its image. Before we do this, we need to recall some definitions from [23].

Definition 3.6

a) *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *be a function. We define the* domain *of* f *as the set* $\mathrm{dom}(f) := \{x \in \mathbb{R}^n \mid f(x) < \infty\}$ *where it has finite values.*

b) *We call a function* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ proper, *if* $\mathrm{dom}(f) \neq \emptyset$.

c) *A proper convex function* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *is called* closed, *if its lower semi-continuous, i.e.* $f(\lim_{n\to\infty} x_n) \leq \lim_{n\to\infty} f(x_n)$ *for all convergent sequences* $(x_n)_{n\in\mathbb{N}}$ *in* $\mathbb{R}^n$.

d) *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *be a convex function. We call the map*

$$\partial f : \mathbb{R}^n \to \mathcal{P}\big(\mathbb{R}^n\big), x \mapsto \big\{y \in \mathbb{R}^n \mid \forall z \in \mathbb{R}^n : f(z) \geq f(x) + \langle y, z - x\rangle\big\}$$

the subdifferential *of* f.

As one would expect, the subdifferential is in a strong relationship with the usual gradient.

Lemma 3.7 *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *be a proper convex function. Then we have the following:*

a) *For* $x \notin \mathrm{dom}(f)$, $\partial f(x)$ *is empty.*

b) *There exists at least one* $x \in \mathbb{R}^n$ *such that* $\partial f(x) \neq \emptyset$.

c) $\partial f(x)$ *is a nonempty bounded set if and only if* $x \in \text{int}(\text{dom}(f))$.

d) f *is differentiable in* $x \in \text{dom}(f)$ *if and only if* $\partial f(x)$ *contains exactly one element. In this case* $\partial f(x) = \{\nabla f(x)\}$.

Proof: For the proofs we refer to [23]. Part *a*) and *c*) is [23, Theorem 23.4]. Part *b*) follows from [23, Theorem 6.2] and [23, Theorem 23.4]. The last part is [23, Theorem 25.1]. ■

Definition 3.8 *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *be a closed proper convex function. We define the* convex conjugate *of* f *as the function*

$$f^* : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}, y \mapsto \sup_{x \in \mathbb{R}^n} \big(\langle y, x\rangle - f(x)\big).$$

Lemma 3.9 *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *be a closed proper convex function. Then the convex conjugate* f^* *is a closed proper convex function and it holds that* $f^{**} = f$.

Proof: We refer to [23, Theorem 12.2] for the proof. ■

Lemma 3.10 *For every closed proper convex function* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *and for every* $x, y \in \mathbb{R}^n$ *the following are equivalent:*

i) $x \in \partial f^*(y)$

ii) $f(x) + f^*(y) = \langle x, y\rangle$

iii) $y \in \partial f(x)$

Proof: [23, Theorem 23.5] ■

Definition 3.11 *A map* $\rho : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^n)$ *is called* multivalued mapping.

- *It is called* single-valued, *if* $\rho(x)$ *contains at most one element for each* $x \in \mathbb{R}^n$.
- *It is called* one-to-one, *if* ρ *and the inverse multivalued mapping* $\rho^{-1} : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^n), y \mapsto \{x \mid y \in \rho(x)\}$ *are both single-valued.*

Obviously, the subdifferential is a multivalued mapping. The next theorem presents a nice characterization of functions of Legendre type using the above terminology.

Theorem 3.12 *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *be a closed proper convex function. The subdifferential* ∂f *is one-to-one if and only if* $f : \text{int}(\text{dom}(f)) \to \mathbb{R}$ *is a function of Legendre-type.*

We split the proof into the following two lemmata:

Lemma 3.13 *For every closed proper convex function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ the following are equivalent:*

i) The subdifferential ∂f is single-valued.

(ii) The interior of $\operatorname{dom}(f)$ *is nonempty,* $f : \operatorname{int}(\operatorname{dom}(f)) \to \mathbb{R}$ *is differentiable and it satisfies the condition $(L3)$ in definition 3.1.*

In this case,

- $\partial f(x) = \{\nabla f(x)\}$ *for* $x \in \operatorname{int}\big(\operatorname{dom}(f)\big)$ *and*
- $\partial f(x) = \emptyset$ *for* $x \notin \operatorname{int}(\operatorname{dom}(f))$.

Proof: We recall the proof from [23, Theorem 26.1] to see why the condition $(L3)$ is important. We write C for $\operatorname{int}(\operatorname{dom}(f))$.

First we assume that ii) holds. By lemma 3.7, the subdifferential $\partial f(x)$ is empty for $x \notin \operatorname{dom}(f)$ and $\partial f(x) = \{\nabla f(x)\}$ for $x \in C$. So it is enough to consider the points in $\operatorname{dom}(f) \setminus C$.

By our assumption $C \neq \emptyset$, applying [23, Theorem 6.3] for the convex set $\operatorname{dom}(f)$ we have that $\overline{\operatorname{dom}(f)} = \overline{C}$. From this follows that every point in $\operatorname{dom}(f) \setminus C$ lies on the boundary of C.

Assume that there is a point $x \in \operatorname{dom}(f) \setminus C$ such that $\partial f(x) \neq \emptyset$. By [23, Theorem 25.6] there is a sequence $(x_n)_{n\in\mathbb{N}}$ in C such that $x_n \longrightarrow x$, f is differentiable in x_n for all $n \in \mathbb{N}$ and the limit of $\big(\nabla f(x_n)\big)_{n\in\mathbb{N}}$ is in $\partial f(x)$. By continuouty of the norm we have $\|\nabla f(x_n)\| \longrightarrow \|\lim_{n\to\infty} \nabla f(x_n)\| < \infty$. But this contradicts $(L3)$, thus we have $\partial f(x) = \emptyset$ for all $x \in \operatorname{dom}(f) \setminus C$. In conclusion, ∂f is a single valued map.

Conversely, let ∂f be single-valued. By Lemma 3.7 c), the subdifferential is not empty if $x \in C$. From this follows the differentiability of $f : C \to \mathbb{R}$.

To show that C is nonempty, we use lemma 3.7. By part b), there is at least one y such that $\partial f(y) \neq \emptyset$. Since ∂f is single-valued, $\partial f(y)$ has exactly one element. This implies differentiability of f in y. By [23, Corollary 25.1.1], every point, where f is differentiable, lies in C, so C is not empty.

To get a contradiction, we assume that $f : C \to \mathbb{R}$ does not satisfy $(L3)$, i.e. there is a convergent sequence $(x_n)_{n\in\mathbb{N}}$ in C such that $x := \lim_{n\to\infty} x_n$ lies on the boundary of C and $\big(\nabla F(x_n)\big)_{n\in\mathbb{N}}$ is bounded. This implies that $\big(\nabla F(x_n)\big)_{n\in\mathbb{N}}$ has a convergent subsequence.

To simplify the notation, we assume that $\big(\nabla F(x_n)\big)_{n\in\mathbb{N}}$ is convergent. By [23, Theorem 24.4] the limit of $\big(\nabla F(x_n)\big)_{n\in\mathbb{N}}$ lies in $\partial f(x)$. Together with our assumption on ∂f this implies differentiability in x. Using again [23, Corollary 25.1.1] we see that x lies in C, but this is a contradiction to $x \in \mathrm{bd}(C)$. ■

Lemma 3.14 *For every closed proper convex function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ the following are equivalent:*

i) The inverse subdifferential $(\partial f)^{-1}$ is single-valued.

(ii) f is strictly convex on every convex subset of $\{x \in \mathbb{R}^n \mid \partial f(x) \neq \emptyset\}$.

Proof: We refer to [23, Theorem 26.3] for the proof. ■

Proof of Theorem 3.12: If f satisfies $ii)$ from 3.13, then $\{x \in \mathbb{R}^n \mid \partial f(x) \neq \emptyset\} = \mathrm{int}(\mathrm{dom}(f))$. In this case the condition $ii)$ from 3.14 is equivalent to the strict convexity of f in $\mathrm{int}(\mathrm{dom}(f))$. With this in mind, the theorem follows immediately from 3.13 and 3.14. ■

Notation 3.15 *Let $F : \Omega \to \mathbb{R}$ be a function of Legendre type.*

a) We write Ω^* for the gradient space $\Omega^* := \{\nabla F(x) \mid x \in \Omega\}$.

b) We extend F to a closed proper convex function $\tilde{F} : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ by setting $\tilde{F}(x) := F(x)$ for $x \in \Omega$ and $\tilde{F}(x) := \infty$ for $x \notin \Omega$.

c) We write F^* for the restriction of the convex conjugate of $\tilde{F}$ to $\mathrm{int}(\mathrm{dom}(\tilde{F}^*))$.

Theorem 3.16 *Let $F : \Omega \to \mathbb{R}$ be a function of Legendre type, and let F^* and Ω^* be defined as above. Then we have:*

a) $\mathrm{int}(\mathrm{dom}(\tilde{F}^*)) = \Omega^*$

b) $F^* : \Omega^* \to \mathbb{R}$ is a function of Legendre type.

c) The gradient map $\nabla F : \Omega \to \Omega^*, x \mapsto \nabla F(x)$ is a homeomorphism with inverse map ∇F^*.

Proof: From theorem 3.12, it follows that the subdifferential $\partial\tilde{F}$ is one-to-one. By lemma 3.10, it holds that $(\partial\tilde{F})^{-1} = \partial\tilde{F}^*$. This implies that $\partial\tilde{F}^*$ is one-to-one as well, and $F^* : \mathrm{int}(\mathrm{dom}(\tilde{F}^*)) \to \mathbb{R}$ is a function of Legendre type.
Using 3.13, we have $\mathrm{int}(\mathrm{dom}(\tilde{F}^*)) = \{y \in \mathbb{R}^n \mid \partial\tilde{F}^*(y) \neq \emptyset\} = \Omega^*$.

In conclusion, the induced maps $\Omega \to \Omega^*, x \to \nabla F(x)$ and $\Omega^* \to \Omega, y \mapsto \nabla F^*(y)$ are inverses of each other. Convex differentiability implies continuous differentiability ([23, Theorem 25.5]), so ∇F and ∇F^* are continuous, i.e. ∇F is a homeomorphism. ■

Notation 3.17 *Let $F : \Omega \to \mathbb{R}$ be a function of Legendre type. If F is clear from the context we write x^* for $\nabla F(x)$.*

Theorem 3.18 *Let $F : \Omega \to \mathbb{R}$ be a function of Legendre type. The convex conjugate $F^* : \Omega^* \to \mathbb{R}$ is given by $F^*(x^*) = \langle x^*, x\rangle - F(x)$ for all $x^* \in \Omega^*$.*

Proof: By theorem 3.16, the gradient map $\nabla F : \Omega \to \Omega^*$ is a homeomorphism, so there exists a unique $x \in \Omega$ such that $\nabla F(x) = x^*$.

By definition 3.8, the convex conjugate of x^* is given by $\sup_{y\in\Omega} \big(\langle x^*, y\rangle - F(y)\big)$. Since $g : \Omega \to \mathbb{R}, y \mapsto \langle x^*, y\rangle - F(y)$ is a concave function, the supremum is reached at the unique point where its gradient is zero ([23, Theorem 27.1.e]). The gradient of g is given by $\nabla g(y) = x^* - \nabla F(y)$ for all $y \in \Omega$. Thus we have

$$F^*(x^*) = \sup_{y\in\Omega} \big(\langle x^*, y\rangle - F(y)\big) = \langle x^*, x\rangle - F(x).$$

■

Lemma 3.19 *Let $F : \Omega \to \mathbb{R}$, $G : \Omega' \to \mathbb{R}$ be functions of Legendre type. The convex conjugate of their sum is the sum of the convex conjugates i.e.*

$$(F + G)^* = F^* + G^*.$$

Proof: By proposition 3.3, $F+G : \Omega\times\Omega' \to \mathbb{R}, (x, y) \mapsto F(x)+G(y)$ is a function of Legendre type, and the partial derivatives of $F + G$ in $a = (a_1, \dots, a_{n+m}) \in \Omega \times \Omega'$ are given by

$$\frac{\partial F + G}{\partial x_i}(a) = \begin{cases} \frac{\partial F}{\partial x_i}(a_1, \dots, a_n) & \text{for } 1 \le i \le n \\ \frac{\partial G}{\partial x_i}(a_{n+1}, \dots, a_{n+m}) & \text{for } n + 1 \le i \le n + m \end{cases}$$

From this it follows that $\nabla(F+G)(x, y) = \big(\nabla F(x), \nabla G(y)\big)$ for all $(x, y) \in \Omega\times\Omega'$, so we have $\big(\Omega \times \Omega'\big)^* = \Omega^* \times \Omega'^*$.

From theorem 3.18 it follows immediately that

$$\begin{aligned}(F{+}G)^*(\nabla(F + G)(x, y)) &= \langle\nabla(F + G)(x, y), (x, y)\rangle - (F + G)(x, y) = \\ &= \langle\nabla F(x), x\rangle + \langle\nabla G(y), y\rangle - F(x) - G(y) = F^*(\nabla F(x)) + G^*(\nabla G(y)).\end{aligned}$$ ■

Using theorem 3.18 and lemma 3.19 one can easily compute the convex conjugate of the functions introduced in example 3.4.

$F:\Omega\to\mathbb{R}$	Ω	$\nabla F(x)$	Ω^*	$F^*(x^*)$
half the squared Euclidean norm	$\mathbb{R}^n$	x	$\mathbb{R}^n$	$\frac{1}{2}\|x^*\|^2$
convex Shannon entropy	$\mathbb{R}^n_+$	$\big(ln(x_1),\dots,ln(x_n)\big)$	$\mathbb{R}^n$	$\sum\limits_{i=1}^{n} e^{x_i^*}$
Burg entropy	$\mathbb{R}^n_+$	$\big(-\frac{1}{x_1},\dots,-\frac{1}{x_n}\big)$	$\mathbb{R}^n_-$	$\sum\limits_{i=1}^{n}\big(-2-ln(-x_i^*)\big)$

3.2. Bregman divergences

Bregman divergences measure the vertical distance between the graph of a function and a tangent plane of the graph. In his original work [18], Bregman used strictly convex differentiable functions, we follow [20],[1] and we use functions of Legendre type. This is a smaller class of functions, but it still contains prominent examples and it allows us to show better properties. We recall the results from [20] and [23], which are relevant for our later work.
In this section $F:\Omega\to\mathbb{R}$ denotes a function of Legendre type.

Definition 3.20 *The Bregman divergence associated with F is the map*

$$D_F:\Omega\times\Omega\to\mathbb{R},(x,y)\mapsto D_F(x\|y):=F(x)-(F(y)+\langle\nabla F(y),x-y\rangle).$$

The next propositon presents two alternative characterizations of the Bregman divergence.

Proposition 3.21

a) Let $y\in\Omega$, and $H_y:\Omega\to\mathbb{R},x\mapsto H_y(x):=F(y)+\langle\nabla F(y),x-y\rangle$ be the best linear approximation of F in y, then $D_F(x\|y)=F(x)-H_y(x)$ for all $x\in\Omega$.

b) For all $x,y\in\Omega$, it holds that $D_F(x\|y)=F(x)+F^*(y^*)-\langle y^*,x\rangle$.

Proof: Part *a*) is obvious. Part *b*) is a straightforward computation using Theorem 3.18, i.e. $F^*(y^*)=\langle y^*,y\rangle-F(y)$ and $y^*=\nabla F(y)$. ■

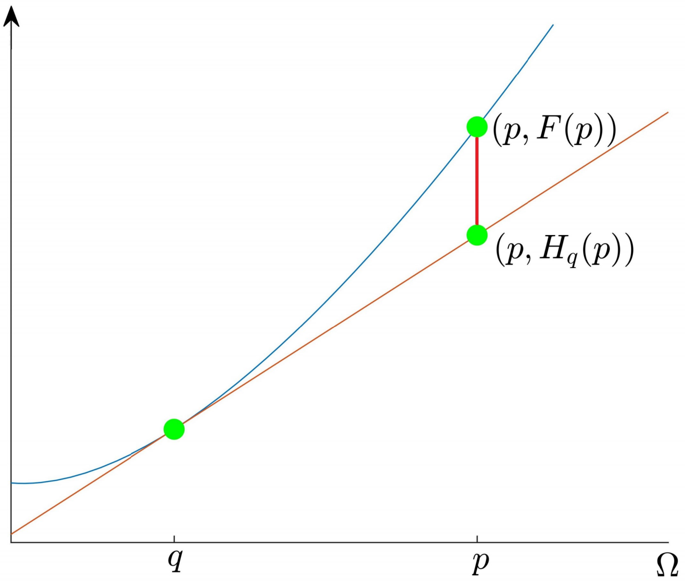


Figure 2: $D_F(p\|q)$ measures the vertical distance between $\big(p, F(p)\big)$ and $\big(p, H_q(p)\big)$

In general, a Bregman divergence does not satisfy the triangle inequality and it is not symmetric. If we want to talk about two explicit points $p, q \in \Omega$, we call $D_F(p\|q)$ the Bregman divergence from p to q. We only have the following metric axiom:

Lemma 3.22 *For all $x, y \in \Omega$, we have that $D_F(x\|y) \geq 0$, and $D_F(x\|y) = 0$ if and only if $x = y$.*

Proof: Since F is differentiable, the subdifferential of F in every point $z \in \Omega$ contains exactly one element, i.e. $\partial F(z) = \{\nabla F(z)\}$ (see lemma 3.7).

The claimed inequality is equivalent to $F(x) \geq F(y) + \langle \nabla F(y), x - y\rangle$. Since F is convex and $\nabla F(y) \in \partial F(y)$, this inequality is satisfied.

Using proposition 3.21, we see that the claimed equality is equivalent to

$$F(x) + F^*(y^*) = \langle y^*, x\rangle.$$

By lemma 3.10, this is satisfied if and only if

$$y^* \in \partial F(x) = \{\nabla F(x)\}.$$

Since ∇F is bijective (see theorem 3.16), this is equivalent to $y = x$. ■

It turns out that the characterization in proposition 3.21 b) is extremely useful to show some properties of D_F.

Lemma 3.23 *The function D_F is strictly convex in its first argument, and continuous in both arguments.*

Proof: The continuouty follows immediately from proposition 3.21 b), since F, F^*, ∇F (see theorem 3.16) and the scalar product are continuous functions.

To show the strict convexity, let us fix a point $y \in \Omega$. The function $D_F(\bullet \| y)$ is the sum of F and the affine linear function $\Omega \to \mathbb{R}, x \mapsto F^*(y^*) - \langle y^*, x \rangle$. Since F is strictly convex, one can easily show that $D_F(\bullet \| y)$ is strictly convex. ■

Informally speaking, one would expect that the 'distance' D_F between a point and 'us' is getting smaller and smaller, if 'we' are moving closer and closer to that point. The next lemma formalizes this.

Corollary 3.24 *Let $(y_n)_n$ be a sequence in Ω such that its limit y lies in Ω. Then $\big(D_F(y\|y_n)\big)_n$ and $\big(D_F(y_n\|y)\big)_n$ are converging to 0.*

Proof: By lemma 3.23, D_F is continuous in both arguments. ■

As we have already seen in theorem 3.16, the gradient map $\nabla F : \Omega \to \Omega^*$ is a homeomorphism. It also preserves the Bregman divergence, however it swaps the arguments.

Lemma 3.25 ***(Legendre duality)***
For all $x, y \in \Omega$, it holds that $D_F(x\|y) = D_{F^*}(y^*\|x^*)$.

Proof: Since $x^* = \nabla F(x)$ and $\nabla F^*(x^*) = x$ for every $x \in \Omega$, the result follows from proposition 3.21 b). ■

At the end of this section, we introduce some important examples.

Proposition 3.26 *Let $G : \Omega' \to \mathbb{R}$ a function of Legendre type. For all $x, y \in \Omega$, $a, b \in \Omega'$ it holds that $D_{F+G}\big((x,a)\|(y,b)\big) = D_F(x\|y) + D_G(a\|b)$.*

Proof: Using lemma 3.19 and the fact that $\nabla(F+G)(x,a) = \big(\nabla F(x), \nabla G(a)\big)$ for all $(x,a) \in \Omega \times \Omega'$, we have:

$$\begin{aligned} D_{F+G}\big((x,a)\|(y,b)\big) &= (F+G)(x,a) + (F+G)^*\big((y,b)^*\big) - \langle(\nabla(F+G)(y,b),(x,a)\rangle = \\ &= F(x) + G(a) + F^*(y^*) + G^*(b^*) - \langle y^*, x\rangle - \langle b^*, a\rangle = D_F(x\|y) + D_G(a\|b). \quad ■ \end{aligned}$$

In example 3.4 we have seen some functions of Legendre type. Using these functions we get the following Bregman divergences:

function of Legendre type	$F(x)$	Bregman divergence	$D_F(x\|y)$
half the squared Euclidean norm	$\frac{1}{2}\|x\|^2$	half the squared Euclidean distance	$\frac{1}{2}\|x-y\|^2$
convex Shannon entropy	$\sum_{i=1}^{n}\big(x_i ln(x_i) - x_i\big)$	Kullback-Leibler divergence	$\sum_{i=1}^{n}\big(x_i ln(\frac{x_i}{y_i}) - x_i + y_i\big)$
Burg entropy	$\sum_{i=1}^{n}\big(1 - ln(x_i)\big)$	Itakura-Saito divergence	$\sum_{i=1}^{n}\big(\frac{x_i}{y_i} - ln(\frac{x_i}{y_i}) - 1\big)$

3.3. Elements of Bregman geometry

In the previous section, we have introduced Bregman divergences. Using this 'distance' function, we can define basic geometrical objects analog to the usual Euclidean geometry.
During this section, $F : \Omega \to \mathbb{R}$ stands for a function of Legendre type.

Definition 3.27 *Let $x \in \Omega$ and $r \geq 0$.*
The primal Bregman ball with center x and radius r is defined as

$$B_F(x, r) := \{y \in \Omega \mid D_F(x\|y) \leq r\}.$$

The dual Bregman ball with center x and radius r is given by

$$B'_F(x, r) := \{y \in \Omega \mid D_F(y\|x) \leq r\}.$$

The primal Bregman sphere with center x and radius r is defined as

$$\partial B_F(x, r) := \{y \in \Omega \mid D_F(x\|y) = r\}.$$

The dual Bregman sphere with center x and radius r is given by

$$\partial B'_F(x, r) := \{y \in \Omega \mid D_F(y\|x) = r\}.$$

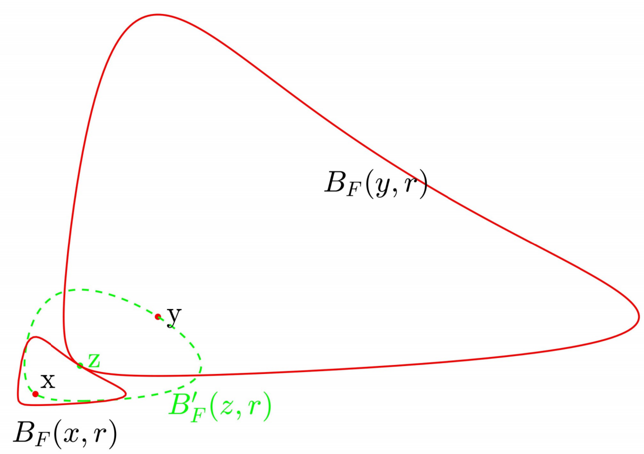


Figure 3: primal and dual Itakura-Saito balls

We follow the method described in [19] and give a geometrical approach to construct dual Bregman balls. First, we fix some notation.

Notation 3.28 *We write*

a) $\hat{\Omega}$ *for the Cartesian product* $\Omega \times \mathbb{R}$,

b) $\mathcal{F}$ *for the graph of* F, *i.e.* $\mathcal{F} := \{(x, F(x)) \in \hat{\Omega} \mid x \in \Omega\}$,

c) $\mathrm{epi}(F)$ *for the points in* $\hat{\Omega}$ *lying above the graph of* F, *i.e.* $\mathrm{epi}(F) := \{(x, z) \in \hat{\Omega} \mid z \geq F(x)\}$,

d) $Proj_{\Omega} : \hat{\Omega} \to \Omega, (x, z) \mapsto x$ *for the natural projection map.*

Definition 3.29

a) A set $\mathcal{H} \subset \mathbb{R}^m$ *is called a* hyperplane, *if there exists an* $\alpha \in \mathbb{R}^m \setminus \{0\}$ *and* $\beta \in \mathbb{R}$ *such that* $\mathcal{H} = \{v \in \mathbb{R}^m \mid \langle \alpha, v \rangle + \beta = 0\}$. *The vector* α *is said to be* normal *to the hyperplane* $\mathcal{H}$.

b) A hyperplane $\mathcal{H}$ *with normal vector* α *defines two* half-spaces $\mathcal{H}^+ := \{v \in \mathbb{R}^m \mid \langle \alpha, v \rangle + b \leq 0\}$ *and* $\mathcal{H}^- := \{v \in \mathbb{R}^m \mid \langle \alpha, v \rangle + b \geq 0\}$.

The choice of the superfixes "+" and "−" seems to be unnatural. The next remark gives an explanation why we should not choose the superfixes the other way around.

Remark 3.30 *Let $H : \mathbb{R}^n \to \mathbb{R}, x \mapsto \langle a, x\rangle + b$ be an affine linear map for some $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$. The graph of H is a hyperplane $\mathcal{H}$ in $\mathbb{R}^{n+1}$ with normal vector $\alpha = (a, -1)$ and $\beta = b$. The set of points above the graph is the half-space $\mathcal{H}^+$, i.e.*

$$\mathcal{H}^+ = \{(x, z) \in \mathbb{R}^{n+1} \mid z \geq H(x)\}.$$

The set of points under the graph H is the half-space $\mathcal{H}^-$, i.e.

$$\mathcal{H}^- = \{(x, z) \in \mathbb{R}^{n+1} \mid z \leq H(x)\}.$$

Lemma 3.31 *For every dual Bregman ball $B' \subset \Omega$ there is a hyperplane $\mathcal{H} \subset \mathbb{R}^{n+1}$ such that $Proj_\Omega(\mathcal{H}^- \cap \operatorname{epi}(F)) = B'$.*

Proof: Let $B' = B'_F(c, r)$ be a dual Bregman ball with center $c \in \Omega$ and radius $r \geq 0$. Recall that the Bregman divergence from a point $x \in \Omega$ to c is the difference between $F(x)$ and $H_c(x)$, where H_c is the best linear approximation of F in c, i.e. $H_c : \mathbb{R}^n \to \mathbb{R}, x \mapsto H_c(x) := F(c) + \langle \nabla F(c), x - c\rangle$.

Let $\mathcal{H}_{c,r}$ be the graph of $\mathbb{R}^n \to \mathbb{R}, x \mapsto H_c(x) + r$, i.e. $\mathcal{H}_{c,r}$ is a vertical translate of the graph of H_c with height r.

A point $x \in \Omega$ lies in $Proj_\Omega(\mathcal{H}^-_{c,r} \cap \operatorname{epi}(F))$ if and only if there is a $z \in \mathbb{R}$ such that $H_c(x) + r \geq z$ and $z \geq F(x)$. This is equivalent to $r \geq F(x) - H_c(x) = D_F(x\|c)$. ∎

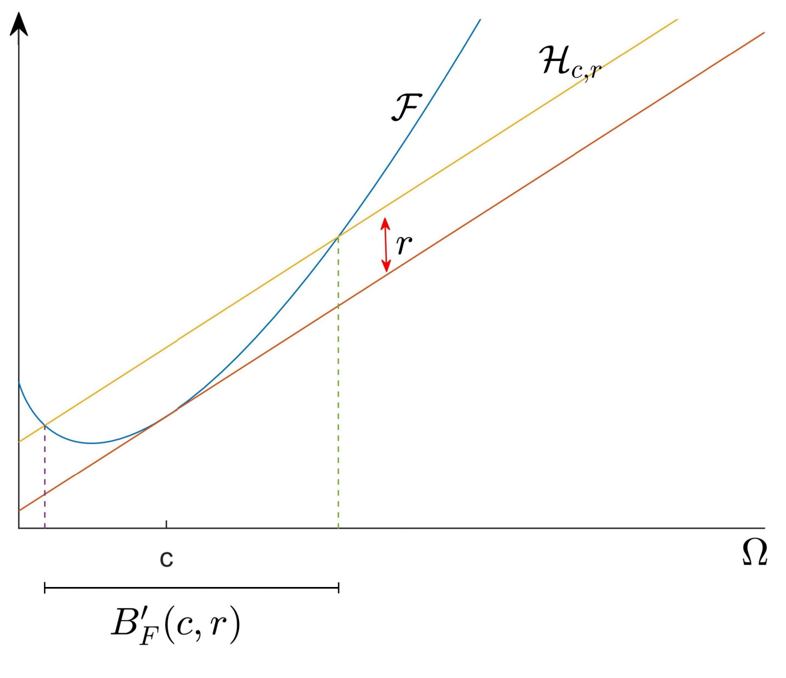


Figure 4: visualization of lemma 3.31

In general, primal Bregman balls are not convex. Fortunately, we have this nice property for dual Bregman balls.

Corollary 3.32 *Dual Bregman balls are convex.*

Proof: By lemma 3.31, every dual Bregman ball can be written as $Proj_\Omega(\mathcal{H}^- \cap \text{epi}(F))$ for a hyperplane $\mathcal{H} \subset \mathbb{R}^{n+1}$. The convexity of F implies that $\text{epi}(F)$ is a convex set. Since intersection of convex sets is convex, every dual Bregman ball is convex as the image of a convex set under a linear map. ■

In lemma 3.31, we have seen that every dual Bregman ball is the projection of $\text{epi}(F)$ intersected with a hyperplane. Conversely, one could ask which hyperplanes give us dual Bregman balls.

Lemma 3.33 *Let $a \in \mathbb{R}^n$, $b \in \mathbb{R}$, and let $\mathcal{H}$ be the graph of the affine linear map $\Omega \to \mathbb{R}, \langle a, x\rangle + b$. If $a \in \Omega^*$, then $Proj_\Omega(\mathcal{H}^- \cap \text{epi}(F))$ is a dual Bregman ball with center $\nabla F^*(a)$ and radius $\langle a, \nabla F^*(a)\rangle - F(\nabla F^*(a)) + b$.*

Proof:
Set $c := \nabla F^*(a)$. Using $\nabla F(\nabla F^*(a)) = a$ (theorem 3.16), we have that

$$\langle a, x\rangle + b = \langle \nabla F(c), x - c\rangle + F(c) + \langle \nabla F(c), c\rangle - F(c) + b$$

Set $r := \langle \nabla F(c), c\rangle - F(c) + b$. A similar argument as in the proof of lemma 3.31 shows that x lies in $Proj_\Omega(\mathcal{H}^- \cap \text{epi}(F))$ if and only if $r \geq D_F(x\|c)$.

So $Proj_\Omega(\mathcal{H}^- \cap \text{epi}(F))$ is a dual Bregman ball with center $\nabla F^*(a)$ and radius $\langle a, \nabla F^*(a)\rangle - F(\nabla F^*(a)) + b$. ■

As we have already mentioned before, using Legendre duality, one can translate primal to dual objects and vice versa.

Proposition 3.34 *For all $x \in \Omega$ and $r \geq 0$ we have that*

$$\nabla F\big(B_F(x,r)\big) = B'_{F^*}(x^*, r) \qquad \textit{and} \qquad \nabla F\big(B'_F(x,r)\big) = B_{F^*}(x^*, r).$$

Proof: It follows immediately from Legendre duality 3.25. ■

Lemma 3.35 *Primal und dual Bregman balls are compact.*

Proof: We start with dual Bregman balls. Let $B'_F(c,r)$ be the dual Bregman ball with center $c \in \Omega$ and radius $r \geq 0$. Because of the Heine–Borel theorem, it is enough to show that $B'_F(c,r)$ is closed and bounded.

Consider the function $D_F(\bullet\|c) : \Omega \to \mathbb{R}, x \mapsto D_F(x\|c)$. By lemma 3.23, this function is continuous and strictly convex.

Since $B'_F(c, r)$ is the preimage of the closed interval $[0, r]$ under $D_F(\bullet\|c)$, it is closed.

By proposition 3.21 b), we have $D_F(x\|c) = F(x) + F^*(c^*) - \langle c^*, x\rangle$ for all $x \in \Omega$. So $B'_F(c, r)$ is bounded if and only if $\{x \in \Omega \mid F(x) - \langle c^*, x\rangle \leq r\}$ is bounded. This set is a sublevel set of $f : \Omega \to \mathbb{R}, x \mapsto F(x) - \langle c^*, x\rangle$.

We may extend the functions F and f to $\mathbb{R}^n$ by setting $F(x) := \infty$, $f(x) := \infty$ for $x \notin \Omega$. One can easily see that $f : \mathbb{R}^n \to \mathbb{R}$ is a closed proper convex function. If we can show that $0 \in \operatorname{int}(\operatorname{dom}(f^*))$, then from [23, Corollary 14.2.2] will follow that the set $\{x \in \Omega \mid f(x) \leq r\}$ is bounded.

By definition 3.8, the convex conjugate is given by

$$f^*(z) = \sup_{x\in\mathbb{R}^n} \big(\langle z, x\rangle - F(x) + \langle c^*, x\rangle\big) = \sup_{x\in\mathbb{R}^n} \big(\langle z + c^*, x\rangle - F(x)\big) = F^*(z + c^*)$$

for all $z \in \mathbb{R}^n$. This implies that $\operatorname{int}(\operatorname{dom}(f^*)) = \{x^* - c^* \mid x \in \operatorname{int}(\operatorname{dom}(F^*))\}$. Since $c^* \in \Omega^* = \operatorname{int}(\operatorname{dom}(F^*))$, 0 lies in $\operatorname{int}(\operatorname{dom}(f^*))$. From this follows that the dual Bregman ball $B'_F(c, r)$ is bounded.

By proposition 3.34, we can write every primal Bregman ball with center $c \in \Omega$ and radius $r \geq 0$ as $B_F(c, r) = \nabla F^*\big(B'_{F^*}(c^*, r)\big)$. We have already shown above that dual Bregman balls are compact. Since ∇F^* is a homeomorphism, the primal Bregman ball $B_F(c, r)$ is compact. ∎

Before we introduce other objects of Bregman geometry, we show a property of primal Bregman balls, which will play an important role in the data analysis part.

Lemma 3.36 *Intersections of primal Bregman balls are either empty or contractible.*

Proof: Let $\big\{B_F(c, r_c)\big\}_{c\in C}$ be a collection of primal Bregman balls. Since ∇F is a homeomorphism, $\bigcap_{c\in C} B_F(c, r_c)$ is contractible if and only if
$\nabla F\Big(\bigcap_{c\in C} B_F(c, r_c)\Big) = \bigcap_{c\in C} B'_{F^*}(c^*, r_c)$ is contractible.

By corollary 3.32, dual Bregman balls are convex, so the intersection is also convex. It is a well-known result in algebraic topology that nonempty convex sets are contractible. ∎

In the rest of this section we follow [19] and introduce some basic objects in Bregman geometry.

Definition 3.37 *Let x and y be points in Ω. The primal Bregman bisector of x and y is defined as*

$$BB_F(x,y) := \{z \in \Omega \mid D_F(x\|z) = D_F(y\|z)\}.$$

The dual Bregman bisector of x and y is defined as

$$BB'_F(x,y) := \{z \in \Omega \mid D_F(z\|x) = D_F(z\|y)\}.$$

Proposition 3.38 *A dual Bregman bisector $BB'_F(x,y) \subset \Omega$ is a hyperplane, if and only if $x \neq y$. In this case, x and y lie on different sides of the hyperplane $BB'_F(x,y)$.*

Proof: A point $z \in \Omega$ lies on $BB'_F(x,y) \in \Omega$ if and only if $D_F(z\|x) = D_F(z\|y)$. Using the characterization from proposition 3.21, a straightforward computation shows that this is equivalent to $\langle y^* - x^*, z\rangle + F^*(x^*) - F^*(y^*) = 0$. This is an equation of a hyperplane if and only if $x^* \neq y^*$ or equivalently $x \neq y$.

The points x and y lie on different sides since
$\langle y^* - x^*, x\rangle + F^*(x^*) - F^*(y^*) = D_F(x\|x) - D_F(x\|y) < 0$ and
$\langle y^* - x^*, y\rangle + F^*(x^*) - F^*(y^*) = D_F(y\|x) - D_F(y\|y) > 0$. ■

Primal Bregman bisectors do not have to be hyperplanes, however we can use Legendre duality again.

Proposition 3.39 *For all $x, y \in \Omega$, we have that*

$$\nabla F\big(BB_F(x,y)\big) = BB'_{F^*}(x^*,y^*) \qquad \text{and} \qquad \nabla F\big(BB'_F(x,y)\big) = BB_{F^*}(x^*,y^*).$$

Proof: It follows from Legendre duality 3.25. ■

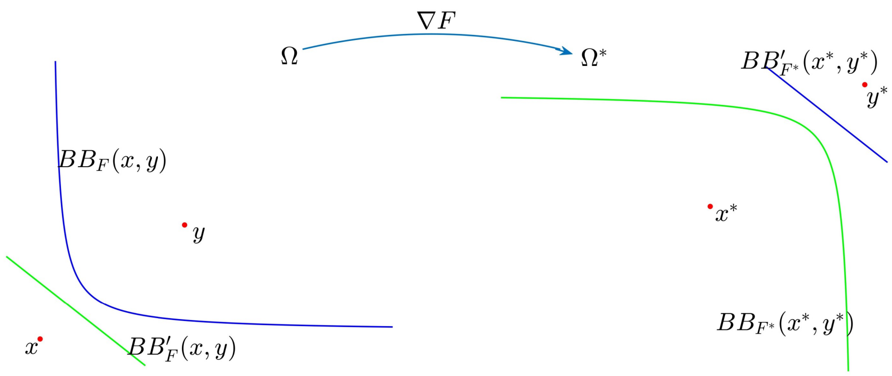


Figure 5: primal and dual Bregman bisectors

One can see the next property as a weak version of the triangle inequality. Using this result one can compute which triplet of points violates the triangle inequality.

Proposition 3.40 ***(Three-point property)*** *For all $x, y, z \in \Omega$ it holds that*

$$D_F(x\|z) + \langle z^* - y^*, x - y\rangle = D_F(x\|y) + D_F(y\|z).$$

Proof: The proof is a straightforward computation using definition 3.20 and the linearity of the scalar product. ■

Proposition 3.41 *Let $A \subset \Omega$ be a nonempty closed convex subset. For every $y \in \Omega$, there exists a unique point $x \in A$ that minimizes the Bregman divergence from x to y. This point is denoted by y_A and called the* <u>*Bregman projection*</u> *of y onto A.*

Proof: Consider the function $f := D_F(\bullet\|y) : A \to \mathbb{R}, a \mapsto D_F(a\|y)$. This is strictly convex and continuous by lemma 3.23.

Pick a point $a \in A$ and consider the dual Bregman ball with center y and radius $r := D_F(a\|y)$. By lemma 3.35, this dual Bregman ball is closed and bounded, from this follows that $A \cap B'_F(y, r)$ is closed and bounded, i.e. it is a compact set.

Since f is continuous, $f \restriction_{A \cap B'_F(y,r)}$ has a minimum point. We denote this minimum point by y_A. For every $x \in A \setminus B'_F(y, r)$ we have $D_F(x\|y) > r \geq D_F(y_A\|y)$, so y_A is a minimum point of f.

The uniqueness of y_A follows by a standard argument of convex analysis. We recall it for completeness.

Let $z \in A$ another minimum point sucht that $z \neq y_A$. Since A is convex, $\frac{y_A+z}{2}$ lies an A. The strict convexity of f implies that $f\left(\frac{y_A+z}{2}\right) < \frac{1}{2}f(y_A) + \frac{1}{2}f(z) = f(y_A)$ but this contradicts to the minimality of $f(y_A)$. ■

The Bregman Pythagoras inequality is a well-known result in Bregman geometry. Several slightly different proofs are known in the literature e.g. [19, Property 6], [24, Proposition 1], [20, Proposition 3.16]. We recall the proof from [24, Proposition 1] for the sake of completeness.

Lemma 3.42 ***(Bregman Pythagoras inequality)*** *Let $A \subset \Omega$ be a closed convex subset, $y \in \Omega$ and y_A its Bregman projection onto A. For every $a \in A$ we have that*

$$D_F(a\|y) \geq D_F(a\|y_A) + D_F(y_A\|y)$$

with equality if A is an affine subspace.

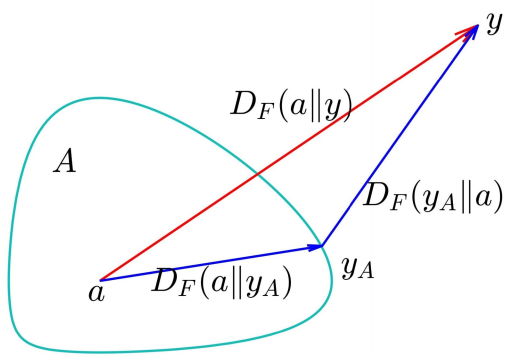


Figure 6: Bregman Pythagoras inequality

Proof: By the three-point property (proposition 3.40), it is enough to show that

$$\langle \nabla F(y) - \nabla F(y_A), a - y_A \rangle \leq 0 \tag{1}$$

holds for all $a \in A$ with equality if A is an affine subspace.

First, we assume that $F(y) = 0$ and F has its minimum at $y \in \Omega$. In this case, we have that

$$\nabla F(y) = 0 \qquad \text{and} \qquad D_F(x\|y) = F(x)$$

for every $x \in \Omega$. Thus, y_A is the minimum point of $A \to \mathbb{R}, a \mapsto F(a)$.
By [23, Theorem 27.4], $-\nabla F(y_A)$ is normal to A at y_A i.e.

$$\langle a - y_A, -\nabla F(y_A) \rangle \leq 0$$

for every $a \in A$. This implies (1), since $\nabla F(y) = 0$. In the case where A is an affine subspace, from [23, Theorem 27.4] it follows that $\nabla F(y_A)$ is orthogonal to A. Thus, it holds equality in (1).

For an arbitrary F we consider

$$G : \Omega \to \mathbb{R}, x \to G(x) := F(x) - (F(y) + \langle \nabla F(y), x - y \rangle).$$

It is easy to check that

- $G(y) = 0$,
- $\nabla G(y) = 0$,
- $D_G(v\|w) = D_F(v\|w)$ for every $v, w \in \Omega$.

Thus, we can apply the above argument for G and establish (1) for F. ■

Definition 3.43

a) *For $x, y, z \in \Omega$, the triplet (x, y, z) is said to be* <u>*Bregman orthogonal*</u> *in Ω, if* $D_F(x\|y) + D_F(y\|z) = D_F(x\|z)$.

b) *We say that $X \subset \Omega$ is Bregman orthogonal to $Z \subset \Omega$ if for every $x \in X$ and $z \in Z$ there exists a $y \in X \cap Z$ such that (x, y, z) is Bregman orthogonal.*

Proposition 3.44 *A triplet (x, y, z) of points in Ω is Bregman orthogonal if and only if (z^*, y^*, x^*) is Bregman orthogonal in Ω^*.*

Proof: Using the three-point property (proposition 3.40) we see that (x, y, z) is Bregman orthogonal if and only if $\langle z^* - y^*, x - y\rangle = 0$ or equivalently $\langle x - y, z^* - y^*\rangle = 0$. But this is equivalent to say that the triplet (z^*, y^*, x^*) is Bregman orthogonal in Ω^*. ■

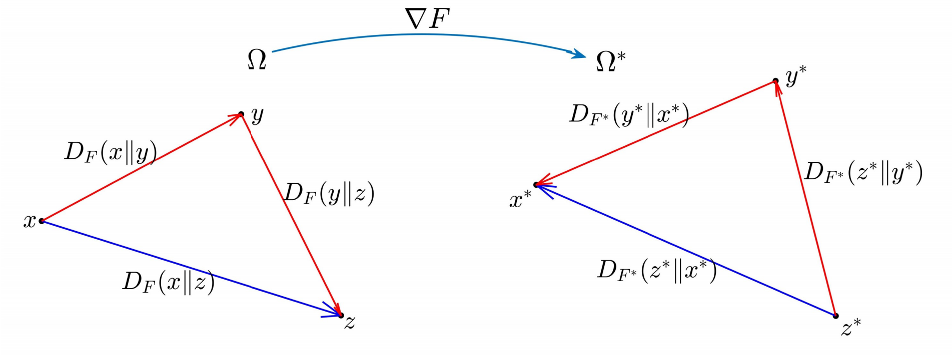


Figure 7: Bregman orthogonal points in Ω and in Ω^*

Remark 3.45 *Bregman orthogonality is not symmetric. To present this behaviour, we give a counterexample using the Itakura-Saito divergence.*
Let $x := (2, 2)$,$y := (1, 1)$ and $z := (5, \frac{5}{9})$. An easy computation shows that

$$D_F(x\|y) + D_F(y\|z) = 1.6354 = D_F(x\|z)$$
$$D_F(z\|y) + D_F(y\|x) = 2.9202 \neq 2.5339 = D_F(z\|x)$$

Thus, $X := \{x, y\}$ is Bregman orthogonal to $Z := \{y, z\}$ but Z is not Bregman orthogonal to X.

Notation 3.46 *Let x, y be in Ω. We write*

a) $\Gamma_F(x, y) := \{z \in \Omega \mid z^* = (1-t)x^* + ty^*, t \in [0, 1]\}$ *for the geodesic arc joining x to y,*

b) $\Lambda_F(x, y) := \{z \in \Omega \mid z = (1-t)x + ty, t \in [0, 1]\}$ *for the line segment between x and y.*

Proposition 3.47 *For all* $x, y \in \Omega$ *we have* $\nabla F\big(\Gamma_F(x,y)\big) = \Lambda_{F^*}(x^*, y^*)$ *and* $\nabla F\big(\Lambda_F(x,y)\big) = \Gamma_{F^*}(x^*, y^*)$

Proof: It follows immediately from the above definition using that $x^{**} = x$ for all $x \in \Omega$. ■

Lemma 3.48 *For all points* x, y *in* Ω,

a) the dual Bregman bisector $BB'_F(x,y)$ *is Bregman orthogonal to* $\Gamma(x,y)$,

b) and the line segment $\Lambda(x,y)$ *is Bregman orthogonal to the primal Bregman bisector* $BB_F(x,y)$.

Proof: A complete proof is given in [19, Lemma 7], we recall it for completness.

For $x = y$ is the lemma trivial. So we assume that $x \neq y$. In this case $BB'_F(x,y)$ is a hyperplane and x, y lies on different sides of it (see proposition 3.38). Using the duality result from propisiton 3.47 we see that $\Gamma(x,y)$ is a continuous path between x and y. This implies that the intersection $\Gamma(x,y) \bigcap BB'_F(x,y)$ is not empty.

Consider points $a \in BB'_F(x,y)$, $c \in \Gamma(x,y)$, $b \in \Gamma(x,y) \bigcap BB'_F(x,y)$ and show that (a,b,c) is Bregman othongonal. By the three point property 3.40, it is enough to show that $\langle c^* - b^*, a - b\rangle = 0$.

Both b and c are in $\Gamma(x,y)$, so there are some $s, t \in [0,1]$ such that

$$b^* = (1-s)x^* + sy^* \qquad \text{and} \qquad c^* = (1-t)x^* + ty^*$$

From this follows that

$$c^* - b^* = (s-t)(x^* - y^*).$$

If $s = t$ then $b = c$, and the triplet (a,c,c) is Bregman orthogonal, since $D_F(c\|c) = 0$. In the following, we assume that $s \neq t$, this allows us to divide by $(s-t)$. We have $x^* - y^* = \frac{1}{s-t}(c^* - b^*)$.

Since $a, b \in BB'_F(x,y)$, we have

$$-F(x) - \langle x^*, a-x\rangle = -F(y) - \langle y^*, a-y\rangle \text{ and } -F(x) - \langle x^*, b-x\rangle = -F(y) - \langle y^*, b-y\rangle.$$

Subtracting the first equality from the second, we get

$$\langle x^*, a-x\rangle - \langle x^*, b-x\rangle = \langle y^*, a-y\rangle - \langle y^*, b-y\rangle.$$

This is equivalent to $\langle x^* - y^*, a-b\rangle = 0$. After substituting $x^* - y^* = \frac{1}{s-t}(c^* - b^*)$ we have $\frac{1}{s-t}\langle c^* - b^*, a-b\rangle = 0$.

This shows part a). The second part follows using duality. Let $v \in \Lambda(x,y)$ and $w \in BB_F(x,y)$. By proposition 3.39 and 3.47, we have $\nabla F\big(BB_F(x,y)\big) = BB'_{F^*}(x^*,y^*)$ and $\nabla F\big(\Lambda(x,y)\big) = \Gamma_{F^*}(x^*,y^*)$.

By part a) there is a $z^* \in BB'_{F^*}(x^*,y^*) \bigcap \Gamma_{F^*}(x^*,y^*)$ such that (w^*,z^*,v^*) is Bregman orthogonal. From propositon 3.44 it follows that (v,z,w) is a Bregman orthogonal triplet. Since ∇F is a homeomorphism it is clear that $z \in \Lambda(x,y) \bigcap BB_F(x,y)$. ■

3.4. Smallest including dual balls

In rest of this chapter, we will study geometrical objects, which will play an essential role in the next chapter where we will build simplicial complexes. We start with the *smallest including dual ball* and its center.
In this section $F : \Omega \to \mathbb{R}$ denotes a function of Legendre type.

Definition 3.49 *Let $P \subset \Omega$ be a finite subset. We call a dual Bregman ball $B' \subset \Omega$ including dual Bregman ball of P if $P \subset B'$.*

Proposition 3.50 *Let $P \subset \Omega$ a finite subset and $B'_F(c,r)$ a dual Bregman ball with center $c \in \Omega$ and radius $r \geq 0$.
The center c lies in the intersection $\bigcap_{p \in P} B_F(p,r)$ if and only if $B'_F(c,r)$ is an including dual Bregman ball of P.*

Proof: The statement follows from definition 3.27 and 3.49:

$$c \in \bigcap_{p \in P} B_F(p,r) \Leftrightarrow \forall p \in P : D_F(p\|c) \leq r \Leftrightarrow P \subset B'_F(c,r) \qquad ■$$

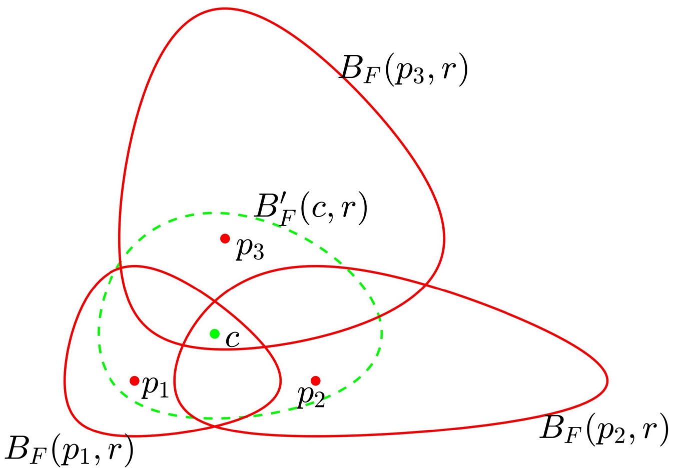


Figure 8: intersection of primal Bregman balls

Definition 3.51 *The smallest including dual ball of a finite subset $P \subset \Omega$ is an including dual Bregman ball of P such that there is no other including dual Bregman ball of P with smaller radius.*

In definition 3.51, we were shameless and talked about 'the' smallest including dual ball without knowing the uniqueness and the existence of it. The next proposition fills these gaps.

Proposition 3.52 *Let $P \subset \Omega$ a finite subset. The smallest including dual ball of P exists and it is unique.*

Proof: For each fixed $c_0 \in \Omega$ the smallest radius r such that $B'_F(c_0, r)$ includes P is given by $\max_{p\in P} D_F(p\|c_0)$. So in order to find the smallest including dual ball of P we have to find a minimum point of the function $f : \Omega \to \mathbb{R}, x \mapsto \max_{p\in P} D_F(p\|x)$.

Using the fact that maximum of continuous function is continuous and lemma 3.23 we see that f is a continuous function. We want to use that continuous functions on compact sets have a minimum point. Unfortunately, Ω is not compact, so we have to find an other compact set.

Pick an arbitrary point $q \in \Omega$. Using 3.35, we can see that the set $K := \bigcap_{p\in P} B_F(p, f(q))$ is compact, so there is a $c \in K$ such that $f\restriction_K$ has a minimum at c.

From proposition 3.50 it follows that every point $x \in \Omega$ with $f(x) \leq f(q)$ lies in $\bigcap_{p\in P} B_F(p, f(x)) \subset \bigcap_{p\in P} B_F(p, f(q)) = K$. This implies that c is a minimum point of f.

Now, we show the uniqueness. Assume there are two smallest including dual balls. Let $c, d \in \Omega$ be their centers and let $r \geq 0$ be their radius. Using the duality property (lemma 3.25) and using that Bregman divergences are strictly convex in the first argument (lemma 3.23) we have for all $p \in P$ that

$$D_F\Big(p\Big\|\Big(\frac{c^*+d^*}{2}\Big)^*\Big) = D_{F^*}\Big(\frac{c^*+d^*}{2}\Big\|p^*\Big) < \frac{1}{2}D_{F^*}(c^*\|p^*) + \frac{1}{2}D_{F^*}(d^*\|p^*) =$$
$$= \frac{1}{2}D_F(p\|c) + \frac{1}{2}D_F(p\|d) \leq \frac{1}{2}r + \frac{1}{2}r = r.$$

But this contradicts to the minimality of r. ■

Notation 3.53 *Let $P \subset \Omega$ a finite subset. We write $c_F(P)$ for the center of the smallest including dual ball of P and $\varrho_F(P)$ for its radius. We follow [24] and call $c_F(P)$ the Chernoff point of P.*

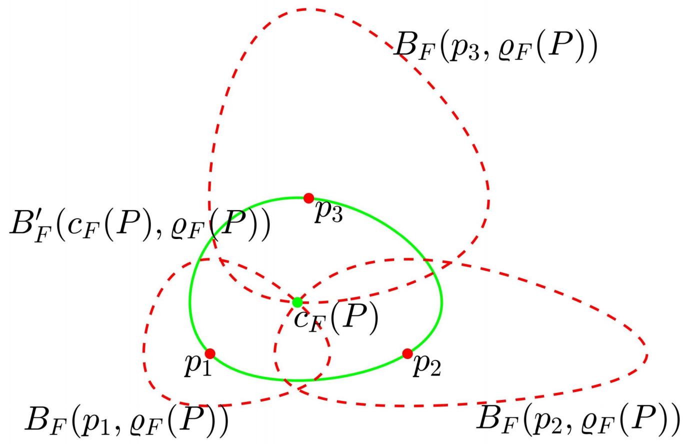


Figure 9: smallest including dual ball of a finite set $P = \{p_1, p_2, p_3\}$

In [25] and [26], Nielsen and Nock presented an algorithm to compute the center of the smallest including *primal* Bregman ball of a finite set. Using Legendre duality 3.25, we can apply this algorithm to compute the Chernoff point of a finite set. However, it will be advantageous to have some theoretical properties about Chernoff points. In [24], Edelsbrunner et al. studied the location of Chernoff points. Now we recall their result.

Lemma 3.54
The Chernoff point of a finite subset $P \subset \Omega$ lies in the convex hull of P.

Proof: Let $c := c_F(P)$ be the Chernoff point and $r := \varrho_F(P)$ the radius of the smallest including dual ball of P.

We write A for the complex hull of P. This is a convex closed subset of Ω, so we can apply the Bregman Pythagoras inequality 3.42.

Assume that c is not in A. In this case, the Bregman projection of c onto A and c must be different points, this implies that $D_F(c_A \| c) > 0$. Using this and the Bregman Pythagoras inequality 3.42 we have that

$$D_F(p \| c_A) < D_F(p \| c_A) + D_F(c_A \| c) \leq D_F(p \| c) \leq r$$

for all $p \in P$. Thus, there is an including dual Bregman ball of P with center c_A and radius less than r, but this is a contradiction to the minimality of r. ■

At the end of this section, we present a result about the behaviour of smallest including dual balls with Cartesian products.

Lemma 3.55 *Let $G : \Omega' \to \mathbb{R}$ be an other function of Legendre type. Let $P \subset \Omega$ and $Q \subset \Omega'$ finite subsets. We have the following relations for the radii of smallest including dual balls:*

$$\varrho_{F+G}(P \times Q) \leq \varrho_F(P) + \varrho_G(Q) \leq 2\varrho_{F+G}(P \times Q)$$

Proof: In order to show the first inequality, we show that the dual Bregman ball centered at $(w_1, w_2) := (c_F(P), c_G(Q))$ with radius $\varrho_F(P) + \varrho_G(Q)$ includes $P \times Q$.

By proposition 3.26, it holds for every $(p, q) \in P \times Q$ that

$$D_{F+G}\big((p, q)\|(w_1, w_2)\big) = D_F(p\|w_1) + D_G(q, w_2) \leq \varrho_F(P) + \varrho_G(Q)$$

The first inequality follows from the minimality of $\varrho_{F+G}(P \times Q)$.

For the second inequality it is enough to show that $\varrho_F(P) \leq \varrho_{F+G}(P \times Q)$. Write (c_1, c_2) for $c_{F+G}(P \times Q) \in \Omega \times \Omega'$. Since the Bregman divergence is always non-negative, for any $q \in Q$ we have:

$$D_F(p\|c_1) \leq D_F(p\|c_1) + D_G(q\|c_2) = D_{F+G}\big((p, q)\|(c_1, c_2)\big) \leq \varrho_{F+G}(P \times Q)$$

Thus, the dual Bregman ball centered at c_1 with radius $\varrho_{F+G}(P \times Q)$ includes P. This implies $\varrho_F(P) \leq \varrho_{F+G}(P \times Q)$. ■

3.5. Circumballs

In chapter 4, we want to compute the smallest including dual ball of each simplex of a simplicial complex. As we have already mentioned, using the algorithm from [25], [26] and Legendre duality one can compute the smallest including dual balls. However, if we had to call this algorithm for each simplex, it would cost too much time.

In [1], Edelsbrunner and Wagner presented a somewhat more efficient algorithm using *circumballs* and *discrete Morse theory*. In this section, we discuss how to compute *circumballs*. Here again, $F : \Omega \to \mathbb{R}$ denotes a function of Legendre type.

Definition 3.56 *Let $P \subset \Omega$ be a finite subset, $c \in \Omega$ and $r \geq 0$. We say*

a) the primal Bregman ball $B_F(c, r)$ is a primal circumball of P, if $P \subset \partial B_F(r, c)$.

b) the dual Bregman ball $B'_F(c, r)$ is a dual circumball of P, if $P \subset \partial B'_F(r, c)$.

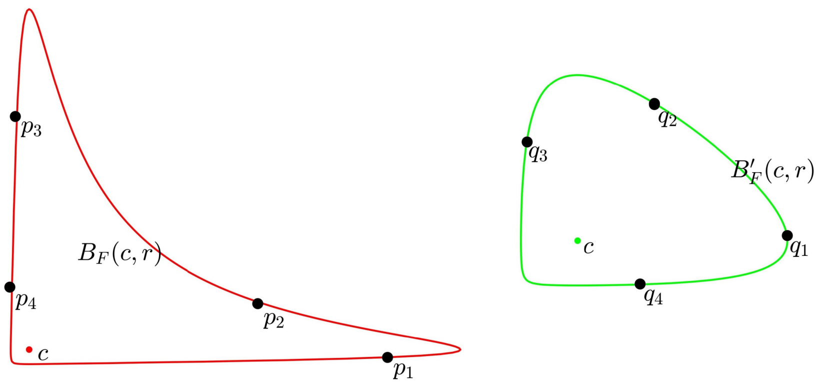


Figure 10: primal and dual circumballs

Proposition 3.57 *Let $P \subset \Omega$ a finite subset, $c \in \Omega$ and $r \geq 0$.*

a) The point c is a center of a primal circumball of P if and only if $c \in \bigcap_{p,q \in P} BB'(p,q)$.

b) The point c is a center of a dual circumball of P if and only if $c \in \bigcap_{p,q \in P} BB(p,q)$.

Proof: By definition c is a center of a primal circumball of P if and only if $D_F(c\|p) = D_F(c\|q)$ for all $p, q \in P$. This is exactly the same as saying that $c \in BB'_F(p,q)$ for all $p, q \in P$. Part b) follows by a similar argument. ■

By proposition 3.38, dual Bregman bisectors are hyperplanes, so to determine whether a set has a primal circumball, it is enough to compute the intersection of some hyperplanes, which can be easily done using linear algebra techniques.

Lemma 3.58 *Let $P = \{p_0, \ldots, p_k\} \subset \Omega$ be a finite subset. A point $c \in \Omega$ is a center of a primal circumball of P if and only if c is a solution of the linear system:*

$$\langle p_1^* - p_0^*, x \rangle + F^*(p_0^*) - F^*(p_1^*) = 0$$

$$\vdots$$

$$\langle p_k^* - p_0^*, x \rangle + F^*(p_0^*) - F^*(p_k^*) = 0$$

Proof: By proposition 3.38, the hyperplane $BB'(p,q)$ is given by the equation

$$\langle q^* - p^*, x \rangle + F^*(p^*) - F^*(q^*) = 0$$

for all $p, q \in \Omega$.

Fixing $p_0 \in P$, we have $\bigcap_{p,q \in P} BB'(p,q) = \bigcap_{p \in P} BB'(p_0, p)$. Now, the result follows immediately from proposition 3.57. ■

Note that not every solution of the above linear system is a center of a primal circumball. In general, there is no guarantee that a solution lies in Ω. Using duality we can translate the above result to dual circumballs.

Proposition 3.59 *Let $P \subset \Omega$ be a finite subset, $c \in \Omega$ and $r \geq 0$. Then $B'_F(c, r)$ is a dual cirumball of P if and only if $B_{F^*}(c^*, r)$ is a primal cirumball of $P^* := \{p^* \mid p \in P\}$.*

Proof: By lemma 3.25, we have $D_F(p\|c) = D_{F^*}(c^*\|p^*)$ for all $p \in P$. The result follows using definition 3.56. ■

Lemma 3.60 *Let $P = \{p_0, \dots, p_k\} \subset \Omega$ be a finite subset, $c \in \Omega$ and $r \geq 0$. There is a dual circumball of P with center c if and only if $c^* \in \Omega^*$ is a solution of the linear system:*

$$\langle p_1 - p_0, x\rangle + F(p_0) - F(p_1) = 0$$

$$\vdots$$

$$\langle p_k - p_0, x\rangle + F(p_0) - F(p_k) = 0$$

Proof: By proposition 3.59, there exists a dual circumball of P with center c if and only if there is a primal circumball of P^* with center c^*.

Using $F = F^{**}$, $p = p^{**}$ and lemma 3.58, it follows that P^* has a primal circumball with center c^* if and only of c^* is a solution of the above linear equation system. ■

The solvability of the linear system does not imply the existence of a dual circumball, since a solution does not have to be in Ω^*. However, if the linear system is not solvable, there is no dual circumball. So the existence of circumballs is not guaranteed. It is not a surprise, since even in the usual Euclidean geometry not every set of points has a circumball.

Example 3.61 *To represent this phenomenon, let us consider the points $p_0 := (0, 0), p_1 := (1, 0), p_2 := (-1, 0)$ in $\mathbb{R}^2$ and let F be the half the squared Euclidean norm. The linear system*

$$\langle (1, 0), x\rangle + 0 - \frac{1}{2} = 0$$

$$\langle (-1, 0), x\rangle + 0 - \frac{1}{2} = 0$$

has no solution. This implies that there exists no circumball of $\{p_0, p_1, p_2\}$.

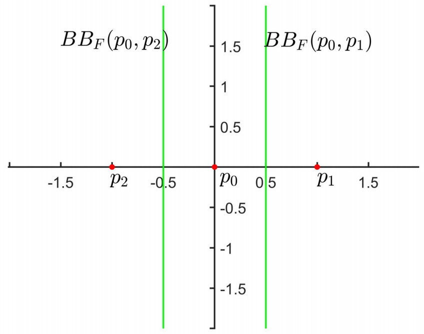


Figure 11: visualization of example 3.61

If the set $P \subset \Omega$ is affinely independent, there exists always a dual circumball. In their paper [1], Edelsbrunner and Wagner gave an algorithm for computing the circumball of an affinely independent set.

Algorithm 1: `CircumBall`

Input : P
let $\mathcal{A}$ be the affine hull of the points $(p, F(p))$, $p \in P$;
find $(q, \psi) \in \mathcal{A}$ minimizing $\mathbb{R}^{n+1} \to \mathbb{R}, (a, \alpha) \mapsto F(a) - \alpha$ over $\mathcal{A}$;
Output: $(q, \psi - F(q))$

Lemma 3.62 *Let $P \subset \Omega$ be an affinely independent set. The* `CircumBall` *algorithm computes the center and the radius of a dual circumball of P.*

Proof: Let $P = \{p_0, \ldots, p_k\} \subset \Omega$ be an affinely independent subset. In the first part of the proof we show that $g : \mathbb{R}^{n+1} \to \mathbb{R}, (a, \alpha) \mapsto F(a) - \alpha$ has always a minimum over $\mathcal{A} := \operatorname{Aff}\big((p_0, F(p_0)), \cdots, (p_k, F(p_k))\big)$.

The epigraph of F is a closed convex set, since F is a continuous convex function. By definition 3.28, for every $(a, \alpha) \in \mathcal{A}$ it holds that

$$F(a) - \alpha \begin{cases} \leq 0 & \text{if } (a, \alpha) \in \operatorname{epi}(F) \\ > 0 & \text{if } (a, \alpha) \notin \operatorname{epi}(F) \end{cases}$$

From this follows that the minimum point of g lies in $\mathcal{A} \cap \operatorname{epi}(F)$, if it exists.

Since g is continuous, it is enough to show that $\mathcal{A} \cap \operatorname{epi}(F)$ is compact, or equivalently, it is bounded and closed. The sets $\mathcal{A}$ and $\operatorname{epi}(F)$ are closed so their intersection is closed as well.

The recession cone of the convex set $\mathcal{A} \cap \operatorname{epi}(F)$ consists of the zero vector alone. From [23, Theorem 8.4] follows that it is bounded.

In conclusion, g has always a minimum. Denote $(q, \psi) \in \mathcal{A}$ the minimum point of g over $\mathcal{A}$. Since g is strictly convex, the minimum point (q, ψ) is unique.

In the second part of the proof we show that P has a dual circumball with center q and radius $\psi - F(q)$. In the rest of this proof we consider the vectors in $\mathbb{R}^n$ as column vectors.

One can easily show that $\left\{ \begin{pmatrix} p_0 \\ F(p_0) \end{pmatrix}, \dots, \begin{pmatrix} p_k \\ F(p_k) \end{pmatrix} \right\}$ is affinely independent as well. This implies that the matrix

$$M := \begin{pmatrix} p_1 - p_0 & \dots & p_k - p_0 \\ F(p_1) - F(p_0) & \dots & F(p_k) - F(p_0) \end{pmatrix} \in \mathbb{R}^{(n+1)\times k}$$

has rank k, and the affine linear map $\phi : \mathbb{R}^k \to \mathbb{R}^{n+1}, \mu \mapsto M\mu + b$ is injective, where $b = \begin{pmatrix} p_0 \\ F(p_0) \end{pmatrix} \in \mathbb{R}^{(n+1)}$.

The image of ϕ is $\mathcal{A}$ by construction. There is a unique $\lambda \in \mathbb{R}^k$ such that $\phi(\lambda) = \begin{pmatrix} q \\ \psi \end{pmatrix}$. This λ minimizes the function

$$f := g \circ \phi : \mathbb{R}^k \to \mathbb{R}, \mu \mapsto F\Big((1 - \sum_{i=1}^{k} \mu_i) p_0 + \sum_{i=1}^{k} \mu_i p_i\Big) - \big(1 - \sum_{i=1}^{k} \mu_i\big) F(p_0) - \sum_{i=1}^{k} \mu_i F(p_i)$$

Since ϕ is an injective affine linear function and g is strictly convex, one can easily show that f is strictly convex. The minimum point of a strictly convex function is unique. A point λ is a minimum point of f if and only if

$$0 = J_f(\lambda) = J_g(\phi(\lambda)) \cdot J_\phi(\lambda) = (\nabla F(q)^t, -1) \cdot M.$$

This is equivalent to $\langle p_i - p_0, \nabla F(q) \rangle + F(p_0) - F(p_i) = 0$ for all $i = 1, \dots, k$. From lemma 3.60 it follows that q is a center of a dual circumball of P.

Using $q = \sum_{j=1}^{k} \lambda_j (p_j - p_0) + p_0$ and $\psi = \sum_{j=1}^{k} \lambda_j \big(F(p_i) - F(p_0)\big) + F(p_0)$ one can compute the radius of this dual circumball. We give the computation for completeness.

$$\begin{aligned}
D_F(p_0\|q) &= F(p_0) - F(q) - \langle \nabla F(q), p_0 - q\rangle = \\
&= F(p_0) - F(q) + \langle \nabla F(q), \sum_{j=1}^{k} \lambda_j (p_j - p_0)\rangle = \\
&= F(p_0) - F(q) + \sum_{j=1}^{k} \lambda_j \langle \nabla F(q), (p_j - p_0)\rangle = \\
&= F(p_0) - F(q) + \sum_{j=1}^{k} \lambda_j \big(F(p_j) - F(p_0)\big) = \psi - F(q).
\end{aligned}$$

■

Remark 3.63 *The constrained n-dimenisonal optimization problem*

$$\begin{aligned}
&\underset{(a,\alpha)\in\mathbb{R}^{n+1}}{\text{minimize}} && F(a) - \alpha \\
&\text{subject to} && (a,\alpha) \in Aff\Big(\big(p_0, F(p_0)\big), \cdots, \big(p_k, F(p_k)\big)\Big)
\end{aligned}$$

can be very demanding if n is big. In the proof above we have seen that it is enough to solve the unconstrained k-dimenisonal convex optimization problem

$$\underset{\mu\in\mathbb{R}^k}{\text{minimize}} \quad F\Big((1 - \sum_{i=1}^{k} \mu_i) p_0 + \sum_{i=1}^{k} \mu_i p_i\Big) - \big(1 - \sum_{i=1}^{k} \mu_i\big) F(p_0) - \sum_{i=1}^{k} \mu_i F(p_i)$$

which could be done more effectively.

In Matlab, there are several built-in functions for solving unconstrained convex optimization problems. For the implementation of the `CircumBall` algorithm, we used the function `fminunc`. One can find the corresponding Matlab code in appendix A.1.

Definition 3.64 *Let $P \subset \Omega$ be a finite subset. We call a dual (resp. primal) circumball of P smallest dual (resp. primal) circumball, if there is no other dual (resp. primal) circumball P with smaller radius.*

Proposition 3.65 *If there is at least one dual (resp. primal) circumball of P, then exists the smallest dual (resp. primal) circumball of P and it is unique.*

Proof: First, we consider primal circumballs. By lemma 3.57, the center of each primal circumball of P lies in $\bigcap_{p,q\in P} BB'(p,q)$. So from the assumption follows that

$\bigcap_{p,q\in P} BB'(p,q)$ is a nonempty set. From proposition 3.38 it follows that $\bigcap_{p,q\in P} BB'(p,q)$ is closed and convex.

Pick a point $p \in P$. The radius of a primal circumball with center c is given by $D_F(c\|p)$. So, the center of the smallest primal circumball of P is the Bregman projection of p onto $\bigcap_{p,q\in P} BB'(p,q)$. In proposition 3.41, we have already seen that the Bregman projection onto a nonempty closed convex set always exists and it is unique.

The result for dual circumballs follows with duality. If P has a dual circumball, then P^* has a primal circumball by proposition 3.59 and therefore there is a unique smallest primal circumball $B_{F^*}(c^*, r)$ of P^*.

Using lemma 3.25 one can show that $\nabla F^*\big(B_{F^*}(c^*, r)\big)$ is a dual circumball of P. It is the smallest dual circumball and it is unique, since every dual circumball of P corresponds to a primal circumball of P^*. ■

In [24] Edelsbrunner et al. studied the location of the center of smallest dual circumballs. We recall their result.

Proposition 3.66 *Let $P \subset \Omega$ be a finite subset such that there is a dual circumball of P. The center of the smallest dual circumball lies in the affine hull of P.*

Proof: By proposition 3.65, the smallest dual circumball of P exists if P has a dual circumball. Let $c \in \Omega$ be the center of the smallest dual circumball of P and r its radius.

Assume that $c \notin \mathrm{Aff}(P)$ and let $c_A \in \mathrm{Aff}(P)$ be the Bregman projection of c onto $\mathrm{Aff}(P)$. Since $\mathrm{Aff}(P)$ is an affine subspace, lemma 3.42 implies that

$$D_F(p\|c_A) = D_F(p\|c) - D_F(c_A\|c) = r - D_F(c_A\|c)$$

for each $p \in P \subset \mathrm{Aff}(P)$. Thus, there is a dual circumball of P with center $c_A \in \mathrm{Aff}(P)$ and radius $r - D_F(c_A\|c) < r$. But this contradicts to the minimality of r. ■

Lemma 3.67 *Let $P \subset \Omega$ be an affinely independent subset. The* `CircumBall` *algorithm computes the center and the radius of the smallest dual circumball of P.*

Proof: We follow the notation used in the proof of lemma 3.62 and we write $\{p_0, \ldots, p_k\}$ for P. Lemma 3.62 tells us that the `CircumBall` algorithm computes a dual circumball with center in $\mathrm{Aff}(P)$. However, we do not know yet whether it is the smallest dual circumball or not.

Let $c \in \Omega$ be the center of the smallest dual circumball of P. By proposition 3.66, c lies in $\text{Aff}(P)$. Thus, there is a $\mu \in \mathbb{R}^k$ such that $c = \big(1 - \sum_{i=1}^{k} \mu_i\big)p_0 + \sum_{i=1}^{k} \mu_i p_i$.

By lemma 3.60, $\nabla F(c)$ satisfies

$$\langle p_i - p_0, \nabla F(c)\rangle + F(p_0) - F(p_i) = 0$$

for all $i = 1, \ldots, k$. If we recall the proof of lemma 3.62, we see that this implies that μ is a minimum point of the function f.

Since f is strictly convex, so it has a unique minimum point. The `CircumBall` algorithm computes this minimum point. In conclusion, the first output of the `CircumBall` algorithm is c. ■

The next corollary seems to be a bit unnecessary, but in the next chapter it will turn out that it is extremely useful to have such a statement.

Corollary 3.68 *Let $P \subset \Omega$ be an affinely independent subset and $B'_F(c, r)$ be a dual circumball of P. If c lies in $\text{Aff}(P)$, then $B'_F(c, r)$ is the smallest dual circumball of P.*

Proof: Write $\{p_0, \ldots, p_k\}$ for P. If c lies in $\text{Aff}(P)$ then there is a $\mu \in \mathbb{R}^k$ such that

$$c = \big(1 - \sum_{i=1}^{k} \mu_i\big)p_0 + \sum_{i=1}^{k} \mu_i p_i.$$

Analog to the proof above, it follows from lemma 3.60 that μ is a minimum point of the function f from the proof of lemma 3.62. Since the minimum point of f is unique, c is the center of the smallest dual circumball of P. ■

3.6. Voronoi diagrams

Last but not least, we introduce *Voronoi diagrams.* They will help us to build lower-dimensional simplicial complexes for *Bregman point clouds.* In this section, let $F : \Omega \to \mathbb{R}^n$ be a function of Legendre type.

Definition 3.69 *Let $X \subset \Omega$ be a finite subset. The primal Voronoi domain of $x \in X$ is defined as*

$$V_F(x) := \{a \in \Omega \mid \forall y \in X : D_F(x\|a) \leq D_F(y\|a)\}.$$

The dual Voronoi domain of x is defined as

$$V'_F(x) := \{a \in \Omega \mid \forall y \in X : D_F(a\|x) \leq D_F(a\|y)\}.$$

Proposition 3.70 *Let $X \subset \Omega$ be a finite subset. The dual Voronoi domain $V'_F(x)$ is convex for all $x \in X$.*

Proof: Before we go into the proof we recall some notation.
The best linear approximation of F in y is given by

$$H_y : \Omega \to \mathbb{R}, a \mapsto H_y(a) := F(y) + \langle \nabla F(y), a - y \rangle.$$

The halfspace above the graph of H_y is given by $\mathcal{H}_y^+ = \{(a, \alpha) \in \hat{\Omega} \mid H_y(a) \leq \alpha\}$.

Fix a point $x \in X$. By proposition 3.21, $D_F(a\|x) \leq D_F(a\|y)$ is equivalent to $H_y(a) \leq H_x(a)$ for all $a \in \Omega$ and $y \in X$.

This implies that $a \in \Omega$ lies in the dual Voronoi domain $V_F'(x)$ if and only if $\big(a, H_x(a)\big)$ is in the convex set $\bigcap\limits_{y \in X} \mathcal{H}_y^+ \bigcap \mathcal{H}_x$.

Thus, $V_F'(x)$ is the image of the convex set $\bigcap\limits_{y \in X} \mathcal{H}_y^+ \bigcap \mathcal{H}_x$ under $Proj_\Omega : \hat{\Omega} \to \Omega$. Since the projection map is linear, it follows that $V_F'(x)$ is a convex set. ■

In general, primal Voronoi domains are not convex.

Proposition 3.71 *Let $X \subset \Omega$ be a finite subset. For all $x \in X$ it holds that*

$$V_F(x) = \nabla F^*\big(V_{F^*}'(x^*)\big) \text{ and } V_F'(x) = \nabla F^*\big(V_{F^*}(x^*)\big).$$

Proof: It follows from definition 3.69 and Legendre duality (lemma 3.25). ■

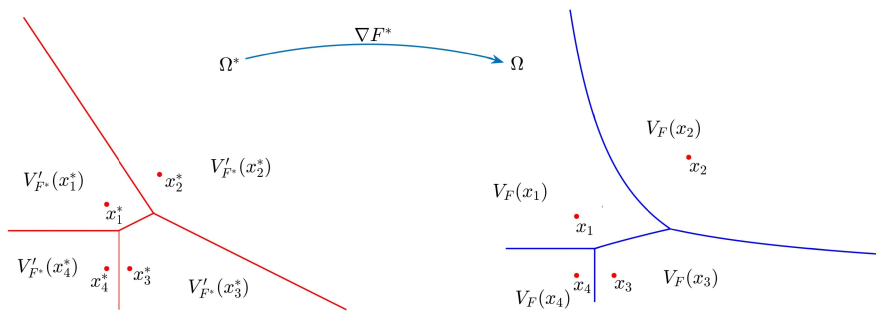


Figure 12: dual and primal Voronoi domains

Lemma 3.72 *Intersection of primal (resp. dual) Voronoi domains is either empty or contractible.*

Proof: The proof is similar to the proof of lemma 3.36. From proposition 3.71 it follows that the intersection of dual Voronoi domains is convex, this implies that their intersection is either empty or contractible.

The result for primal Voronoi domains follows from proposition 3.71 and the fact that ∇F^* is a homeomorphism. ■

Definition 3.73 *Let $X \subset \Omega$ be a finite subset. The* <u>*primal Voronoi diagram*</u> *of X is defined as the set of all primal Voronoi domains of the points in X, i.e.*

$$\mathrm{Vor}_F(X) := \{V_F(x) \mid x \in X\}.$$

We define the <u>*dual Voronoi diagram*</u> *of X as*

$$\mathrm{Vor}'_F(X) := \{V'_F(x) \mid x \in X\}.$$

Proposition 3.74 *Let $X \subset \Omega$ be a finite subset. Then $\bigcup_{x \in X} V_F(x) = \Omega$.*

Proof: Since X is a finite set, for every $a \in \Omega$ there is a $p \in X$ such that $D_F(p\|a) = \min_{x \in X} D_F(x\|a)$. By construction, $a \in V_F(p)$. ∎

Lemma 3.75 *Let $X \subset \Omega$ be a finite subset and $X^* := \{x^* \mid x \in X\}$, it holds that*

$$\mathrm{Vor}_F(X) = \nabla F^*\big(\mathrm{Vor}'_{F^*}(X^*)\big) \text{ and } \mathrm{Vor}'_F(X) = \nabla F^*\big(\mathrm{Vor}_{F^*}(X^*)\big).$$

Proof: It follows from proposition 3.71. ∎

In [19], Boissonnat and Nielsen proved that dual Voronoi diagrams are identical to *power diagrams*, also known as *Euclidean weighted Voronoi diagrams*. There are several algorithms that compute power diagrams see e.g. [27]. Furthermore, one can find an implemented algorithm on MATLAB Central File Exchange [28]. Using these observations and lemma 3.75, we can also compute primal Voronoi diagrams.

In the rest of this section, we introduce power diagrams and show the correspondence between dual Voronoi diagrams and power diagrams.

Definition 3.76 *Let U be a finite subset in some Euclidean space $\mathbb{R}^m$. To each point $u \in U$ we order a number $w_u \in \mathbb{R}$ and say w_u is the* <u>*weight*</u> *of u.*

a) The <u>*power*</u> *of a point $z \in \mathbb{R}^m$ is defined as $\pi_u(z) := \|z-u\|^2 - w_u$.*

b) The <u>*power cell*</u> *of $u \in U$ is the set of points $z \in \mathbb{R}^m$ with $\pi_u(z) \le \pi_v(z)$ for all $v \in U$. We write $\mathrm{pow}(u)$ for this set.*

c) The <u>*power diagram*</u> *of U with weights $(w_u)_{u \in U}$ is the set of all power cells.*

Lemma 3.77 *Let $X \subset \Omega$ be a finite subset. The dual Voronoi diagram of $X^* := \{x^* = \nabla F(x) \mid x \in \Omega\}$ corresponds to the power diagram of X with weigths $w_x = \|x\|^2 - 2F(x)$, $x \in X$. This correspondence is given by*

$$V'_{F^*}(x^*) = \mathrm{pow}(x) \cap \Omega^*$$

Proof: The proof is a straightforward computation. Let $x, y \in X$ and $a^* \in \Omega^*$. Using proposition 3.21 b), one can observe that $D_{F^*}(a^* \| x^*) \leq D_{F^*}(a^* \| y^*)$ is equivalent to $2F(x) - 2\langle x, a^* \rangle \leq 2F(y) - 2\langle y, a^* \rangle$.

Adding and subtracting $\|a^* - x\|^2$ to the left and $\|a^* - y\|^2$ to the right side we have

$$\|a^* - x\|^2 - \|a^*\|^2 - \|x\|^2 + 2F(x) \leq \|a^* - y\|^2 - \|a^*\|^2 - \|y\|^2 + 2F(y),$$

and adding $\|a^*\|^2$ to both sides we get

$$\|a^* - x\|^2 - \Big(\|x\|^2 - 2F(x)\Big) \leq \|a^* - y\|^2 - \Big(\|y\|^2 - 2F(y)\Big) .$$

■

4. Topological data analysis with Bregman divergences

After we have studied the basic objects of Bregman geometry, we are ready to start with data analysis. First, we introduce *Čech* and *Delaunay complexes* of *Bregman point clouds*. These constructions are well understood for metric spaces [12] [29]. We can generalize these results to our Bregman setting using the statements from the previous chapter. We follow [1] and describe algorithms that compute these filtered simplicial complexes of Bregman point clouds.

4.1. Bregman point clouds and their simplicial complexes

Definition 4.1 *Let $F : \Omega \to \mathbb{R}$ be a function of Legendre type and $X \subset \Omega$ a finite subset. We call the pair (X, F) a Bregman point cloud.*

Definition 4.2 *Given a Bregman point cloud (X, F) and $r \geq 0$, the Čech complex of (X, F) and $r \geq 0$ is defined as the nerve of the collection of primal Bregman balls $\big\{B_F(x, r)\big\}_{x \in X}$, i.e.*

$$\text{Čech}_F(X, r) = \{Q \subset X \mid \bigcap_{q \in Q} B_F(q, r) \neq \emptyset\}.$$

A Čech complex can be extremely high-dimensional even when the data lives in a low dimensional space. In the usual Euclidean setting there is another well known simplicial complex of the data, the Delaunay complex, or also called alpha complex [12]. The dimension of such an alpha complex is generically less or equal than the dimension of the containing space. This motivates us to introduce *Delaunay complexes* for Bregman point clouds.

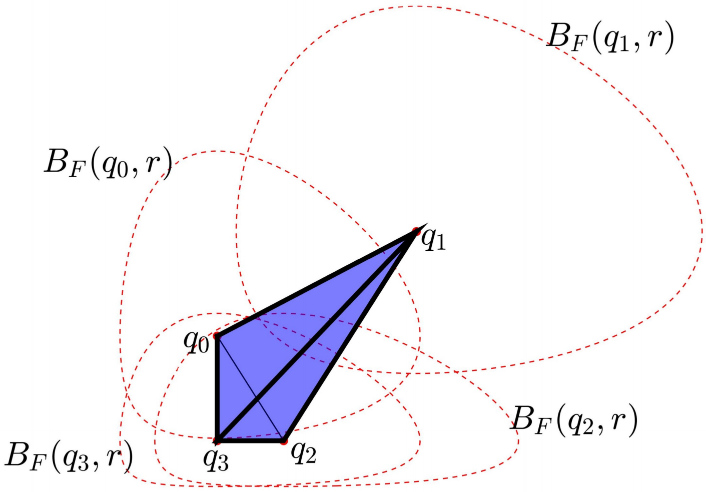


Figure 13: a simplex spanned by 4 points in the Čech complex

Definition 4.3 *The nerve of the primal Voronoi domains is called the* <u>*Delaunay triangulation*</u> *of* (X, F)*, denoted by* $\mathrm{Del}_F(X)$*, with other words*

$$\mathrm{Del}_F(X) := \{Q \subset X \mid \bigcap_{q \in Q} V_F(q) \neq \emptyset\}.$$

Definition 4.4 *Let* U *be a finite set of points in some Euclidean space* $\mathbb{R}^m$ *with real weights* $(w_u)_{u \in U}$*. The* <u>*Euclidean weighted Delaunay triangulation*</u> *of* U *is the nerve of the power diagram of* U*.*

Lemma 4.5 *Let* (X, F) *a Bregman point cloud such that* $\Omega^* = \mathbb{R}^n$*. In this case, the Delaunay triangulation of* (X, F) *is the same as the Euclidean weighted Delaunay triangulation of* X *with weights* $w_x = \|x\|^2 - 2F(x)$*,* $x \in X$*.*

Proof: Since ∇F^* is a bijection, from proposition 3.71 it follows for each $Q \subset X$ that $\bigcap_{q \in Q} V_F(q) \neq \emptyset$ if and only if $\bigcap_{q \in Q} V'_{F^*}(q^*) \neq \emptyset$.
By lemma 3.77, the second intersection is equal to $\bigcap_{q \in Q} \mathrm{pow}(q) \cap \Omega^*$. The assumption $\Omega^* = \mathbb{R}^n$ implies that the Delaunay triangulation of (X, F) and the Euclidean weighted Delaunay triangulation of X with weights $w_x = \|x\|^2 - 2F(x)$ are the same. ∎

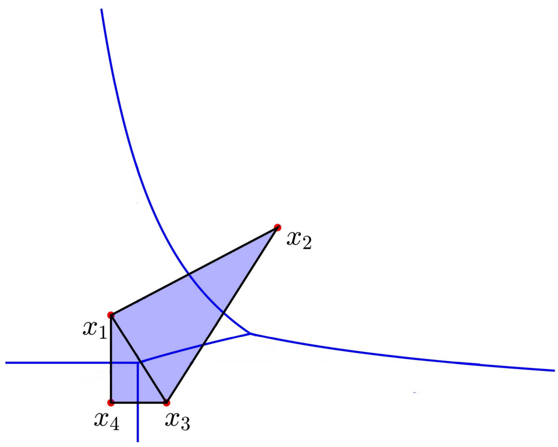


Figure 14: Delaunay triangulation of $\{x_1, x_2, x_3, x_4\}$

Unfortunately, the Burg entropy does not satisfy the condition of 4.5. Nevertheless, we can use the above statement for our two other examples, for the Shannon entropy, and the squared Euclidean norm.
In these cases, one can use the Matlab function `WeightedDelTriangulation` (see appendix A.4) for the computation of the Delaunay triangulation of a Bregman point cloud. The function `WeightedDelTriangulation` is a modified version of [28].

Definition 4.6 *The Delaunay complex of (X, F) and $r \geq 0$ is the nerve of the collection $\{B_F(x,r) \cap V_F(x)\}_{x \in X}$, i.e.*

$$\mathrm{Del}_F(X,r) = \{Q \subset X \mid \bigcap_{q \in Q} B_F(q,r) \cap V_F(q) \neq \emptyset\}.$$

Lemma 4.7 *For a Bregman point cloud (X, F) and $r \geq 0$ we have that*

$$|\,\check{\mathrm{C}}\mathrm{ech}_F(X,r)\,| \simeq |\,\mathrm{Del}_F(X,r)\,| \simeq \bigcup_{x \in X} B_F(x,r).$$

Proof: By lemma 3.36, the intersection of two primal Bregman balls is either empty or contractible. From the Nerve Theorem 2.4, it follows that $|\,\check{\mathrm{C}}\mathrm{ech}_F(X,r)\,|$ has the same homotopy type as the union of the primal Bregman balls.

From lemma 3.36 and lemma 3.72 it follows that the intersection of two elements of the collection $\{B_F(x,r) \cap V_F(x)\}_{x \in X}$ is either empty or contractible. The Nerve Theorem 2.4 implies that $|\,\mathrm{Del}_F(X,r)\,| \simeq \bigcup_{x \in X} B_F(x,r) \cap V_F(x)$.

From proposition 3.74 it follows that
$\bigcup_{x \in X} B_F(x,r) \cap V_F(x) = \bigcup_{x \in X} B_F(x,r) \cap \bigcup_{x \in X} V_F(x) = \bigcup_{x \in X} B_F(x,r)$. ∎

Remark 4.8 *In the setting of metric spaces, the Vietoris-Rips complex approximates the Čech complex of the data. In [1], Edelsbrunner and Wagner gave an example that there is no interleaving between the Čech and Vietoris-Rips complex of a Bregman point cloud. That is the reason why we do not consider Vietoris-Rips complexes of Bregman point clouds.*

In [29], Bauer and Edelsbrunner showed that the Delaunay and Čech complexes of points living in a Euclidean space arise as sublevel sets of *generalized discrete Morse functions.* We can translate this result to our Bregman setting. It will play a significant role in computing simplicial complexes of Bregman point clouds.
In order to handle Čech and Delaunay complexes at the same time, we generalize these constructions.

Definition 4.9 *Let (X, F) be a Bregman point cloud, $S \subset X$ and $r \geq 0$.*

a) The primal Voronoi ball of $x \in X$ is defined as

$$V_F(x, r \mid S) := \big\{a \in \Omega \mid D_F(x\|a) \leq r \text{ and } \forall s \in S : D_F(x\|a) \leq D_F(s\|a)\big\}.$$

b) The selective Delaunay complex $\mathrm{Del}_F(X, r \mid S)$ is the nerve of the collection $\big\{V_F(x, r \mid S)\big\}_{x \in X}$ i.e.

$$\mathrm{Del}_F(X, r \mid S) = \{P \subset X \mid \bigcap_{x \in P} V_F(x, r \mid S) \neq \emptyset\}.$$

c) The selective Delaunay triangulation is given by

$$Del_F(X \mid S) := \bigcup_{r \geq 0} Del_F(X, r \mid S).$$

Proposition 4.10 *Let (X, F) be a Bregman point cloud, $S \subset X$ and $r \geq 0$. It holds that*

a) $V_F(x, r \mid \emptyset) = B_F(x, r)$,

b) $\mathrm{Del}_F(X, r \mid \emptyset) = \check{\mathrm{C}}\mathrm{ech}_F(X, r)$,

c) $\mathrm{Del}_F(X \mid \emptyset) = \Delta(X)$,

d) $V_F(x, r \mid X) = B_F(x, r) \cap V_F(x)$,

e) $\mathrm{Del}_F(X, r \mid X) = \mathrm{Del}_F(X, r)$,

f) $\mathrm{Del}_F(X \mid X) = \mathrm{Del}_F(X)$.

Proof: All of the statements follow immediately from the definitions. ■

Proposition 4.11 *Let (X, F) be a Bregman point cloud, $S \subset X$ and $r \geq 0$.*

a) For every $r \leq r'$ it holds that $\mathrm{Del}_F(X, r \mid S) \subset \mathrm{Del}_F(X, r' \mid S)$.

b) For every $E \subset S$ it holds that $\mathrm{Del}_F(X, r \mid S) \subset \mathrm{Del}_F(X, r \mid E)$.

Proof: Note that $V_F(x, r \mid S) \subset V_F(x, r' \mid S)$ for all $x \in X$ and $r \leq r'$. From this, part a) follows immediately.

If $E \subset S$, then $V_F(x, r \mid E)$ has less constraints than $V_F(x, r \mid S)$ and this implies that $V_F(x, r \mid S) \subset V_F(x, r \mid E)$ for all $x \in X$. ■

Definition 4.12 *Let (X, F) be a Bregman point cloud and $S \subset X$. We say that a dual Bregman ball $B'_F(z, r)$ excludes S if $D_F(s \| z) \geq r$ for all $s \in S$.*

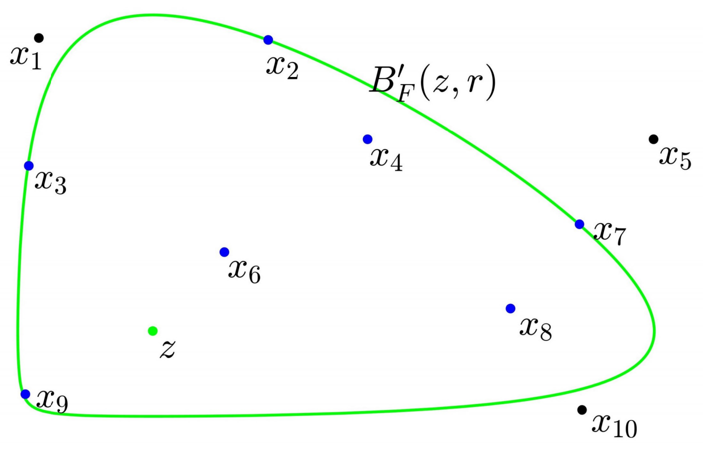


Figure 15: the dual Bregman ball $B'_F(z, r)$ excludes $S = \{x_1, x_5, x_{10}\}$

Proposition 4.13 *Let (X, F) be a Bregman point cloud and $P, S \subset X$, $P \neq \emptyset$. There exists an including dual Bregman ball of P that excludes S if and only if $P \in \mathrm{Del}_F(X \mid S)$.*
In this case, there is a unique smallest including dual Bregman ball of P that excludes S. i.e. there exists a unique including dual Bregman ball of P that excludes S such that there is no other including dual Bregman ball of P excluding S with smaller radius.

Proof: If there is a dual Bregman ball $B'_F(y \| r)$ that includes P and excludes S then y is in $\bigcap_{p \in P} V_F(p, r \mid S)$ and therefore $P \in \mathrm{Del}_F(X \mid S)$ by definition 4.9.

Conversely, assume that $P \in \mathrm{Del}_F(X \mid S)$. This implies that $\bigcap_{p \in P} V_F(p, r \mid S)$ is nonempty, i.e. there is an $a \in \Omega$ such that $D_F(p\|a) \leq r$ and $D_F(p\|a) \leq D_F(s\|a)$ for all $p \in P$ and $s \in S$. Note that the dual Bregman ball $B'_F(a, r)$ includes P but it does not necessarily exclude S.

We can write each primal Voronoi ball $V_F(p, r \mid S)$ as the intersection of the compact set $B_F(p, r)$ and the closed set $\{a \in \Omega \mid \forall s \in S : D_F(p\|a) \leq D_F(s\|a)\} =$
$= \bigcap_{s \in S} \Big(D_F(s\|\bullet) - D_F(p\|\bullet)\Big)^{-1}([0, \infty))$.

From this follows that $\bigcap_{p \in P} V_F(p, r \mid S)$ is a compact set. Since this intersection is nonempty by our assumption, the continuous function

$$f : \bigcap_{p \in P} V_F(p, r \mid S) \to \mathbb{R}, x \mapsto \max_{p \in P} D_F(p\|x)$$

has a minimum point. Let z denote a minimum point of f and ρ the value of f at z. Since P is a finite set, there is a $p_0 \in P$ such that $D_F(p_0\|z) = \rho$. From $z \in V_F(p_0, r \mid S)$ it follows that $\rho = D_F(p_0\|z) \leq D_F(s\|z)$ for all $s \in S$.
In conclusion, the dual Bregman ball $B'_F(z, \rho)$ includes P and excludes S.

To show uniqueness, it is enough to show that f has a unique minimum point. In order to show this, we use Legendre duality 3.25 and consider the strictly convex function $g : \bigcap_{p \in P} V'_{F^*}(p^*, r \mid S^*) \to \mathbb{R}, a \mapsto \max_{p \in P} D_F(a^*\|p^*)$, where for each $p \in P$

$$V'_{F^*}(p^*, r \mid S^*) := \Big\{a^* \in \Omega^* \mid D_{F^*}(a^*\|p^*) \leq r \text{ and } \forall s \in S : D_{F^*}(a^*\|p^*) \leq D_{F^*}(a^*\|s^*)\Big\}.$$

Since ∇F preserves the divergence, $\nabla F\big(V_F(p, r \mid S)\big) = V'_{F^*}(p^*, r \mid S^*)$ and a point y is a minimum point of f if and only if y^* is a minimum point of g. Because of the strict convexity, g has at most one minimum point, so the minimum point of f is unique. ■

Corollary 4.14 *Let (X, F) be a Bregman point cloud. A simplex $P \in \Delta(X)$ is in the Delaunay triangulation $\mathrm{Del}_F(X)$ if and only if it has a unique smallest empty dual circumball $B'_F(z, r)$, i.e. a dual circumball such that its interior $\big(B'_F(z, r) \backslash \partial B'_F(z, r)\big)$ contains no point of X, and no other such dual circumball has smaller radius.*

Proof: Note that an including dual Bregman ball that excludes X is an empty dual circumball and vice versa. By proposition 4.10 $\mathrm{Del}_F(X) = \mathrm{Del}_F(X \mid X)$. Now, the corollary follows from proposition 4.13 by $S = X$. ■

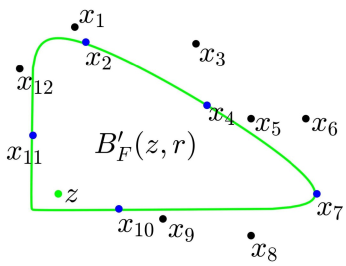


Figure 16: $B'_F(z,r)$ is an empty dual circumball of $\{x_2, x_4, x_7, x_{10}, x_{11}\}$

Definition 4.15 *Let (X,F) be a Bregman point cloud and $S \subset X$. The function*

$$\varrho_F^S : \mathrm{Del}_F(X \mid S) \to \mathbb{R}, P \mapsto \min\Big\{r \in [0,\infty) \Big| P \in \mathrm{Del}_F(X, r, \mid S)\Big\}$$

is called the radius function of (X,F) and S.

We write $\varrho_F^{\text{Čech}}$ for $\varrho_F^{\emptyset}$ and call it the Čech radius function of (X,F).

We write ϱ_F^{Del} for ϱ_F^{X} and call it the Delaunay radius function of (X,F).

Once we know the radius function, we can build a so-called explicit simplex stream in Matlab using the JavaPlex package [15]. This means that we can put the filtered simplicial complex $\mathrm{Del}_F(X, \bullet \mid S)$ in a form that our computer can understand and applying functions from JavaPlex we can compute persistent homology and persistence diagrams.

Lemma 4.16 *Let (X,F) be a Bregman point cloud and $S \subset X$.*
For all $P \in \mathrm{Del}_F(X \mid S)$, $\varrho_F^S(P)$ is the radius of the smallest including dual Bregman ball of P that excludes S.

Proof: By proposition 4.13, each $P \in \mathrm{Del}_F(X \mid S)$ has a smallest including dual Bregman ball that excludes S. Let $z \in \Omega$ be its center and $\rho \geq 0$ its radius. It is enough to show that $P \in \mathrm{Del}_F(X, r \mid S)$ if and only if $r \geq \rho$.

If $P \in \mathrm{Del}_F(X, r \mid S)$, then there is a $y \in \Omega$ sucht that $y \in V_F(p, r \mid S)$ for all $p \in P$. From this follows that $r' := \max_{p \in P} D_F(p \| y) \leq r$.

The dual Bregman ball $B'_F(y, r')$ includes P and excludes S. By minimality of ρ we have $r \geq r' \geq \rho$.

If $r \geq \rho$, then $P \in \mathrm{Del}_F(X, \rho \mid S) \subset \mathrm{Del}_F(X, r, \mid S)$ by proposition 4.11. ■

Corollary 4.17 *Let (X, F) be a Bregman point cloud.*

a) For every $P \in \Delta(X)$, $\varrho_F^{\text{Čech}}(P)$ is the radius of the smallest including dual ball of P.

b) For every $P \in \mathrm{Del}_F(X)$, ϱ_F^{Del} is the radius of the smallest empty dual circumball of P.

Proof: Note that the smallest including dual Bregman ball of P that excludes the empty set is the smallest including dual ball of P.

Similarly, the smallest including dual Bregman ball of P that excludes X is the smallest empty dual circumball of P. ■

4.2. Discrete Morse theory

In this section, we show that the radius function of a Bregman point cloud is a *generalized discrete Morse function* if we assume that the points are *in general position.* This property will help us introduce efficient algorithms to compute the radius functions. This section is based on [1] and [29]. First, we recall some definition from discrete Morse theory.

Definition 4.18 *Let K be a simplicial complex.*

a) For two simplices $P, R \in K$ we call the set of simplices

$$[P, R] := \{Q \in K \mid P \subset Q \subset R\}$$

the interval with lower bound P and upper bound R.

b) A generalized discrete vector field is a partition of K into intervals.

c) We call K a generalized discrete gradient if K is a generalized discrete vector field and there is a function $f : K \to \mathbb{R}$ such that for every $P, Q \in K$ with $P \subset Q$ it holds that $f(P) \leq f(Q)$, with equality if and only if P and Q belong to a common interval. The function f is called a generalized discrete Morse function.

Before we can establish that the Čech and Delaunay radius functions are generalized discrete Morse functions, we need a couple of preliminary statements. The main ingredient will be the (Karush)-Kuhn-Tucker Theorem [23, Corollary 28.3.1] from convex optimization theory.

Theorem 4.19 ***(KKT-Theorem)*** *Let $C \subset \mathbb{R}^n$ be a nonempty convex set and $f_0, \ldots, f_k : C \to \mathbb{R}$ differentiable convex functions. A vector $z \in C$ is an optimal solution to the convex optimization problem*

$$\begin{aligned} &\underset{x \in C}{\textit{minimize}} && f_0(x) \\ &\textit{subject to} && f_i(x) \leq 0 \quad \forall i = 1, \ldots, k \end{aligned}$$

if and only if there are nonnegative real numbers $\lambda_1, \ldots, \lambda_k \in [0, \infty)$ satisfying

(KKT0) $f_i(z) \leq 0$ *for all* $i = 1, \ldots, k$

(KKT1) $\lambda_i f_i(z) = 0$ *for all* $i = 1, \ldots, k$

(KKT2) $\nabla f_0(z) + \sum_{i=1}^{k} \lambda_i \cdot \nabla f_i(z) = 0.$

Lemma 4.20 *Let* (X, F) *be a Bregman point cloud and* $Q, S \subset X$*. Let* $B'_F(z, r)$ *be an including dual Bregman ball of* Q *that excludes* S*.*

a) Assume that there is an $y \in Q \cap S$*. The dual Bregman ball* $B'_F(z, r)$ *is the smallest including dual Bregman ball of* Q *that excludes* S *if and only if there are nonnegative real numbers* $(\lambda_q)_{q\in Q}$, $(\mu_s)_{s\in S}$ *such that*

- $\lambda_q = 0$ *if* $q \in Q \setminus \partial B'_F(z, r)$
- $\mu_s = 0$ *if* $s \in S \setminus \partial B'_F(z, r)$
- $\sum_{q\in Q} \lambda_q = 1$
- $z = \sum_{q\in Q\setminus S} \lambda_q \cdot q + \left(\lambda_y + \sum_{s\in S\setminus\{y\}} \mu_s\right) \cdot y + \sum_{x\in Q\cap S\setminus\{y\}} (\lambda_x - \mu_x) \cdot x + \sum_{s\in S\setminus Q} (-\mu_s) \cdot s.$

b) If $S = \emptyset$*, we have that* $B'_F(z, r)$ *is the smallest including dual Bregman ball of* Q *if and only if there are nonnegative real numbers* $(\lambda_q)_{q\in Q}$ *such that*

- $\lambda_q = 0$ *if* $q \in Q \setminus \partial B'_F(z, r)$
- $\sum_{q\in Q} \lambda_q = 1$
- $z = \sum_{q\in Q} \lambda_q \cdot q.$

Proof: First, we consider part a). Notice that $y \in Q \cap S$ lies on the boundary of every dual Bregman ball that includes Q and excludes S. With this in mind, we can see that $B'_F(z, r)$ is the smallest including dual Bregman ball of Q that excludes S if and only if z is an optimal solution of the following optimization problem.

$$\begin{aligned} &\underset{a\in\Omega}{\text{minimize}} && \max_{q\in Q} D_F(q\|a) \\ &\text{subject to} && D_F(y\|a) \leq D_F(s\|a) \quad \forall s \in S \end{aligned}$$

This problem is equivalent to its epigraph problem form [30]

$$\begin{aligned} &\underset{(a,\alpha)\in\Omega\times\mathbb{R}}{\text{minimize}} && \alpha \\ &\text{subject to} && \max_{q\in Q} D_F(q\|a) \leq \alpha \\ & && D_F(y\|a) \leq D_F(s\|a) \quad \forall s \in S \end{aligned}$$

Using proposition 3.21 we can see that this is equivalent to

$$\begin{aligned}
&\underset{(a^*,\alpha)\in\Omega^*\times\mathbb{R}}{\text{minimize}} && \alpha \\
&\text{subject to} && F^*(a^*) + F(q) - \langle q, a^*\rangle - \alpha \le 0 && \forall q \in Q \\
& && F(y) - \langle y, a^*\rangle - F(s) + \langle s, a^*\rangle \le 0 && \forall s \in S
\end{aligned}$$

Since it is a convex optimization problem and the functions are differentiable, we can apply the KKT-Theorem 4.19. Since $B'_F(z,r)$ is an including dual Bregman ball of Q that excludes S, the vector (z^*, r) satisfies the condition $(KKT0)$. Thus, (z^*, r) is an optimal solution if and only if there are nonnegative real numbers $(\lambda_q)_{q\in Q}$, $(\mu_s)_{s\in S}$ such that

- $\lambda_q \cdot \big(D_F(q\|z) - r\big) = 0$ for all $q \in Q$,
- $\mu_s \cdot \big(D_F(y\|z) - D_F(s\|z)\big) = 0$ for all $s \in S$,
- $(0,1) + \sum\limits_{q\in Q} \lambda_q \cdot (z - q, -1) + \sum\limits_{s\in S} \mu_s \cdot (s - y, 0) = (0,0)$.

One can easily rewrite these into the form given in part a).

Part b) follows from a similar argument. Since $S = \emptyset$, we do not have to consider its points. $B'_F(z,r)$ is the smallest including dual Bregman ball of Q if and only if (z,r) is an optimal solution of the optimization problem

$$\begin{aligned}
&\underset{(a,\alpha)\in\Omega\times\mathbb{R}}{\text{minimize}} && \alpha \\
&\text{subject to} && \max_{q\in Q} D_F(q\|a) \le \alpha
\end{aligned}$$

Using proposition 3.21 and the KKT-Theorem 4.19, it follows that (z,r) is optimal if and only if there are nonnegative real numbers $(\lambda_q)_{q\in Q}$ such that

- $\lambda_q \cdot \big(D_F(q\|z) - r\big) = 0$ for all $q \in Q$,
- $(0,1) + \sum\limits_{q\in Q} \lambda_q \cdot (z - q, -1) = (0,0)$.

These conditions are equivalent to the conditions given in part b). ■

Definition 4.21 *A Bregman point cloud (X, F) is in general position if for every $P \subset X \subset \mathbb{R}^n$ of cardinality at most $n+1$,*

(G1) the points in P are affinely independent and

(G2) no points of $X \setminus P$ lie on the boundary of the smallest dual circumball of P.

Proposition 4.22 *Let (X, F) be a Bregman point cloud in general position such that $X \subset \mathbb{R}^n$. In this case,*

a) every subset of X of cardinality at most $n+1$ has a dual circumball,

b) no subset of X of cardinality at least $n+2$ has a dual circumball.

Proof: By lemma 3.62, every affinely independent subset of X has a dual circumball, so part a) follows from $(G1)$.

To get a contradiction, we assume that there is a $Q \subset X$ of cardinality at least $n+2$ such that Q has a dual circumball $B'_F(z, r)$. Without loss of generality we can assume that $B'_F(z, r)$ is the smallest dual circumball of Q (see proposition 3.65).

Let $P \subset Q$ be a subset of cardinality $n+1$. By assumption $(G1)$, the set P is affinely independent. From this follows that $\mathrm{Aff}(P) = \mathrm{Aff}(Q)$.

By proposition 3.66, z lies in $\mathrm{Aff}(Q)$ and therefore in $\mathrm{Aff}(P)$. The smallest dual circumball of Q is also a dual circumball of P. Since $z \in \mathrm{Aff}(P)$, it follows from corollary 3.68 that $B'_F(z, r)$ is the smallest dual circumball of P. This is a contradiction to $(G2)$. ■

Corollary 4.23 *The dimension of the Delaunay complex of a Bregman point cloud in general position is bounded above by the dimension of the including space.*

Proof: Let (X, F) be a Bregman point cloud in general position such that $X \subset \mathbb{R}^n$. By corollary 4.14, every simplex in $\mathrm{Del}_F(X)$ has an empty dual circumball.
From proposition 4.22 b), it follows that every simplex in $\mathrm{Del}_F(X)$ has cardinality at most $n+1$. ■

Definition 4.24 *Let (X, F) be a Bregman point cloud in general position and $B'_F(z, r)$ the smallest dual circumball of some subset $P \subset X$. Let $(\nu_p)_{p \in P}$ a collection of real numbers such that $z = \sum_{p \in P} \nu_p \cdot p$ and $\sum_{p \in P} \nu_p = 1$ We define*

a) the <u>front face</u> of P as $\mathrm{Front}\left(B'_F(z, r)\right) := \{p \in P \mid \nu_p > 0\}$

b) and the <u>back face</u> of P as $\mathrm{Back}\left(B'_F(z, r)\right) := \{p \in P \mid \nu_p < 0\}$.

Remark 4.25 *In proposition 4.22 we have shown that every simplex $P \subset X$ with a dual circumball has cardinality at most $n+1$. By proposition 3.66, the center of the smallest dual circumball of $P \subset X$ lies in the affine hull of P, so such an affine combination always exists and it is unique since P is affinely independent by $(G1)$.*

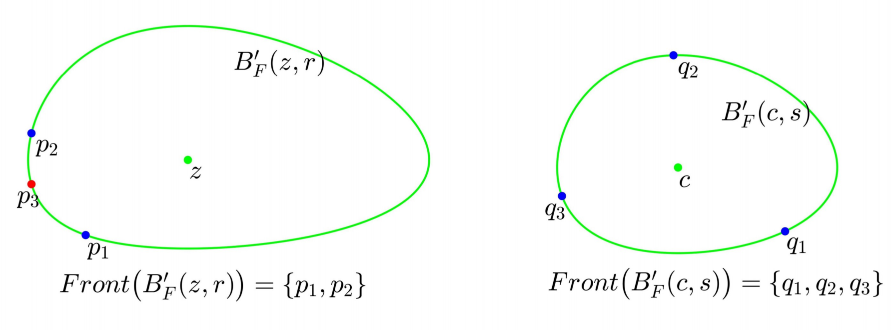


Figure 17: Examples for front face of a simplex

Proposition 4.26 *Let (X, F) be a Bregman point cloud in general position and $B'_F(z,r)$ the smallest dual circumball of a subset $P \subset X$. Then it holds that*

$$\text{Front}\left(B'_F(z,r)\right) \bigcup \text{Back}\left(B'_F(z,r)\right) = P.$$

Proof: Let $(\nu_p)_{p\in P}$ be a collection of real numbers such that $z = \sum_{p\in P} \nu_p \cdot p$ and $\sum_{p\in P} \nu_p = 1$. It is enough to show that $\nu_p \neq 0$ for all $p \in P$.

To get a contradiction, we assume that there is a $p_0 \in P$ such that $\nu_{p_0} = 0$. In this case z is an affine combination of $P \setminus \{p_0\}$ and $B'_F(z,r)$ is a dual circumball of $P \setminus \{p_0\}$. From corollary 3.68 it follows that $B'_F(z,r)$ is the smallest dual circumball of $P \setminus \{p_0\}$. This contradicts $(G2)$. ■

Theorem 4.27 *Let (X, F) be a Bregman point cloud in general position and $Q \in \Delta(X)$. Let $B'_F(z,r)$ be an including dual Bregman ball of Q. It is the smallest including dual Bregman ball of Q if and only if*

- *it is the smallest dual circumball of $\partial B'_F(z,r) \cap X$,*
- Front $\left(B'_F(z,r)\right) \subset Q$,
- *and* Back $\left(B'_F(z,r)\right) \subset \emptyset$.

Proof: Define $\tilde{Q} := Q \cap \partial B'_F(z,r)$.
First, we assume that $B'_F(z,r)$ is the smallest including dual Bregman ball of Q. By lemma 4.20 b) there are nonnegative real numbers $(\lambda_q)_{q\in\tilde{Q}}$ such that

- $\sum_{q\in\tilde{Q}} \lambda_q = 1$ and
- $z = \sum_{q\in\tilde{Q}} \lambda_q \cdot q.$

Define $\alpha_q := \lambda_q$ and $\beta_q := 0$ for all $q \in \tilde{Q}$ and fix a point $y \in \tilde{Q}$. So $(\alpha_q)_{q\in\tilde{Q}}$, $(\beta_q)_{q\in\tilde{Q}}$ are nonnegative and they satisfy

- $\sum\limits_{q\in\tilde{Q}} \alpha_q = 1$ and
- $z = \big(\alpha_y + \sum\limits_{s\in\tilde{Q}\setminus\{y\}} \beta_s\big) \cdot y + \sum\limits_{q\in\tilde{Q}\setminus\{y\}} (\alpha_q - \beta_q) \cdot q.$

By lemma 4.20 a), $B'_F(z,r)$ is the smallest including dual Bregman ball of $\tilde{Q}$ that excludes $\tilde{Q}$, i.e. it is the smallest dual circumball of $\tilde{Q}$. By $(G2)$ we have that $\tilde{Q} = \partial B'_F(z,r) \cap X$.

By definition 4.24, we have Front $\big(B'_F(z,r)\big) \subset \tilde{Q} \subset Q$. Since $\alpha_q \geq 0$ and $\beta_q = 0$ for all $q \in \tilde{Q}$, we also have Back $\big(B'_F(z,r)\big) = \emptyset$.

Conversely, let $B'_F(z,r)$ the smallest dual circumball of $\partial B'_F(z,r) \cap X$ such that Front $\big(B'_F(z,r)\big) \subset Q$ and Back $\big(B'_F(z,r)\big) \subset \emptyset$.

From proposition 4.26, it follows that

$$\partial B'_F(z,r) \cap X = \text{Front}\,\big(B'_F(z,r)\big) \cup \text{Back}\,\big(B'_F(z,r)\big) \subset Q \cup \emptyset,$$

and therefore $\partial B'_F(z,r) \cap X = \tilde{Q}$.

Since $B'_F(z,r)$ the smallest dual circumball of $\tilde{Q}$, from lemma 4.20 a) it follows the existence of nonnegative real numbers $(\alpha_q)_{q\in\tilde{Q}}$, $(\beta_q)_{q\in\tilde{Q}}$ such that

- $\sum\limits_{q\in\tilde{Q}} \alpha_q = 1,$
- $z = \big(\alpha_y + \sum\limits_{s\in\tilde{Q}\setminus\{y\}} \beta_s\big) \cdot y + \sum\limits_{q\in\tilde{Q}\setminus\{y\}} (\alpha_q - \beta_q) \cdot q.$

for a fixed $y \in \tilde{Q}$.

Define $\lambda_y := \alpha_y + \sum\limits_{s\in\tilde{Q}\setminus\{y\}} \beta_s \geq 0$ and $\lambda_q := \alpha_q - \beta_q$ for $q \in \tilde{Q} \setminus \{y\}$.
The numbers $(\lambda_q)_{q\in\tilde{Q}}$ are nonnegative since Back $\big(B'_F(z,r)\big) = \emptyset$.

For $q \in Q \setminus \partial B'_F(z,r)$ we set $\lambda_q := 0$. So $(\lambda_q)_{q\in Q}$ satisfy the conditions of 4.20 b), and therefore $B'_F(z,r)$ is the smallest including dual Bregman ball of Q. ■

Notation 4.28 *Let (X,F) a Bregman point cloud, $S \subset X$ and $P \in \mathrm{Del}_F(X \mid S)$. We write $B'_F(P \mid S)$ for the smallest including dual Bregman ball of P that excludes S.*

Remark 4.29

- $B'_F(P \mid \emptyset)$ *is the smallest including dual ball of* P.
- $B'_F(P \mid P)$ *is the smallest dual circumball of* P.
- $B'_F(P \mid X)$ *is the smallest empty dual circumball of* P.

Definition 4.30 *Let* (X, F) *be a Bregman point cloud in general position. We define* $\mathrm{Low}(X, F)$ *as the set of simplices* $\{P \in \Delta(X) \mid B'_F(P \mid P) = B'_F(P \mid \emptyset)\}$.

Corollary 4.31 ***(Čech Morse Function Corollary)***
The Čech radius function of a Bregman point cloud (X, F) *in general position is a generalized discrete Morse function. Its generalized discrete gradient consists the intervals* $[P, B'_F(P, P) \cap X]$, $P \in \mathrm{Low}(X, F)$.

Proof: First, we show that the intervals $[P, B'_F(P, P) \cap X]$, $P \in \mathrm{Low}(X, F)$ build a partition of $\Delta(X)$.

Let $Q \in \Delta(X)$ be a simplex. By theorem 4.27, the smallest including dual Bregman ball $B'_F(Q \mid \emptyset)$ is the smallest dual circumball of $P := \partial B'_F(Q \mid \emptyset) \cap X$, $\mathrm{Front}\left(B'_F(Q \mid \emptyset)\right) \subset Q$ and $\mathrm{Back}\left(B'_F(Q \mid \emptyset)\right) \subset \emptyset$.

Since $\mathrm{Front}\left(B'_F(Q \mid \emptyset)\right) \subset P$, we can use theorem 4.27 again and establish that $B'_F(Q \mid \emptyset)$ is also the smallest including dual Bregman ball of P.
In conclusion, $P \in \mathrm{Low}(X, F)$ and $Q \in [P, B'_F(Q \mid \emptyset) \cap X]$.

If there is a $P' \in \mathrm{Low}(X, F)$ such that $Q \in [P', B'_F(P', P') \cap X]$, then
$P' \subset Q \subset B'_F(P', \emptyset)$.

By theorem 4.27, $B'_F(P', \emptyset)$ is the smallest dual circumball of $\partial B'_F(P', \emptyset) \cap X$, $\mathrm{Front}\left(B'_F(P', \emptyset)\right) \subset P' \subset Q$ and $\mathrm{Back}(B'_F(P', \emptyset)) \subset \emptyset$.

Using the other direction of theorem 4.27, we see that $B'_F(P', \emptyset)$ is the smallest including dual Bregman ball of Q. Which implies that $P = \partial B'_F(P' \mid \emptyset) \cap X =$
$= \partial B'_F(P' \mid P') \cap X$. By $(G2)$, we have that $P = P'$. So the intervals are disjoint.

In the rest of the proof let $\emptyset \neq R \subset Q$. By corollary 4.17, $\varrho_F^{\text{Čech}}(Q)$ is the radius of the smallest including dual ball of Q. From this follows immediately that $\varrho_F^{\text{Čech}}(R) \leq \varrho_F^{\text{Čech}}(Q)$.

If there is a $P \in \mathrm{Low}(X, F)$ such that both R and Q belong to the interval $[P, B'_F(P, P) \cap X]$, then both $B'_F(R \mid \emptyset)$ and $B'_F(Q \mid \emptyset)$ are the smallest dual circumballs of P, since it is unique, it follows that $\varrho_F^{Čech}(R) = \varrho_F^{Čech}(Q)$.

If $\varrho_F^{\text{Čech}}(R) = \varrho_F^{\text{Čech}}(Q)$, then no other including dual Bregman ball of R has smaller radius than the radius of $B'_F(Q \mid \emptyset)$, so $B'_F(Q \mid \emptyset)$ is the smallest including dual Bregman ball of R. This implies that $\partial B'_F(Q \mid \emptyset) \cap X = \partial B'_F(R \mid \emptyset) \cap X$, i.e. R and Q belong to the same interval. ■

Theorem 4.32 *Let (X, F) be a Bregman point cloud in general position and let $Q, S \subset X$ such that $Q \cap S \neq \emptyset$. Let $B'_F(z, r)$ be an including dual Bregman ball of Q that excludes S. It is the smallest including dual Bregman ball of Q that excludes S if and only if*

- *it is the smallest dual circumball of $\partial B'_F(z, r) \cap X$,*
- $\text{Front}\big(B'_F(z, r)\big) \subset Q$ *and*
- $\text{Back}\big(B'_F(z, r)\big) \subset S$.

Proof: Fix a point $y \in Q \cap S$ and define $\tilde{Q} := \big(Q \cap \partial B'_F(z, r)\big) \setminus S$, $\tilde{S} := \big(S \cap \partial B'_F(z, r)\big) \setminus Q$, $W := (S \cap Q) \setminus \{y\}$.
By construction, we have $(Q \cup S) \cap \partial B_F(z, r) = \tilde{Q} \,\dot\cup\, \{y\} \,\dot\cup\, W \,\dot\cup\, \tilde{S} =: P$.

Notice that the smallest dual circumball of P is the smallest including dual Bregman ball of P that excludes P. By lemma 4.20 a), $B'_F(z, r)$ is the smallest dual circumball of P if and only if there are nonnegative real numbers $(\alpha_p)_{p \in P}$, $(\beta_p)_{p \in P}$ such that

$$(*) \quad \sum_{p \in P} \alpha_p = 1 \text{ and}$$

$$(**) \quad z = \big(\alpha_y + \sum_{s \in P \setminus \{y\}} \beta_s\big) \cdot y + \sum_{x \in P \setminus \{y\}} (\alpha_p - \beta_p) \cdot p.$$

First, we assume that $B'_F(z, r)$ is the smallest including dual Bregman ball of Q that excludes S. By lemma 4.20, there are nonnegative real numbers $(\lambda_q)_{q \in \tilde{Q} \dot\cup \{y\} \dot\cup W}$, $(\mu_s)_{s \in \tilde{S} \dot\cup \{y\} \dot\cup W}$ such that

$$(\circ) \quad \sum_{q \in \tilde{Q} \dot\cup \{y\} \dot\cup W} \lambda_q = 1 \text{ and}$$

$$(\circ\circ) \quad z = \sum_{q \in \tilde{Q}} \lambda_q \cdot q + \big(\lambda_y + \sum_{s \in \tilde{S} \cup W} \mu_s\big) \cdot y + \sum_{x \in W} (\lambda_x - \mu_x) \cdot x + \sum_{s \in \tilde{S}} (-\mu_s) \cdot s.$$

Set $\alpha_q := \lambda_q$, $\beta_q := 0$ for all $q \in \tilde{Q}$,
$\alpha_x := \lambda_x$, $\beta_x := \mu_x$ for all $x \in W$,
$\alpha_s := 0$, $\beta_s := \mu_s$ for all $s \in \tilde{S}$
and $\alpha_y := \lambda_y$, $\beta_y := 0$
Then $(\alpha_p)_{p \in P}$, $(\beta_p)_{p \in P}$ are nonnegative real numbers satisfiying $(*)$ and $(**)$, so $B'_F(z, r)$ is the smallest dual circumball of P.

By $(G2)$, we have $P = \partial B'_F(z,r) \cap X$. And from $(\circ\circ)$ it follows that $\text{Front}\big(B'_F(z,r)\big) \subset Q$ and $\text{Back}\big(B'_F(z,r)\big) \subset S$.

Conversely, assume that $B'_F(z,r)$ is the smallest dual circumball of $\partial B'_F(z,r) \cap X$, $\text{Front}\big(B'_F(z,r)\big) \subset Q$ and $\text{Back}\big(B'_F(z,r)\big) \subset S$.

From proposition 4.26 it follows that $\partial B'_F(z,r) \cap X \subset Q \cup S$, so $\partial B'_F(z,r) \cap X = P$. Since $B'_F(z,r)$ is the smallest dual circumball of P, there are nonnegative real numbers $(\alpha_p)_{p\in P}$, $(\beta_p)_{p\in P}$ satisfying $(*)$ and $(**)$.

For $q \in \tilde{Q}$ define $\lambda_q := \alpha_q - \beta_q$. This is nonnegative, since $\text{Back}\big(B'_F(z,r)\big) \subset S$.
For $s \in \tilde{S}$ define $\mu_s := \beta_s - \alpha_s$. It is nonnegative, since $\text{Front}\big(B'_F(z,r)\big) \subset Q$.
For $x \in W$ set $\lambda_x := \alpha_x \geq 0$ and $\mu_x := \beta_x \geq 0$
and we set $\lambda_y := \alpha_y + \sum_{s\in\tilde{S}} \alpha_s + \sum_{q\in\tilde{Q}} \beta_q \geq 0$.

By construction we have

$$\lambda_y + \sum_{s\in\tilde{S}\cup W} \mu_s = \alpha_y + \sum_{s\in\tilde{S}} \alpha_s + \sum_{q\in\tilde{Q}} \beta_q + \sum_{s\in W} \beta_s + \sum_{s\in\tilde{S}} (\beta_s - \alpha_s) = \alpha_y + \sum_{s\in P\setminus\{y\}} \beta_s$$

and

$$\sum_{q\in\tilde{Q}\cup W\cup\{y\}} \lambda_q = \alpha_y + \sum_{s\in\tilde{S}} \alpha_s + \sum_{q\in\tilde{Q}} \beta_q + \sum_{x\in W} \alpha_x + \sum_{q\in\tilde{Q}} (\alpha_q - \beta q) = 1$$

We set $\lambda_q := 0$ for $q \in Q \setminus \partial B'_F(z,r)$ and $\mu_s := 0$ for $s \in S \setminus \partial B'_F(z,r)$. Then $(\lambda_q)_{q\in Q}$, $(\mu_s)_{s\in S}$ satisfy the conditions from lemma 4.20 a). In conclusion, $B'_F(z,r)$ is the smallest including dual Bregman ball of Q that excludes S. ∎

Corollary 4.33 *Let (X,F) be a Bregman point cloud in general position and let $Q \subset X$. Let $B'_F(z,r)$ be an empty dual circumball of Q. It is the smallest empty dual circumball of Q if and only if*

- *it is the smallest dual circumball of $\partial B'_F(z,r) \cap X$ and*
- $\text{Front}\big(B'_F(z,r)\big) \subset Q$.

Proof: Apply the theorem with $S = X$. ∎

Definition 4.34 *Let (X,F) be a Bregman point cloud in general position. We define $\text{Upp}(X,F)$ as the set of simplices $\big\{P \in \text{Del}_F(X) \mid B'_F(P \mid P) = B'_F(P \mid X)\big\}$.*

Corollary 4.35 ***(Delaunay Morse Function Corollary)***
The Delaunay radius function of a Bregman point cloud (X,F) in general position is a generalized discrete Morse function. Its generalized discrete gradient consists the intervals $[\text{Front}\big(B'_F(P,P)\big), P]$, $P \in \text{Upp}(X,F)$.

Proof: The arguments are very similiar to the proof of the Čech Morse Function Corollary 4.31. We give the proof for the sake of completeness.

First, we show that $\{[\text{Front}\left(B'_F(P,P)\right), P]\}_{P \in \text{Upp}(X,F)}$ is a partition of $\text{Del}_F(X)$.

Let $Q \in \text{Del}_F(X)$ be a simplex. By corollary 4.14, the smallest empty dual circumball $B'_F(Q \mid X)$ exists.

By corollary 4.33, $B'_F(Q \mid X)$ is the smallest dual circumball of $P := \partial B'_F(Q \mid X) \cap X$ and $\text{Front}\left(B'_F(Q \mid X)\right) \subset Q$.
Since $\text{Front}\left(B'_F(Q \mid X)\right) \subset P$, we can use corollary 4.33 again and see that $B'_F(Q \mid X)$ is the smallest empty dual circumball of P. It means that $P \in \text{Upp}(X, F)$ and $Q \in [\text{Front}\left(B'_F(Q \mid X)\right), P]$.

Let $P' \in \text{Upp}(X, F)$ be an other upper bound such that $Q \in [\text{Front}\left(B'_F(P' \mid X)\right), P']$. By corollary 4.33, the smallest empty dual circumball $B'_F(P' \mid X)$ is the smallest dual circumball of $\partial B'_F(P' \mid X) \cap X$ and $\text{Front}\left(B'_F(P' \mid X)\right) \subset P'$.

Since $\text{Front}\left(B'_F(P' \mid X)\right) \subset Q \subset P'$, from Corollary 4.33 it follows that $B'_F(P' \mid X)$ is the smallest empty dual circumball of Q. This implies that $P = \partial B'_F(Q \mid X) \cap X = \partial B'_F(P' \mid X) \cap X$. The assumption $(G2)$ implies that $P = P'$. In conclusion, the intervals are disjoint.

Let $R, Q \in \text{Del}_F(X)$ such that $R \subset Q$. By corollary 4.17, $\varrho_F^{\text{Del}}(Q)$ is the radius of the smallest empty dual circumball of Q. This implies that $\varrho_F^{\text{Del}}(R) \leq \varrho_F^{\text{Del}}(Q)$.

If R and Q belong to the same interval then both $B'_F(R \mid X)$ and $B'_F(Q \mid X)$ are the smallest dual circumball of $\partial B'_F(R \mid X) \cap X = \partial B'_F(Q \mid X) \cap X$, indeed $\varrho_F^{\text{Del}}(R) = \varrho_F^{\text{Del}}(Q)$.

If $\varrho_F^{\text{Del}}(R) = \varrho_F^{\text{Del}}(Q)$, then the smallest empty dual circumballs of R and Q are the same, since $B'_F(Q, X)$ is an empty dual circumball of R and there are no other empty dual circumball of R with smaller radius. From this follows that $\partial B'_F(Q \mid X) \cap X = \partial B'_F(R \mid X) \cap X$ and therefore they belong to the same interval. ∎

4.3. Radius function algorithms

This section concludes the results of the previous one. Here, (X, F) stands for a Bregman point cloud in general position.

We start with Čech complexes. By the Čech Morse Function Corollary 4.31, $\varrho_F^{\text{Čech}} : \Delta(X) \to \mathbb{R}$ is a generalized discrete Morse function. A simplex $P \in \Delta(X)$ is a lower bound of an interval if and only if the smallest including dual Bregman ball and the smallest dual circumball of P are the same. Using the `CircumBall` algorithm, we can compute the center and the radius $\varrho_F^{\check{\mathrm{C}}\mathrm{ech}}(P)$ of the smallest including dual circumball of P. The upper bound of the interval is easily computable, we just collect the points of X lying in the smallest dual circumball of P. Once we have the lower and the upper bound of the interval, we mark each simplex of the interval with $\varrho_F^{\check{\mathrm{C}}\mathrm{ech}}(P)$. This can be done with the following algorithm.

Algorithm 2: `MarkTheInterval`

```
Input  : an interval [P, R] of simplices and a real number ϱ
let 𝒬 be the set containing all subsets of R \ P;
forall Q ∈ 𝒬 do
    mark P ∪ Q with ϱ;
```

It is still a question how we can decide whether a simplex is a lower bound. Our strategy is quite simple, we consider the simplicies in the order of increasing dimension. If we find an unmarked simplex, it must be a lower bound, so we mark the whole interval. An implementation of the `CechRadiusFunction` algorithm can be find in appendix A.2.

Algorithm 3: `CechRadiusFunction`

```
Input  : a Bregman point cloud (X, F)
let Δ(X) be the set containing all nonempty subsets of X;
for i = 0 to dim Δ(X) do
    foreach P ∈ Δ(X) with dim P = i do
        if P is unmarked then
            (z, ϱ_F^Čech(P)) ⟵ CircumBall (P) ;
            R ⟵ P ;
            forall x ∈ X \ P do
                if D_F(x‖z) ≤ ϱ_F^Čech(P) then
                    R ⟵ R ∪ {x};
            MarkTheInterval (P, R, ϱ_F^Čech(P));
Output: filtered simplicial complex (Δ(X), ϱ_F^Čech)
```

The case of Delaunay complexes is similar. By the Delaunay Morse Function Corollary 4.35, $\varrho_F^{\mathrm{Del}} : \mathrm{Del}_F(X) \to \mathbb{R}$ is a generalized discrete Morse function. In this case, we know the upper bounds of the intervals. Using the `CircumBall` algorithm we can compute $\varrho_F^{\mathrm{Del}}(P)$ for each $P \in \mathrm{Upp}(X, F)$. To compute the lower bound, we must determine $\mathrm{Front}\big(B'_F(P, P)\big)$. In order to do so, we modify the `CircumBall` algorithm.

Algorithm 4: `CircumBallFront`

Input : $P = \{p_0, \dots p_k\}$
let $\mathcal{A}$ be the affine hull of the points $(p, F(p))$, $p \in P$;
find $(q, \psi) \in \mathcal{A}$ minimizing $\mathbb{R}^{n+1} \to \mathbb{R}, (a, \alpha) \mapsto F(a) - \alpha$ over $\mathcal{A}$;
find $\lambda_0, \dots, \lambda_k \in \mathbb{R}$ such that $q = \sum_{i=0}^{k} \lambda_i p_i$;
$Q \longleftarrow \emptyset$;
for $i = 0$ **to** k **do**
  if $\lambda_i > 0$ **then**
    $Q \longleftarrow Q \cup \{p_i\}$;
Output: $(q, \psi - F(q), Q)$

With Delaunay complexes, we have another difficulty, namely not every subset of X is in $\mathrm{Del}_F(X)$. Using the `WeightedDelTriangulation` function, we can solve this problem. After $\mathrm{Del}_F(X)$ is computed, we consider the simplicies in order of decreasing dimension and mark the unmarked simplices. A Matlab implementation of the following algorithm is given in appendix A.5.

Algorithm 5: `DelaunayRadiusFunction`

Input : a Bregman point cloud (X, F)
foreach $x \in X$ **do**
  $w_x \longleftarrow \|x\|^2 - 2F(x)$;
$\mathrm{Del}_F(X) \longleftarrow$ `WeightedDelTriangulation` $\big(X, (w_x)_{x \in X}\big)$;
for $i = \dim \mathrm{Del}_F(X)$ **to** 0 **do**
  foreach $P \in Del_F(X)$ *with* $\dim P = i$ **do**
    if P *is unmarked* **then**
      $(z, \varrho_F^{\mathrm{Del}}(P), Q) \longleftarrow$ `CircumBallFront` (P) ;
      `MarkTheInterval` $(Q, P, \varrho_F^{\mathrm{Del}}(P))$;
Output: filtered simplicial complex $\big(\mathrm{Del}_F(X), \varrho_F^{\mathrm{Del}}\big)$

For the `CechRadiusFunction` algorithm, we get the triangulation almost for free. The Delaunay triangulation requires some preliminary computations. In corollary 4.23, we have seen that the dimension of the Delaunay complex is bounded above by the dimension of the including space if the points are in general position. Thus, it may be worth doing some preliminary job to make the computation faster.

To test this, we did some experiments using synthetic data sets. One can find the Matlab code in appendix A.7. The functions `CechPersistenceDiagram` (see appendix A.3) and the `DelaunayPersistenceDiagram` (see appendix A.6) compute the persistence diagram of a Bregman point cloud. They use the `CechRadiusFunction` resp. `DelaunayRadiusFunction` functions for the computation of the filtered simplicial complexes. After that, they apply functions from the JavaPlex package [15] to get the persistence diagrams. Figure 4.3 concludes the results.

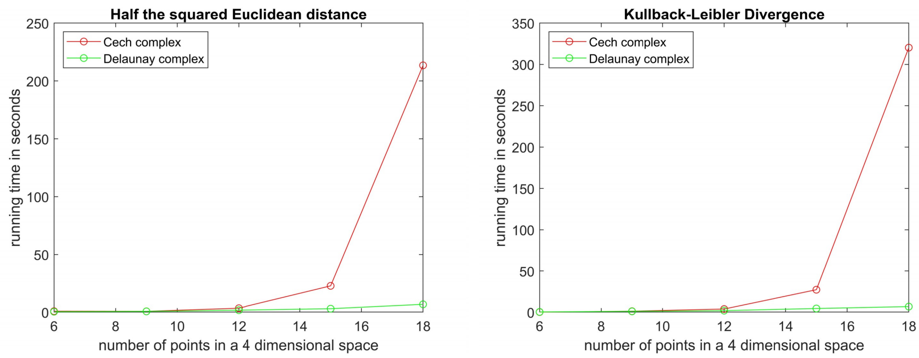


Figure 18: Performance of the `CechPersistenceDiagram` and the `DelaunayPersistenceDiagram` functions

In conclusion, we can say that it is worth to compute the Delaunay triangulation. The running time of the computation of the persistence diagram increases much slower, if we use the `DelaunayRadiusFunction` algorithm. It must be noticed that our implementation of the `DelaunayRadiusFunction` algorithm only works if the dimension of the containing space is smaller than the number of points in the Bregman point cloud minus two and Ω^* is the whole space $\mathbb{R}^n$.

5. Stability of persistence diagrams

In real-life applications, it is almost impossible to get a dataset without any noise. The measured data approximates the real data, but due to experimental errors, they will never be the same, after each measurement, the dataset will look a bit different. A good descriptor of a dataset should be closer to the descriptor of the real data if the measured data approximate better the real one. For a point cloud measured with a metric function, this property is ensured by the following theorem.

Theorem 5.1 ***[31, Thm. 5.2.]*** *Let* (M, d_M),(N, d_N) *be totally bounded metric spaces. Then*

$$\mathrm{d_B}\Big(\mathrm{dgm}\left(\mathrm{H}_n(\mathrm{Rips}(M,\bullet))\right), \mathrm{dgm}\left(\mathrm{H}_n(\mathrm{Rips}(N,\bullet))\right)\Big) \leq 2\,\mathrm{d_{GH}}(M,N),$$

$$\mathrm{d_B}\Big(\mathrm{dgm}\left(\mathrm{H}_n(\check{\mathrm{C}}\mathrm{ech}(M,\bullet))\right), \mathrm{dgm}\left(\mathrm{H}_n(\check{\mathrm{C}}\mathrm{ech}(N,\bullet))\right)\Big) \leq 2\,\mathrm{d_{GH}}(M,N)$$

for all $n \in \mathbb{N}_0$.

Here, the *bottleneck distance* $\mathrm{d_B}$ is used to compare persistence diagrams and we use the *Gromov-Hausdorff distance* $\mathrm{d_{GH}}$ [31, Section 4.2] for metric spaces.

If we want to trust the persistence diagrams of Bregman point clouds, we need something similar. As far as I know, there is no such theorem in the literature. We present the first statement in this direction. As our result is much weaker than the one for metric spaces, further improvements are still desirable.

5.1. Metrics

In this section, we introduce a dissimilarity measure for each collection of objects in the usual TDA workflow. Following [12], we use the *bottleneck distance* for persistence diagrams.

Definition 5.2 *Let $\mathcal{D}, \mathcal{D}'$ be two persistence diagrams. The* <u>*bottleneck distance*</u> *between $\mathcal{D}$ and $\mathcal{D}'$ is defined as:*

$$\mathrm{d_B}(\mathcal{D}, \mathcal{D}') := \inf_{\eta : \mathcal{D} \xrightarrow{1:1} \mathcal{D}'} \sup_{d \in \mathcal{D}} \|d - \eta(d)\|_\infty$$

where the infimum is taken over all bijections $\eta : \mathcal{D} \xrightarrow{1:1} \mathcal{D}'$ between the multisets $\mathcal{D}$ and $\mathcal{D}'$.

By [12, Chapter VIII.2], the bottleneck distance between persistence diagrams satisfies the axioms of a metric function.

Notice that by definition 2.12, a persistence diagram includes the points of the diagonal with infinite multiplicity and it has only a finite number of points above the diagonal. The bottleneck distance describes the best matching between two persistence diagrams. If a point lying above the diagonal is unmatched, we match it with the nearest point on the diagonal. To get an intuitive feeling, look at the figure below. For further details we recommend [13, Chapter 3.1].

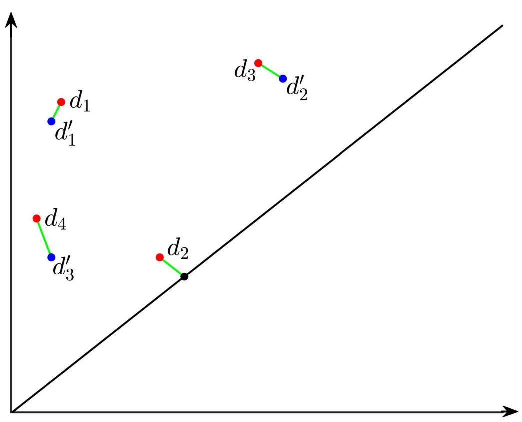
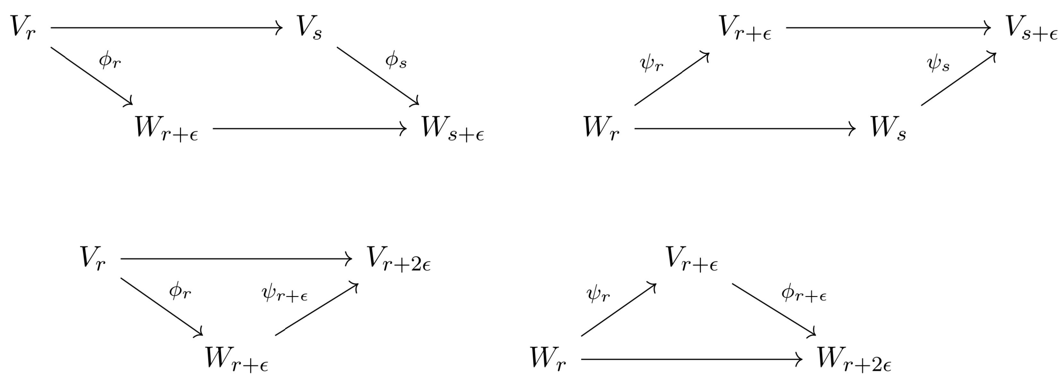


Figure 19: best matching between $\mathcal{D}$ and $\mathcal{D}'$

Using the JavaPlex package, one can easily write a Matlab code that computes the bottleneck distance between two persistence diagrams. For completeness, this code is given in appendix A.8.

We follow [13] to define a distance between two persistence $\mathbf{k}$-modules.

Definition 5.3 *Let $V_\bullet, W_\bullet$ be two persistence $\mathbf{k}$-modules and $\epsilon \geq 0$. An* <u>*ϵ-interleaving*</u> *between $V_\bullet$ and $W_\bullet$ is given by two families of homomorphisms $\{\phi_r : V_r \to W_{r+\epsilon}\}_{r\in[0,\infty)}$ and $\{\psi_r : W_r \to V_{r+\epsilon}\}_{r\in[0,\infty)}$ such that the following diagrams commute for every $r, s \in [0,\infty)$ with $r \leq s$:*



Definition 5.4 *The* <u>*interleaving distance*</u> *between two persistence $\mathbf{k}$-modules $V_\bullet$ and $W_\bullet$ is defined as*

$$\mathrm{d_I}(V_\bullet, W_\bullet) := \inf\{\epsilon \geq 0 \mid \text{ there is an } \epsilon\text{-interleaving between } V_\bullet \text{ and } W_\bullet\ \}$$

By [32, proposition 4.3], the interleaving distance satisfies the triangle inequality. It is easy to see that it is symmetric and non-negative. If we restrict it to pointwise finite-dimensional persistence modules, then the interleaving distance is zero if and only if the modules are isomorphic. Thus, in this case, we have a metric function. This last property is implied by the Isometry Theorem.

Theorem 5.5 ***(Isometry Theorem) [32, Thm. 4.11]*** *Let $V_\bullet$ and $W_\bullet$ be pointwise finite-dimensional persistence $\mathbf{k}$-modules, then it holds that*

$$\mathrm{d_B}\big(\mathrm{dgm}(V_\bullet), \mathrm{dgm}(W_\bullet)\big) = \mathrm{d_I}(V_\bullet, W_\bullet).$$

Last but not least, we introduce a dissimilarity measure for Bregman point clouds. We are interested in the question of what happens if we add some noise to the dataset. That is why we compare the underlying set of the Bregman point clouds using the usual Euclidean distance and not the Bregman divergence. Our constructions are based on the definitions in [31].

Definition 5.6

a) *Let X and Y be two sets. We call a subset C of the Cartesian product $X \times Y$ a correspondence if the canonical projection maps are surjective. In this case, we write $C : X \rightrightarrows Y$.*

b) *Let X, Y be subsets of $\mathbb{R}^n$. The distortion of a correspondence $C : X \rightrightarrows Y$ is defined as* $\operatorname{dis}(C) := \sup\{\|x - y\| \mid (x, y) \in C\}$.

b) *Let $(X, F), (Y, F)$ be Bregman point clouds. We define the Gromov-Hausdorff distance as*

$$\mathrm{d_{GH}}(X, Y) := \min\{\operatorname{dis}(C) \mid C : X \rightrightarrows Y \textit{is a correspondence} \}.$$

5.2. A stability result

We are going to present our stability result. Throughout this section, (X, F) denotes a Bregman point cloud. We consider this as the real data, i.e. the data without any noise. The Bregman point cloud (Y, F) will play the role of the noisy, measured data.

Further, we want to assume that our measured data is not completely wrong; it is not too far away from X, i.e $\|x - y\| \leq \max\{\|x - x'\| \mid x, x' \in X\} =: d$ for all $x \in X$, $y \in Y$.

In the following, we construct a convex, compact subset $\mathcal{X} \subset \Omega = \operatorname{dom}(F)$ such that $\|x - y\| \leq d$ for all $x \in X$, $y \in \mathcal{X}$. The union of the Euclidean balls centered at the points of X with radius d may lie outside of Ω. If we cut the union of these balls with Ω, it could happen, that the closure of this intersection is not contained in Ω. This phenomenon is illustrated in the next figure.

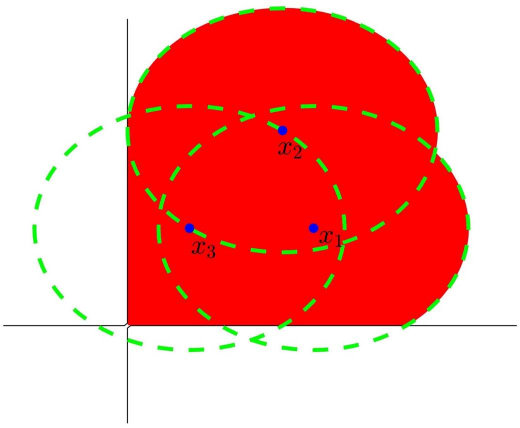


Figure 20: union of the Euclidean balls lies outside of $\Omega = \mathbb{R}^2_+$

To avoid the problems with the boundary of Ω, we put a small tube around it. Let us define $T := \bigcup_{b \in \mathrm{bd}(\Omega)} B_t(b)$ where $t := \frac{1}{2} \inf\{\|b - x\| \mid b \in \mathrm{bd}(\Omega), x \in X\}$.

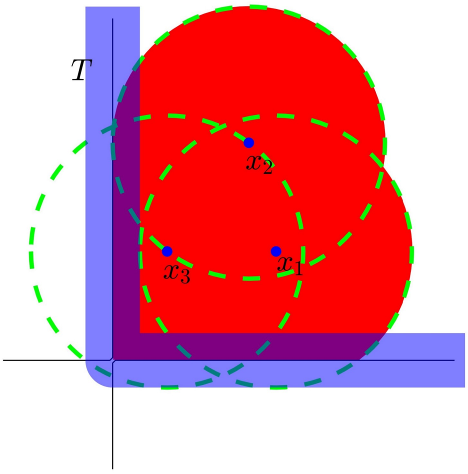


Figure 21: a small tube T around the boundary of $\Omega = \mathbb{R}^2_+$

By construction $\overline{\Omega \setminus T}$ is a closed subset of Ω. Thus, $\left(\bigcup_{x \in X} B_d(x)\right) \cap \left(\overline{\Omega \setminus T}\right)$ is a compact subset of Ω. Since Ω is convex, the set $\mathcal{X} := \mathrm{Conv}\left(\left(\bigcup_{x \in X} B_d(x)\right) \cap \left(\overline{\Omega \setminus T}\right)\right)$ is contained in Ω. One can easily show that $\mathcal{X}$ is a compact set.

In the rest of this section, we assume that the measured data Y is contained in $\mathcal{X}$. The choice of $\mathcal{X}$ is quite arbitrary. In fact, every convex, compact subset of Ω containing X and Y would do the job.

Theorem 5.7 *Assume that the partial derivatives of F are continuously differentiable and $Y \subset \mathcal{X}$. Then there are constants $\alpha, \beta \in \mathbb{R}$ depending only on (X, F) such that $\mathrm{d_{GH}}(X, Y) \leq \epsilon$ implies*

$$\mathrm{d_B}\left(\mathrm{dgm}\left(\mathrm{H}_n(\check{\mathrm{C}}\mathrm{ech}_F(X, \bullet))\right), \mathrm{dgm}\left(\mathrm{H}_n(\check{\mathrm{C}}\mathrm{ech}_F(Y, \bullet))\right)\right) \leq \epsilon^2 \cdot \alpha + \epsilon \cdot \beta$$

for all $n \in \mathbb{N}_0$.

The functions from example 3.4 satisfy the assumption of the theorem. Thus, in these cases, the bottleneck distance between the persistence diagrams converges to zero if the distance between X and Y converges to zero. We can conclude that the information of the persistence diagram from a noisy dataset describes the persistence diagram of the real dataset better if we make a better measurement.

The proof of the theorem is quite technical. We split it into several lemmata. The key ingredient will be the *Dowker interleaving lemma* from [31]. First, we recall the definition of a *Dowker complex.*

Definition 5.8 *Let L, W be two sets and $\Lambda : L \times W \to \mathbb{R}$ a function. The Dowker complex of Λ at scale $r \geq 0$ is the simplicial complex:*

$$\operatorname{Dow}(\Lambda, r) := \{P \subset L \mid \exists w \in W \forall l \in P : \Lambda(l, w) \leq r\}.$$

Definition 5.9 *Let L, L', W, W' be sets with functions $\Lambda : L \times W \to \mathbb{R}$ $\Lambda' : L' \times W' \to \mathbb{R}$ and let $C : L \rightrightarrows L'$, $D : W \rightrightarrows W'$ be two correspondences. We define the distortion of the pair C, D as*

$$\operatorname{dis}(C, D) := \sup_{(l,l') \in L} \sup_{(w,w') \in D} \mid \Lambda(l, w) - \Lambda'(l', w') \mid .$$

Lemma 5.10 ***(Dowker interleaving)*** *Let L, L', W, W' be sets with functions $\Lambda : L \times W \to \mathbb{R}$ and $\Lambda' : L' \times W' \to \mathbb{R}$. If $C : L \rightrightarrows L'$ and $D : W \rightrightarrows W'$ are correspondences such that $\epsilon \geq \operatorname{dis}(C, D)$ then there is an ϵ-interleaving between the persistence* **k***-modules* $\mathrm{H}_n\left(\operatorname{Dow}(\Lambda, \bullet)\right)$ *and* $\mathrm{H}_n\left(\operatorname{Dow}(\Lambda', \bullet)\right)$ *for all $n \in \mathbb{N}_0$.*

To show theorem 5.7, we want to use the Dowker interleaving lemma. The choice of L, L' and Λ, Λ' is quite obvious. But we have to choose W, W' carefully. To build the Čech complex we only need the Chernoff points of the simplices i.e. the centers of the smallest including dual balls.

Definition 5.11 *For a Bregman point cloud (U, F) we define*

$$W_U := \left\{c_F(P) \mid P \in \Delta(U)\right\}$$

as the set containing the Chernoff points of the simplicies in $\Delta(U)$.

Proposition 5.12 *Let $\mathcal{X}$ be defined as at the beginning of this section and assume that $Y \subset \mathcal{X}$. Then W_X and W_Y are contained in $\mathcal{X}$.*

Proof: By lemma 3.54, the center of the smallest including dual ball of each simplex $P \in \Delta(X)$, $Q \in \Delta(Y)$ is contained in the convex hull of P resp. Q. Since $\mathcal{X}$ is convex, it contains $\operatorname{Conv}(P)$ and $\operatorname{Conv}(Q)$. ■

Lemma 5.13 *Let $P = \{p_0, \ldots, p_k\}$ and $Q = \{q_0, \ldots, q_k\}$ subsets of Ω such that $\|p_i - q_i\| \leq \epsilon$ for some $\epsilon \geq 0$ and all $i = 0, \ldots, k$. Let z be the center of the smallest including dual ball of Q. Then $\|z\| \leq \epsilon + \max_{i=0,\ldots,k} \|p_i\|$*

Proof: By lemma 4.20, there are nonnegative real numbers $\lambda_0, \ldots, \lambda_k$ such that $z = \sum_{i=0}^{k} \lambda_i \cdot q_i$ and $\sum_{i=0}^{k} \lambda_i = 1$. From this it follows that

$$\|z\| = \|\sum_{i=0}^{k} \lambda_i \cdot (q_i - p_i + p_i)\| \leq \|\sum_{i=0}^{k} \lambda_i \cdot (q_i - p_i)\| + \|\sum_{i=0}^{k} \lambda_i \cdot p_i\| \leq \sum_{i=0}^{k} \lambda_i \|q_i - p_i\| + \sum_{i=0}^{k} \lambda_i \|p_i\|$$

We can overestimate each term in the first sum by $\lambda_i \epsilon$ and in the second sum by $\lambda_i \cdot \max_{i=0,\ldots,k} \|p_i\|$. Using that $\sum_{i=0}^{k} \lambda_i = 1$, the statement follows. ■

Corollary 5.14 *For all $w \in W_X$, the Euclidean length of w is bounded by $\max_{x \in X} \|x\|$.*

Proof: For every $P \subset X$, use the above lemma with $P = Q$ and $\epsilon = 0$ and notice that $\max_{p \in P} \|p\| \leq \max_{x \in X} \|x\|$. ■

Lemma 5.15

a) Every continuously differentiable function $f : \mathcal{X} \to \mathbb{R}$ is Lipschitz continuous, i.e. there is a $K \in \mathbb{R}_+$ such that $| f(x) - f(y) | \leq K \cdot \|x - y\|$ for all $x, y \in \mathcal{X}$.

b) $F : \mathcal{X} \to \mathbb{R}$ is Lipschitz continuous.

c) $\nabla F : \mathcal{X} \to \Omega^*$ is Lipschitz continuous, if the partial derivatives of F are continuously differentiable.

Proof: By definition $\mathcal{X}$ is a convex, compact set. Since f is differentiable, by the mean value theorem, for all $x, y \in \mathcal{X}$ there is a $z \in \mathcal{X}$ lying on the line segment between x and y such that

$$| f(x) - f(y) | \leq \|\nabla f(z)\| \cdot \|x - y\| \leq \sup_{w \in \mathcal{X}} \|\nabla f(w)\| \cdot \|x - y\|.$$

Since $\|\nabla f(\bullet)\| : \mathcal{X} \to \mathbb{R}$ is a continuous function on a compact set, $K := \sup_{w \in \mathcal{X}} \|\nabla f(w)\|$ is a real number. Indeed, f is a Lipschitz continuous function with Lipschitz constant K.

Convex differentiable functions are continuously differentiable. Thus, part b) follows immediately from part a).

By our assumption and by part a), the partial derivatives $\frac{\partial F}{\partial x_i} : \mathcal{X} \to \mathbb{R}$ $i = 1, \ldots, n$ are Lipschitz continuous. Thus, for all $i = 1, \ldots, n$ there is a $K_i \in \mathbb{R}$ such that $| \frac{\partial F}{\partial x_i}(x) - \frac{\partial F}{\partial x_i}(y) | \leq K_i \cdot \|x - y\|$ for all $x, y \in \mathcal{X}$. This implies

$$\|\nabla F(x) - \nabla F(y)\|^2 = \sum_{i=1}^{n} \big(\frac{\partial F}{\partial x_i}(x) - \frac{\partial F}{\partial x_i}(y)\big)^2 \leq \sum_{i=1}^{n} K_i^2 \cdot \|x - y\|^2$$

In conclusion, ∇F is Lipschitz continuous with Lipschitz constant $\sqrt{\sum_{i=1}^{n} K_i^2}$. ■

Lemma 5.16 *Let $W := W_X \cup W_Y$. Assume that the partial derivatives of F are continuously differentiable and $Y \subset \mathcal{X}$. There are constants $\alpha, \beta \in \mathbb{R}$ depending only on (X, F) such that $\max_{w \in W} \|\nabla F(w)\| \leq \alpha \cdot \epsilon + \beta$ for all $\epsilon \geq \mathrm{d_{GH}}(X, Y)$.*

Proof: Since W is the union of W_X and W_Y, we have

$$\max_{w \in W} \|\nabla F(w)\| = \max\big\{ \max_{w \in W_X} \|\nabla F(w)\|, \max_{z \in W_Y} \|\nabla F(z)\|\big\}.$$

By definition, for all $z \in W_Y$ there is a $Q = \{q_0, \dots, q_k\} \in \Delta(Y)$ such that z is the center of the smallest including dual ball of Q.

Since $\epsilon \geq \mathrm{d_{GH}}(X, Y)$, there are $p_0, \dots, p_k \in X$ such that $\|p_i - q_i\| \leq \epsilon$ for all $i = 0, \dots, k$. Denote c_z the center of the smallest including dual ball of $P := \{p_0, \dots, p_k\}$.

From lemma 5.13 it follows that

$$\|z - c_z\| \leq \|z\| + \|c_z\| \leq \epsilon + \max_{x \in X} \|x\| + \|c_z\| \leq \epsilon + \max_{x \in X} \|x\| + \max_{c \in W_X} \|c\|.$$

By lemma 5.15, ∇F is Lipschitz continuous. Thus, there is a $K > 0$ independent from Y such that

$$\begin{aligned}\|\nabla F(z)\| = \|\nabla F(z) - \nabla F(c_z) + \nabla F(c_z)\| &\leq \|\nabla F(z) - \nabla F(c_z)\| + \|\nabla F(c_z)\| \leq \\ \leq K \cdot \|z - c_z\| + \|\nabla F(c_z)\| &\leq K \cdot \big(\epsilon + \max_{x \in X} \|x\| + \max_{c \in W_X} \|c\|\big) + \max_{c \in W_X} \|\nabla F(c)\|.\end{aligned}$$

From $K > 0$ it follows that

$$\|\nabla F(w)\| \leq K \cdot \big(\epsilon + \max_{x \in X} \|x\| + \max_{c \in W_X} \|c\|\big) + \max_{c \in W_X} \|\nabla F(c)\|.$$

for every $w \in W_X$. In conclusion, we have that

$$\max_{w \in W} \|\nabla F(w)\| \leq K \cdot \big(\epsilon + \max_{x \in X} \|x\| + \max_{c \in W_X} \|c\|\big) + \max_{c \in W_X} \|\nabla F(c)\|.$$

The scalars $\alpha := K$, $\beta := K \cdot \max_{x \in X} \|x\| + K \cdot \max_{c \in W_X} \|c\| + \max_{c \in W_X} \|\nabla F(c)\|$ depend only on (X, F). ■

Now we are ready to show theorem 5.7.

Proof of Theorem 5.7:
If $d_{GH}(X, Y) \leq \epsilon$, then $C := \{(x, y) \in X \times Y \mid \|x - y\| \leq \epsilon\}$ is a correspondence.

Let $W = W' := W_X \cup W_Y$ and $\Lambda := D_F \upharpoonright_{X \times W}$, $\Lambda' := D_F \upharpoonright_{Y \times W}$. If we compare the definition of a Dowker complex 5.2 and corollary 4.17, we see that $\check{\mathrm{C}}\mathrm{ech}_F(X, r) = \mathrm{Dow}(\Lambda, r)$ and $\check{\mathrm{C}}\mathrm{ech}_F(Y, r) = \mathrm{Dow}(\Lambda', r)$ for all $r \geq 0$.

Obviously, $D := \{(w, w) \in W \times W \mid w \in W\}$ is a correspondence. By the Dowker interleaving lemma 5.10, there is a δ-interleaving between $\mathrm{H}_n\left(\check{\mathrm{C}}\mathrm{ech}_F(X, \bullet)\right)$ and $\mathrm{H}_n\left(\check{\mathrm{C}}\mathrm{ech}_F(Y, \bullet)\right)$ for all $n \in \mathbb{N}_0$, if $\delta \geq dis(C, D)$.

Using the above lemmata, we can compute an upper bound of $\mathrm{dis}(C, D)$.

$$
\begin{aligned}
\mathrm{dis}(C, D) &= \sup_{(x,y) \in C} \sup_{w \in W} \mid D_F(x\|w) - D_F(y\|w) \mid = \\
&= \sup_{(x,y) \in C} \sup_{w \in W} \mid F(x) - F(y) + \langle \nabla F(w), y - x \rangle \mid \leq \\
&\leq \sup_{(x,y) \in C} \sup_{w \in W} \mid F(x) - F(y) \mid + \mid \langle \nabla F(w), y - x \rangle \mid
\end{aligned}
$$

Since F is Lipschitz continuous on $\mathcal{X}$ (lemma 5.15), there is a constant K such that

$$\mid F(x) - F(y) \mid \leq K \cdot \|x - y\| \leq K \cdot \epsilon \text{ for every } (x, y) \in C.$$

Using the Cauchy-Schwarz inequality and lemma 5.16, we have for every $(x, y) \in C$, $w \in W$ that

$$\mid \langle \nabla F(w), y - x \rangle \mid \leq \|\nabla F(w)\| \cdot \|y - x\| \leq (\alpha \cdot \epsilon + \beta) \cdot \epsilon$$

for some constants $\alpha, \beta \in \mathbb{R}$ independent from Y.

In conclusion: $\mathrm{dis}(C, D) \leq K \cdot \epsilon + (\alpha \cdot \epsilon + \beta) \cdot \epsilon = (K + \beta) \cdot \epsilon + \alpha \cdot \epsilon^2$. Thus, there is a $\delta := (K + \beta) \cdot \epsilon + \alpha \cdot \epsilon^2$ interleaving between the persistence **k**-modules.

This implies that $d_I\Big(\mathrm{H}_n\left(\check{\mathrm{C}}\mathrm{ech}_F(X, \bullet)\right), \mathrm{H}_n\left(\check{\mathrm{C}}\mathrm{ech}_F(Y, \bullet)\right)\Big) \leq \delta$.

From the Isometry Theorem 5.5 it follows that
$d_B\Big(\mathrm{dgm}\left(\mathrm{H}_n(\check{\mathrm{C}}\mathrm{ech}_F(X, \bullet))\right), \mathrm{dgm}\left(\mathrm{H}_n(\check{\mathrm{C}}\mathrm{ech}_F(Y, \bullet))\right)\Big) \leq \delta.$ ■

Using the algorithms discussed in the previous chapter, we do some experiments with synthetic data sets. First, we compute a random Bregman point cloud. After that, we create a sequence of Bregman point clouds converging to the first one in the Hausdorff distance. For each Bregman point cloud in the sequence, we call the function `CechPersistenceDiagram` (see appendix A.3) to compute the persistence diagram, and we call `BottleneckDistance` to compute the bottleneck distance. To measure the Hausdorff distance, we use the function `HausdorffDist` [33]. The Matlab code of the experiment is given in appendix A.9. Figure 5.2 concludes the results.

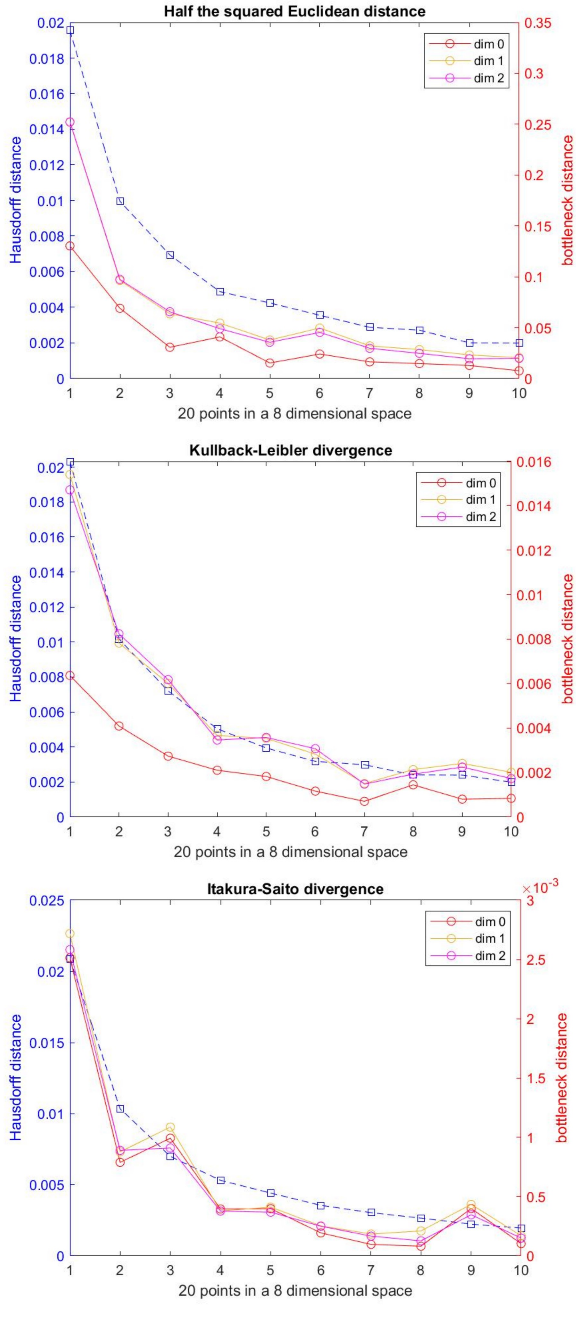


Figure 22: Experiment - Stability of persistence diagrams

We can clearly observe in figure 5.2 what theorem 5.7 tells us. If the Bregman point clouds are converging in the Hausdorff distance, then the corresponding persistence diagrams will converge in the bottleneck distance.

6. Künneth-type formula

In algebraic topology Künneth-type formulas describe a relation between the homology of the product space and the homology of its factors.

Theorem 6.1 ***(topological Künneth formula)***
Let T and T' be topological spaces and let R be a principal ideal domain. There exists a natural short exact sequence

$$0 \to \bigoplus_{i+j=n} \big(\mathrm{H}_i(T;R) \otimes_R \mathrm{H}_j(T';R)\big) \to \mathrm{H}_n(T \times T';R) \to$$

$$\to \bigoplus_{i+j=n} Tor_R\big(\mathrm{H}_i(T;R), \mathrm{H}_{j-1}(T';R)\big) \to 0$$

which splits, but not naturally.

One can find the proof of the theorem in several introductory books on algebraic topology, see e.g. [16, Theorem 3B.6] or [34, Theorem 59.3].

If we take the coefficients from a field $\mathbf{k}$, the Tor term is zero. Indeed, we have a natural isomorphism

$$\bigoplus_{i+j=n} \big(\mathrm{H}_i(T;\mathbf{k}) \otimes_{\mathbf{k}} \mathrm{H}_j(T';\mathbf{k})\big) \xrightarrow{\sim} \mathrm{H}_n(T \times T';\mathbf{k})$$

In this case, we can understand the homology of the product spaces, which can be quite complicated in terms of simpler objects. It would be nice to have a similar result for persistent homology. Some recent papers considered this question. In [35], Gakhar and Perea showed a Künneth formula for Vietoris-Rips complexes.

Theorem 6.2 ***[35, Corollary 4.6]***
Let (M, d_M),(N, d_N) be finite metric spaces. If we equip the Cartesian product $M \times N$ with the maximum metric, then there is an isomorphism

$$\mathrm{H}_n\left(\mathrm{Rips}(M \times N, \bullet)\right) \cong \bigoplus_{i+j=n} \mathrm{H}_i\left(\mathrm{Rips}(M, \bullet)\right) \otimes \mathrm{H}_j\left(\mathrm{Rips}(N, \bullet)\right)$$

for all $n \in \mathbb{N}_0$.

In [36, Theorem 4.1], Lim et al. gave an alternative proof of theorem 6.2. Carlsson and Filippenko [37] studied a similar question, but they equipped the Cartesian product with the sum metric.

Theorem 6.3 ***[37, Prop. 4.3]*** *Let (M, d_M) and (N, d_N) be finite metric spaces and $n \in \mathbb{N}_0$. If we equip the Cartesian product $M \times N$ with the sum metric $d_M + d_N$, then there is a short exact sequence*

$$0 \to \bigoplus_{i+j=n} \mathrm{PH}_i(M, l_M) \otimes_{\mathbf{k}[\mathbb{R}_+]} \mathrm{PH}_j(N, l_N) \to \mathrm{PH}_n(M \times N, l_M + l_N) \to$$

$$\to \bigoplus_{i+j=n} Tor_1\big(\mathrm{PH}_i(M, l_M), \mathrm{PH}_{j-1}(N, l_N)\big) \to 0$$

which is natural with respect to distance non-increasing maps $(M, d_M) \to (M', d_{M'})$ and $(N, d_N) \to (N', d_{N'})$. Moreover, the sequence splits, but not naturally.

Carlsson and Filippenko used a slightly different notation. Since we will not use this notation, we refer to the original paper [37, Section 4] for a detailed explanation. For our purposes, it is enough to note that $\mathrm{PH}_*^{(r)}(M, l_M) = \mathrm{H}_*\left(\mathrm{Rips}(M, r)\right)$ for all $r \geq 0$. Unfortunately, the middle term in the sequence in theorem 6.4 is in general not isomorphic to $\mathrm{PH}_n(M \times N, l_{M\times N})$, where we equip $M \times N$ with the sum metric $d_M + d_N$. For dimension $n \geq 2$, Carlsson and Filippenko gave an upper bound for the interleaving distance between $\mathrm{PH}_n^\bullet(M \times N, l_M + l_N)$ and $\mathrm{PH}_n^\bullet(M \times N, l_{M\times N})$.

Theorem 6.4 *Let (M, d_M) and (N, d_N) be finite metric spaces. Consider the Cartesian product $M \times N$ equipped with the sum metric $d_M + d_N$.*

a) For $n = 0, 1$, there is an isomorphism

$$\mathrm{PH}_n(M \times N, l_M + l_N) \cong \mathrm{PH}_n(M \times N, l_{M\times N}).$$

b) For every $n \in \mathbb{N}_0$, the interleaving distance between the persistence $\mathbf{k}$-modules $\big(r \mapsto \mathrm{PH}_n^{(r)}(M \times N, l_M + l_N)\big)_{r\geq 0}$ and $\big(r \mapsto \mathrm{PH}_n^{(r)}(M \times N, l_{M\times N})\big)_{r\geq 0}$ is less or equal than $\min\big(\mathrm{diameter}(M), \mathrm{diameter}(N)\big)$, where the diameter is given by $\mathrm{diameter}(M) := \max\{d_M(m_0, m_1) \mid m_0, m_1 \in M\}$.

Proof: Part *a*) is [37, Theorem 4.5.], Part *b*) is [37, Theorem 4.9.] ■

Using a statement from [35], we will give a Künneth-type formula for Bregman point clouds in the style of the result of Carlsson and Filippenko. In fact, we consider an object that satisfies a Künneth formula, and we compute an upper bound of the bottleneck distance between the persistence diagram of this object and the persistence diagram of the product space.

6.1. Bregman divergence on the product space

First, we try to find the best way to define a Bregman divergence on the product space. During this section, $F : \Omega \to \mathbb{R}$ and $G : \Omega' \to \mathbb{R}$ stand for functions of Legendre type.

One can easily show that the Cartesian product $\Omega \times \Omega'$ is a nonempty open convex set. According to proposition 3.3, the sum $F + G$ is a function of Legendre type on $\Omega \times \Omega'$. Using proposition 3.26, we can characterize the corresponding Bregman divergence D_{G+H}.

This choice seems to be natural since our main examples 3.4 have this form. But is not clear, whether or not this is the only natural choice. That is why we consider the following general setting.

In order to equip the product space $\Omega \times \Omega'$ with a Bregman divergence using F and G, we need a function $f : \mathrm{im}(F) \times \mathrm{im}(G) \to \mathbb{R}$ such that

$$H := f\big(F(\bullet), G(\bullet)\big) : \Omega \times \Omega' \to \mathbb{R}, (x, y) \mapsto f\big(F(x), G(y)\big)$$

is a function of Legendre type.

Proposition 6.5 $\mathrm{im}(F) \times \mathrm{im}(G)$ *is a nonempty convex set.*

Proof: It is enough to show that $\mathrm{im}(F)$ and $\mathrm{im}(G)$ are nonempty open convex sets. They are clearly nonempty since F and G are well-defined functions and $\Omega \neq \emptyset$, $\Omega' \neq \emptyset$.

The epigraph of a convex function is a convex set. The image of a function is the projection of the epigraph onto its codomain. Since the projection map is linear, the image of a convex function is convex. ■

In the following, we give some conditions on f, which ensures that H is a function of Legendre type.

Proposition 6.6 *If f is convex and strictly increasing in both arguments, then H is strictly convex.*

Proof: For all $(x_1, y_1) \neq (x_2, y_2) \in \Omega \times \Omega'$ it holds that $x_1 \neq x_2$ or $y_1 \neq y_2$. We consider the case $x_1 \neq x_2$. The other case follows by a similar argument.

Since F is strictly convex, we have $F(t \cdot x_1 + (1-t) \cdot x_2) < tF(x_1) + (1-t)F(x_2)$ for all $t \in (0, 1)$. The convexity of G implies that $G(t \cdot y_1 + (1-t) \cdot y_2) \leq tG(y_1) + (1-t)G(y_2)$ for all $t \in (0, 1)$.

By our assumption, f is strictly increasing in both arguments. This implies that

$$\begin{aligned} H\big(t \cdot (x_1, y_1) + (1-t) \cdot (x_2, y_2)\big) &< f\big(tF(x_1) + (1-t)F(x_2), tG(y_1) + (1-t)G(y_2)\big) = \\ &= f\big(t \cdot \big(F(x_1), G(y_1)\big) + (1-t)\big(F(x_2), G(y_2)\big)\big). \end{aligned}$$

Now, we can use the convexity of f to conclude that
$H\big(t\cdot(x_1,y_1)+(1-t)\cdot(x_2,y_2)\big) < tH\big((x_1,y_1)\big)+(1-t)H\big((x_2,y_2)\big)$. ■

Proposition 6.7 *If f is differentiable, then H is differentiable.*

Proof: The Jacobi matrix of $(F,G):\Omega\times\Omega'\to\mathbb{R}\times\mathbb{R},(x,y)\mapsto\big(F(x),G(y)\big)$ at the point (x,y) is given by $\begin{pmatrix}\nabla F(x) & 0\\ 0 & \nabla G(y)\end{pmatrix}\in\mathbb{R}^{2\times(n+m)}$, where $\Omega\subset\mathbb{R}^n$ and $\Omega'\subset\mathbb{R}^m$.

The differentiability of F and G implies that the partial derivatives of (F,G) are continuous. From this it follows that (F,G) is differentiable. It is well-known that the composition of differentiable functions is differentiable. ■

Lemma 6.8 *Let f be convex, differentiable and strictly increasing in both arguments. Furthermore, we assume that for every sequence $\big((a_n,b_n)\big)_{n\in\mathbb{N}}$ in $\operatorname{im}(F)\times\operatorname{im}(G)$ it holds that*

$$\lim_{n\to\infty}\mid\frac{\partial f}{\partial a}(a_n,b_n)\mid>0 \text{ and } \lim_{n\to\infty}\mid\frac{\partial f}{\partial b}(a_n,b_n)\mid>0.$$

Then $H:\Omega\times\Omega',(x,y)\mapsto f\big(F(x),G(y)\big)$ is a function of Legendre type.

Proof: By proposition 6.6 and 6.7, it is enough to show, that H satisfies the condition $(L3)$ in definition 3.1.

Using the chain rule, we can compute the gradient of H at each point $(x,y)\in\Omega\times\Omega'$.

$$\nabla H(x,y)=\big(\tfrac{\partial f}{\partial a}\big(F(x),G(y)\big)\quad\tfrac{\partial f}{\partial b}\big(F(x),G(y)\big)\big)\cdot\begin{pmatrix}\nabla F(x) & 0\\ 0 & \nabla G(y)\end{pmatrix}=$$
$$=\big(\tfrac{\partial f}{\partial a}\big(F(x),G(y)\big)\cdot\nabla F(x),\quad\tfrac{\partial f}{\partial b}\big(F(x),G(y)\big)\cdot\nabla G(y)\big)$$

Let $\big((x_n,y_n)\big)_{n\in\mathbb{N}}$ be a sequence in $\Omega\times\Omega'$ such that $(x_n,y_n)\to(x,y)\in\operatorname{bd}(\Omega\times\Omega')$. As we have seen in proposition 3.3, we can assume without loss of generality that $(x,y)\in\operatorname{bd}(\Omega)\times\overline{\Omega'}$.

Using the above identity, we have:

$$\|\nabla H(x_n,y_n)\|^2=\|\frac{\partial f}{\partial a}\big(F(x_n),G(y_n)\big)\cdot\nabla F(x_n)\|^2+\|\frac{\partial f}{\partial b}\big(F(x_n),G(y_n)\big)\cdot\nabla G(y_n)\|^2\geq$$
$$\geq\mid\frac{\partial f}{\partial a}\big(F(x_n),G(y_n)\big)\mid^2\cdot\|\nabla F(x_n)\|^2.$$

Since F is a function of Legendre type and $x\in\operatorname{bd}(\Omega)$, $\|\nabla F(x_n)\|^2\to\infty$.

By assumption, it holds that $\lim_{n\to\infty}\mid\frac{\partial f}{\partial a}\big(F(x_n),G(y_n)\big)\mid^2>0$.

In conclusion, $\mid\frac{\partial f}{\partial a}\big(F(x_n),G(y_n)\big)\mid^2\cdot\|\nabla F(x_n)\|^2\to\infty$. ■

Example 6.9

a) *Let λ, μ be positive real numbers. The function $f : \mathbb{R} \times \mathbb{R} : (x, y) \mapsto \lambda \cdot a + \mu \cdot b$ satisfies the conditions of lemma 6.8.*
Indeed, $\Omega \times \Omega' \to \mathbb{R}, (x, y) \mapsto \lambda \cdot F(x) + \mu \cdot G(y)$ is a function of Legendre type.

b) *The function $f : \mathbb{R} \times \mathbb{R} : (a, b) \mapsto \exp(a) + \exp(b)$ is convex and differentiable. One can easily check that for every $(a, b) \in \mathbb{R}^2$ it holds that*

$$\frac{\partial f}{\partial a}(a, b) = \exp(a) \qquad \textit{and} \qquad \frac{\partial f}{\partial b}(a, b) = \exp(b).$$

Assume that the image of F and G are bounded from below (e.g. half the squared Euclidean norm or convex Shannon entropy). In this case, f satisfies the conditons of 6.8. Indeed,

$$\Omega \times \Omega' \to \mathbb{R}, (x, y) \mapsto \exp(F(x)) + \exp(G(y))$$

is a function of Legendre type.

6.2. Persistent Künneth formula for parametrised simplicial complexes

In this section, we recall the following statement from [35].

Theorem 6.10 *Let $K_\bullet$ and $K'_\bullet$ be parametrised simplicial complexes. For each $n \in \mathbb{N}_0$ there is an isomorphism of persistence* **k***-modules*

$$\mathrm{H}_n(K_\bullet \times K'_\bullet) \cong \bigoplus_{i+j=n} \mathrm{H}_i(K_\bullet) \otimes \mathrm{H}_j(K'_\bullet).$$

Befor we give a proof, we clarify what kind of product we use.

Definition 6.11 *Let K and K' be simplicial complexes. We define the product $K \times K'$ to be the smallest simplicial complex containing all Cartesian products $\sigma \times \sigma'$ for $\sigma \in K$ and $\sigma' \in K'$.*

One can generalize this definition to parametrised simplicial complexes.

Definition 6.12 *Let $K_\bullet$ and $K'_\bullet$ be parametrised simplicial complexes. Their product is defined to be the pointwise product i.e. $(K_\bullet \times K'_\bullet)_r := K_r \times K'_r$ and $(K_\bullet \times K'_\bullet)(r \leq t) := K_\bullet(r \leq t) \times K'_\bullet(r \leq t)$ for all $r, t \in [0, \infty)$ with $r \leq t$.*

Remark 6.13 *For general simplicial complexes $\mid K \times K' \mid$ and $\mid K \mid \times \mid K' \mid$ are not isomorphic. One can find a counterexample in [35, Chapter 3].*
Consider the standard 1-simplex $\big\{0, 1, \{0, 1\}\big\} =: K = K'$. Then $\mid K \times K' \mid$ is homeomorphic to the standard geometric 3-simplex and $\mid K \mid \times \mid K' \mid$ is homeomorphic to the unit square $[0, 1] \times [0, 1] \subset \mathbb{R}^2$.

Although $| K \times K' |$ and $| K | \times | K' |$ are not isomorphic, they have the same homotopy type. In order to show this, we need some more definitions.

Definition 6.14 *We call a simplicial complex K an ordered simplicial complex if there is a partial order $\preceq$ on $\mathrm{Vert}(K)$ such that each simplex in K is totally ordered with respect to $\preceq$. We denote this object by $K_{\preceq}$.*

A simplicial map between ordered simplicial complexes is said to be order-preserving, if it so between the vertex sets. We denote the resulting category by **oSimp**.

Obviously, there is a forgetful functor **oSimp** $\to$ **Simp**, $K_{\preceq} \mapsto K$.

Definition 6.15 *Let $K_{\preceq}$ and $K'_{\preceq'}$ be ordered simplicial complexes.*

a) The product order $\preceq^x$ on $\mathrm{Vert}(K \times K') = \mathrm{Vert}(K) \times \mathrm{Vert}(K')$ is given by $(v, v') \preceq^x (w, w')$ if and only if $v \preceq v'$ and $w \preceq w'$.

b) We define the product $K_{\preceq} \oslash K'_{\preceq}$ as follows: $\tau \in K_{\preceq} \oslash K'_{\preceq}$ if and only if there exist $\sigma \in K$ and $\sigma' \in K'$ such that τ is a totally ordered subset of $\sigma \times \sigma'$ with respect to $\preceq^x$.

Proposition 6.16 *Let $K_{\preceq}$ and $K'_{\preceq'}$ be ordered simplicial complexes. It holds that*

a) $| K_{\preceq} \oslash K'_{\preceq} |$ is a deformation retract of $| K \times K' |$.

b) $| K_{\preceq} \oslash K'_{\preceq} |$ is homeomorphic to $| K | \times | K' |$.

c) $| K \times K' |$ and $| K | \times | K' |$ have the same homotopy type.

Proof: For the proof, we refer to [38, Chapter 8]. Part *a*) is [38, Lemma 8.11.], part *b*) is [38, lemma 8.9.]. Part *c*) follows immediately from *a*) and *b*). ∎

With this in mind, we can understand the left-hand side of theorem 6.10. On the right-hand side, there are *sums* and *tensor products* of persistence **k**-modules. We recall these notions.

Definition 6.17 *Let $V_\bullet$ and $W_\bullet$ be persistence* **k***-modules.*

a) The sum $V_\bullet \oplus W_\bullet$ is the persistence **k***-module given by $(V_\bullet \oplus W_\bullet)_r := V_r \oplus W_r$ and $V_\bullet \oplus W_\bullet(r \leq t) := V_\bullet(r \leq t) \oplus W_\bullet(r \leq t)$ for all $r, t \in [0, \infty)$ such that $r \leq t$.*

b) The tensor product $V_\bullet \otimes W_\bullet$ is given by $(V_\bullet \otimes W_\bullet)_r := V_r \otimes_{\mathbf{k}} W_r$ and $V_\bullet \otimes W_\bullet(r \leq t) := V_\bullet(r \leq t) \otimes_{\mathbf{k}} W_\bullet(r \leq t)$ for all $[0, \infty)$ such that $r \leq t$.

Proposition 6.18 *Let $\mathbb{1}_I$ and $\mathbb{1}_J$ be interval modules. It holds that $\mathbb{1}_I \otimes \mathbb{1}_J \cong \mathbb{1}_{I \cap J}$.*

Proof: It follows from the definition of interval modules 2.9 and definition 6.17 using that $\mathbf{k} \otimes_{\mathbf{k}} 0 \cong 0$. ■

At the end of this section, we give a proof of theorem 6.10.

Proof of Theorem 6.10:
By the topological Künneth formula 6.1, there is a natural isomorphism for each $r \in [0, \infty)$

$$\mathrm{H}_n\big(\mid K_r \mid \times \mid K'_r \mid \big) \cong \bigoplus_{i+j=n} \big(\mathrm{H}_i(\mid K_r \mid) \otimes_{\mathbf{k}} \mathrm{H}_j(\mid K'_r \mid) \big)$$

By proposition 6.16 *c*), $\mid K_r \times K'_r \mid$ and $\mid K_r \mid \times \mid K'_r \mid$ have the same homotopy type, this implies that $\mathrm{H}_n\big(\mid K_r \times K'_r \mid \big) \cong \mathrm{H}_n(\mid K_r \mid \times \mid K'_r \mid)$.

It is a well known fact that the simplicial homology of a simplicial complex and the singular homology of its geometric realization are isomorphic. In conclusion,there is a natural isomorphism for each $r \in [0, \infty)$

$$\mathrm{H}_n(K_r \times K'_r) \cong \bigoplus_{i+j=n} \big(\mathrm{H}_i(K_r) \otimes_{\mathbf{k}} \mathrm{H}_j(K'_r) \big).$$

Using naturality, this yields an isomorphism as claimed. ■

6.3. Künneth approximation for decomposable Bregman point clouds

In this section, we present our result. Our Künneth-type formula is only an approximation. We show that there is a ln(2)-interleaving between the claimed persistence module and the persistence homology of the product space if *indexed logarithmically.*

In section 6.1, we studied, which is the best way to define a Bregman divergence on the product space. We can conclude that the best way to do this to take the sum as we did in chapter 3. Therefore we will only consider here such functions. We recall a definition from [39].

Definition 6.19 *We call a function $H : \Omega'' \to \mathbb{R}$ decomposable if there exist functions $F : \Omega \to \mathbb{R}$ and $G : \Omega' \to \mathbb{R}$ such that $\Omega'' = \Omega \times \Omega'$ and $H = F + G$.*

If $F : \Omega \to \mathbb{R}$ and $G : \Omega' \to \mathbb{R}$ are functions of Legendre type, then for all $(x, a),(y, b) \in \Omega \times \Omega'$ it holds that $D_{F+G}\big((x, a)\|(y, b)\big) = D_F(x\|y) + D_G(a\|b)$ by proposition 3.26.

Lemma 6.20 *Let (X, F), (Y, G) be two Bregman point clouds. For all $r \in [0, \infty)$ it holds that*

$$\check{\mathrm{C}}\mathrm{ech}_{F+G}(X \times Y, r) \subset \check{\mathrm{C}}\mathrm{ech}_F(X, r) \times \check{\mathrm{C}}\mathrm{ech}_G(Y, r) \subset \check{\mathrm{C}}\mathrm{ech}_{F+G}(X \times Y, 2r).$$

Proof: If $\tau \in \check{C}\text{ech}_{F+G}(X \times Y, r)$, then there is a $(w_1, w_2) \in \Omega \times \Omega'$ such that $D_{F+G}\big((x,a)\|(w_1,w_2)\big) \leq r$ for all $(x,a) \in \tau$. Using proposition 3.26 and the positivity of Bregman divergences (lemma 3.22), we see that

$$\begin{aligned} &D_F(x\|w_1) \leq r \text{ for all } x \in pr_X(\tau) \text{ and} \\ &D_G(a\|w_2) \leq r \text{ for all } a \in pr_Y(\tau). \end{aligned}$$

Indeed, $pr_X(\tau) \in \check{C}\text{ech}_F(X,r)$ and $pr_Y(\tau) \in \check{C}\text{ech}_G(Y,r)$. Thus by definition 6.11, $\tau \in \check{C}\text{ech}_F(X,r) \times \check{C}\text{ech}_G(Y,r)$.

Conversely, let $\tau \in \check{C}\text{ech}_F(X,r) \times \check{C}\text{ech}_G(Y,r)$ By definition 6.11, there are some $\sigma \in \check{C}\text{ech}_F(X,r)$ and $\sigma' \in \check{C}\text{ech}_G(Y,r)$ such that $\tau \subset \sigma \times \sigma'$. From this follows that, there are some $z_1 \in \Omega$ and $z_2 \in \Omega'$ such that

$$\begin{aligned} &D_F(x\|z_1) \leq r \text{ and} \\ &D_G(a\|z_2) \leq r \end{aligned}$$

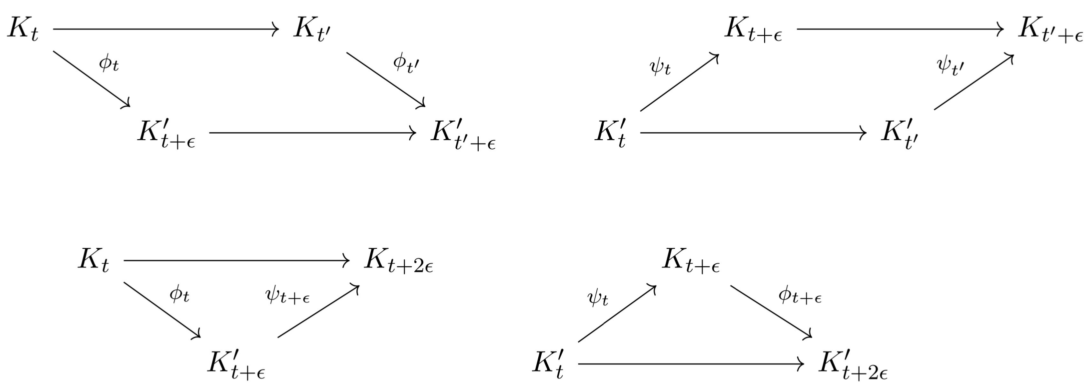
for all $(x,a) \in \tau$. Adding up these two inequalities, from proposition 3.26 it follows that $D_{F+G}\big((x,a)\|(z_1,z_2)\big) \leq 2r$ for all $(x,a) \in \tau$, which implies that $\tau \in \check{C}\text{ech}_{F+G}(X \times Y, 2r)$. ■

Now, we clarify what we mean by *logarithmically indexing.*

Definition 6.21 *The* <u>*logarithm functor*</u> $\mathbf{ln} : \underline{\mathbf{Simp}}^{\mathbb{R}_+} \to \underline{\mathbf{Simp}}^{\mathbb{R}}$ *is defined for objects* $K_\bullet \in \underline{\mathbf{Simp}}^{\mathbb{R}_+}$ *by* $\mathbf{ln}(K_\bullet)_t := K_{e^t}$ *and* $\mathbf{ln}(K_\bullet)(t \leq t') := K_\bullet(e^t \leq e^{t'})$, *for morphisms* $\eta : K_\bullet \to K'_\bullet$ *it is defined by* $\mathbf{ln}(\eta)_t := \eta_{e^t}$.

One can easily modify the definition of ϵ-interleaving 5.3 to objects in $\underline{\mathbf{Simp}}^{\mathbb{R}}$.

Definition 6.22 *Let* $K_\bullet$ *and* $K'_\bullet$ *be two objects in* $\underline{\mathbf{Simp}}^{\mathbb{R}}$ *and* $\epsilon \geq 0$. *An* <u>*ϵ-interleaving*</u> *between* $K_\bullet$ *and* $K'_\bullet$ *is given by two families of simplicial maps* $\{\phi_t : K_t \to K'_{t+\epsilon}\}_{t\in\mathbb{R}}$ *and* $\{\psi_t : K'_t \to K_{t+\epsilon}\}_{t\in\mathbb{R}}$ *such that the following diagrams commute for every* $t, t' \in \mathbb{R}$ *with* $t \leq t'$



Remark 6.23 *Since* H_n *is a functor, an* ϵ*-interleaving between objects* $K_\bullet$ *and* $K'_\bullet$ *in* $\underline{\mathbf{Simp}}^{\mathbb{R}}$ *induces an* ϵ*-interleaving between the persistence* $\mathbf{k}$*-modules* $\mathrm{H}_n(K_\bullet)$ *and* $\mathrm{H}_n(K'_\bullet)$ *for every* $n \in \mathbb{N}_0$.

Lemma 6.24 *Let* $K_\bullet$ *and* $K'_\bullet$ *be two objects in* $\underline{\mathbf{Simp}}^{\mathbb{R}_+}$ *such that the maps* $K_r \to K_s$ *and* $K'_r \to K'_s$ *are inclusions for every* $r, s \in \mathbb{R}_+$ *with* $r \leq s$*. Furthermore, we assume that* $K_r \subset K'_r \subset K_{2r}$ *for all* $r \in \mathbb{R}_+$*. In this case, there is an* $\ln(2)$*-interleaving between* $\mathbf{ln}(K_\bullet)$ *and* $\mathbf{ln}(K'_\bullet)$.

Proof: By our assumption, it holds for all $r \in \mathbb{R}_+$ that $K_r \subset K'_r \subset K_{2r}$. Using $2 \cdot e^t = e^{t+ln(2)}$, we can conclude that

$$\mathbf{ln}(K_\bullet)_t \subset \mathbf{ln}(K'_\bullet)_t \subset \mathbf{ln}(K_\bullet)_{t+\ln(2)} \tag{2}$$

for all $t \in \mathbb{R}$.

For each $t \in \mathbb{R}$ let $\phi_t : \mathbf{ln}(K_\bullet)_t \to \mathbf{ln}(K'_\bullet)_{t+\ln(2)}$ be the first inclusion in (2) concatenated with the map $\mathbf{ln}(K'_\bullet)_t \hookrightarrow \mathbf{ln}(K'_\bullet)_{t+\ln(2)}$. The second inclusion in (2) gives us a simplicial map $\psi_t : \mathbf{ln}(K'_\bullet)_t \to \mathbf{ln}(K_\bullet)_{t+\ln(2)}$. Since all the morphism are inclusions, one can check that the diagrams like in definition 6.22 commute. ■

Example 6.25 *Let* (M, d_M) *be a metric space. It is a well known result that* $\check{\mathrm{C}}\mathrm{ech}(M, r) \subset \mathrm{Rips}(M, r) \subset \check{\mathrm{C}}\mathrm{ech}(M, 2r)$ *(see e.g. [1]). By the above-introduced terminology, one can say that the interleaving distance between* $\mathbf{ln}(\check{\mathrm{C}}\mathrm{ech}(M, \bullet))$ *and* $\mathbf{ln}(\mathrm{Rips}(M, \bullet))$ *is smaller or equal than* $\ln(2)$ *or with other words, the interleaving distance between the Čech and the Vietoris-Rips complex is less or equal than* $\ln(2)$ *if indexed logarithmically.*

It turns out that lemma 6.24 is also useful for Bregman point clouds.

Corollary 6.26 *Let* (X, F) *and* (Y, G) *be Bregman point clouds. We have*

$$\mathrm{d_I}\left(\mathrm{H}_n\left(\mathbf{ln}\big(\check{\mathrm{C}}\mathrm{ech}_{F+G}(X \times Y, \bullet)\big)\right), \mathrm{H}_n\left(\mathbf{ln}\big(\check{\mathrm{C}}\mathrm{ech}_F(X, \bullet) \times \check{\mathrm{C}}\mathrm{ech}_G(Y, \bullet)\big)\right)\right) \leq \ln(2)$$

for every $n \in \mathbb{N}_0$.

Proof: From lemma 6.20 and lemma 6.24 it follows that there is an ln(2)-interleaving between $\mathbf{ln}\big(\check{\mathrm{C}}\mathrm{ech}_{F+G}(X \times Y, \bullet)\big)$ and $\mathbf{ln}\big(\check{\mathrm{C}}\mathrm{ech}_F(X, \bullet) \times \check{\mathrm{C}}\mathrm{ech}_G(Y, \bullet)\big)$. The claimed inequality follows from remark 6.23. ■

Lemma 6.27 *For every* $n \in \mathbb{N}_0$ *there is an isomorphism of persistence* $\mathbf{k}$*-modules*

$$\mathrm{H}_n(\check{\mathrm{C}}\mathrm{ech}_F(X, \bullet) \times \check{\mathrm{C}}\mathrm{ech}(Y, \bullet)) \cong \bigoplus_{i+j=n} \mathrm{H}_i\left(\check{\mathrm{C}}\mathrm{ech}_F(X, \bullet)\right) \otimes \mathrm{H}_j\left(\check{\mathrm{C}}\mathrm{ech}_G(Y, \bullet)\right)$$

Proof: It is a special case of theorem 6.10. ■

Corollary 6.28 *Let (X, F) and (Y, G) be Bregman point clouds, it holds for every $n \in \mathbb{N}_0$ that*

$$\operatorname{dgm}\Big(\mathrm{H}_n\big(\check{\mathrm{C}}\mathrm{ech}_F(X,\bullet) \times \check{\mathrm{C}}\mathrm{ech}(Y,\bullet)\big)\Big) = \\ = \bigcup_{i+j=n} \Big\{(\max\{a,c\}, \min\{b,d\}) \Big| (a,b) \in \operatorname{dgm}\Big(\mathrm{H}_i\big(\check{\mathrm{C}}\mathrm{ech}_F(X,\bullet)\big)\Big), \\ (c,d) \in \operatorname{dgm}\Big(\mathrm{H}_j\big(\check{\mathrm{C}}\mathrm{ech}_G(Y,\bullet)\big)\Big)\Big\}.$$

Proof: It follows from lemma 6.27 and proposition 6.18. ■

Using this corollary, one can easily compute $\operatorname{dgm}\Big(\mathrm{H}_n\big(\check{\mathrm{C}}\mathrm{ech}_F(X,\bullet)\times\check{\mathrm{C}}\mathrm{ech}(Y,\bullet)\big)\Big)$ in terms of $\operatorname{dgm}\Big(\mathrm{H}_*\big(\check{\mathrm{C}}\mathrm{ech}_F(X,\bullet)\big)\Big)$ and $\operatorname{dgm}\Big(\mathrm{H}_*\big(\check{\mathrm{C}}\mathrm{ech}_G(Y,\bullet)\big)\Big)$.
From corollary 6.26 and the isometry theorem 5.5, it follows that the bottleneck distance between $\operatorname{dgm}\Big(\mathrm{H}_n\big(\check{\mathrm{C}}\mathrm{ech}_F(X,\bullet) \times \check{\mathrm{C}}\mathrm{ech}(Y,\bullet)\big)\Big)$ and $\operatorname{dgm}\Big(\mathrm{H}_n\big(\check{\mathrm{C}}\mathrm{ech}_{F+G}(X\times Y,\bullet)\big)\Big)$ is less or equal than $\ln(2)$ if indexed logarithmically. Indeed, if we allow a computational error less or equal than $\ln(2)$, we can use corollary 6.28 for computations. Thus, we get an approximative Künneth formula.

Appendix A Matlab code

A.1 CircumBall

```
function [q,r,Front] = CircumBall(P,F)
%Compute the smallest dual circumball of a Bregman
%point cloud
%INPUTS:
%   P --> a matrix, each row represents a point
%   F --> a function of Legendre type
%OUTPUTS:
%   q --> the center of the smallest dual circumball of P
%   r --> the radius of the smallest dual circumball of P
%   Front --> row indices of the vertices contained in
%             the front face of P, this output is optional

   global optTol
   options = optimoptions(@fminunc,...
             'OptimalityTolerance', ...
             optTol,'StepTolerance',optTol);

    a0 = P(1,:);
    A = P(2:end,:);
    [k,~]= size(A);
    lambdastart = zeros(1,k);

    %Solve the unconstrained convex optimization problem
    [lambda,fval] = fminunc(@(lambda)( ...
       (F((1-sum(lambda))*a0+sum(lambda'.*A,1)) ...
       - (1-sum(lambda))*F(a0)-dot(lambda,Frow(A))) ),...
       lambdastart,options);
    lambda0 = 1 - sum(lambda);
    q = lambda0*a0+sum(lambda'.*A,1);
    r = -fval;

    %Compute the front face, if it is requested
    if nargout > 2
        Front = [];
        if lambda0 > 0
            Front = 1;
        end
        for j = 1:k
            if lambda(j) > 0
                Front = [Front, j+1 ];
```

```
            end
        end
    end


    function B = Frow(A)
        %Apply the function F for each row of a matrix
        [k,~]= size(A);
        B = zeros(1,k);
        for j = 1:k
            B(j) = F(A(j,:));
        end
    end
end
```

A.2 CechRadiusFunction

```
function complex = CechRadiusFunction(X,F,DF)
%Build the filtered Cech-complex of a Bregman point cloud
%INPUTS:
%   X --> a matrix, each row represents a point
%   F --> a function of Legendre type
%   DF --> to F corresponding Bregman divergence
%OUTPUTS:
%   complex --> a structure array,
%               It represents the full simplicial complex
%               of X. Each simplex is represented as the
%               row indices of its vertices. The fields
%               'center' and 'radius' show the center and
%               the radius of the smallest including dual
%               ball of each simplex.

    %To make the computation easier, we consider only the
    %d-skeleton of the full simplicial complex
    global d;
    %tolerance for the radii
    global rTol;

    %cardinality of X
    n = size(X,1);
    %the case d > n-1 is irrelevant, since the full
    %simplicial complex of X has dimension n-1
    d = min(d,n-1);
```

```
%Build the simplicial complex
combi = 1:n;
numOfAllComb = sum(arrayfun(@(k) nchoosek(n,k),...
                    1:(d+1)));

%Set up an empty structure array
complex(numOfAllComb) = struct('vertices',[],...
                              'dimension',NaN,...
                              'center',NaN,...
                              'radius',NaN);
m = 1;  %counter variable

for k = 1:(d+1)
    %sub contains all the simplicies of dimension k-1
    sub = nchoosek(combi,k);

    for j=1:size(sub,1)
        %0-simplicies are easy to handle
        if (k-1) == 0
            complex(m).vertices = sub(j,:);
            complex(m).center =  X(sub(j,:),:);
            complex(m).dimension = k-1;
            complex(m).radius = 0;
            m = m+1;
            continue;
        end

        %Check if the simplex sub(j,:) is unmarked
        unmarked = true;
        for s = 1:(m-1)
                if isequal(complex(s).vertices ,...
                                            sub(j,:) )
                    unmarked = false;
                    break;
                end
        end

        %If sub(j,:) is unmarked, then it is a lower
        %bound of on interval
        if unmarked
            %Compute the smallest dual circumball
            [z,rho] =  CircumBall(X(sub(j,:),:),...
                                  @(x) F(x));
```

```
                complex(m).vertices = sub(j,:);
                complex(m).center = z;
                complex(m).radius = rho;
                complex(m).dimension = ...
                    size(complex(m).vertices,2) -1;
                m = m+1;

                %Compute the upper bound of the interval
                %i.e. collect the points in X, which are
                %in the smallest dual circumball
                %of sub(j,:)
                Upp = sub(j,:);
                XwithoutSubj = combi;
                XwithoutSubj(Upp) = [];
                for j2 = XwithoutSubj
                    if DF( X(j2,:)   , z ) <= rho
                        Upp = [Upp,j2];
                    end
                end

                %MarkTheInterval
                Low = sub(j,:);
                UppwithoutLow  = setdiff(Upp, Low );
                for k2 = 1:size(UppwithoutLow,2)
                    sub2 = nchoosek(UppwithoutLow,k2);
                    for j3 = 1:size(sub2,1)
                        complex(m).vertices =...
                          sort([sub2(j3,:),Low]);
                        complex(m).center = z;
                        complex(m).radius = rho;
                        complex(m).dimension = ...
                          size(complex(m).vertices,2) -1;
                        m = m+1;
                    end
                end
            end
        end
    end

    %To compute persistent homology, we use the function
    %'api.Plex4.createExplicitSimplexStream'  from the
    %Javaplex package. This function can not handle with
    %small computational errors. To avoid difficulties
    %later on, we make almost equal radius equal.
```

```
    %Sort the table by radius
    [~,idx]=sort([complex.radius]);
    complex = complex(idx);

    for j = 2:numOfAllComb
        if abs(complex(j).radius-complex(j-1).radius)<...
                                                        rTol
            complex(j).radius = complex(j-1).radius;
        end
    end
end
```

A.3 CechPersistenceDiagram

```
function [intervals] = CechPersistenceDiagram(X,F,DF)
%Compute the persistence diagram of a Bregman point cloud
%using Cech complex construction
%INPUTS:
%   X --> a matrix, each row represents a point
%   F --> a function of Legendre type
%   DF --> to F corresponding Bregman divergence
%OUTPUTS:
%   intervals --> object produced by the JavaPlex package

    %maximum dimension of a simplex
    global d;
    %characteristic of the coefficient field
    global p;

    complex = ...
        CechRadiusFunction(X,@(x) F(x),@(x,y) DF(x,y));

    import edu.stanford.math.plex4.*;
    stream = api.Plex4.createExplicitSimplexStream(max(...
              [complex.radius])+1);
    [~,m] = size(complex);

    for Q = 1:m
        if complex(Q).dimension == 0
            stream.addVertex(complex(Q).vertices,0);
        else
            stream.addElement(complex(Q).vertices,...
```

```
                                    complex(Q).radius);
        end
    end
    stream.finalizeStream();

    persistence =...
        api.Plex4.getModularSimplicialAlgorithm(d,p);
    intervals = persistence.computeIntervals(stream);
end
```

A.4 WeightedDelTriangulation

```
function complex = WeightedDelTriangulation(X,w)
%Compute the weighted Delaunay triangulation
%INPUTS:
%    X --> matrix, each row represents a point
%    w --> vector containing the weights
%OUTPUTS:
%    complex --> structure array, each simplex is
%                represented as the row indices of
%                its vertices.

    %Lift the points to R^(dim+1)
    [n, dim] = size(X);
    lifted = zeros(n,dim+1);
    for i=1:n
        x = X(i,:);
        lifted(i,:) = [x, x*x' - w(i)];
    end

    %Compute the indices of facets of the lower hull
    C = convhulln(lifted); %error, if n < dim +2
    [m, k] = size(C);
    center = mean(lifted,1);
    for j=1:m
        v = null(bsxfun(@minus, lifted(C(j,1),:),...
                        lifted(C(j,2:end),:)))';
        [mm, kk] = size(v);
        if mm > 1
            V(j,:) = NaN;
            error('nullspace error')
           % possibility of degenerate null vectors
        else
            V(j,:) = v;
```

```
    end
    mid(j,:) = mean(lifted(C(j,:),:),1);
end
dot = sum(bsxfun(@minus, center, mid).*V, 2);
outer = dot < 0;
V(outer,:) = -1*V(outer,:);
ind = V(:,k) > 0;

Del = C(ind, :);

%Add each face to the Delaunay triangulation
m = 1; %counter variable
complex(1) = struct('vertices',[],'dimension',NaN,...
                    'center',NaN,'radius',NaN);

for j = 1:size(Del,1)
%Add the faces of 'Del(j,:)' to 'complex'
    for k = 1:size(Del(j,:),2)
        %'sub' contains all the subsets of
        %'Del(j,:)' of cardinality k
        sub = nchoosek(Del(j,:),k);
        for jj=1:size(sub,1)
           %Check, if the face is already in 'complex'
            notcollected = true;
            for s = 1:(m-1)
                sort(sub(jj,:));
                if isequal(complex(s).vertices,...
                           sort(sub(jj,:)) )
                    notcollected = false;
                    break;
                end
            end

            if notcollected
                complex(m).vertices = sort(sub(jj,:));
                complex(m).dimension = ...
                                 size(sub(jj,:),2)-1;
                complex(m).center = NaN;
                complex(m).radius = NaN;
                m = m+1;
            end
        end
    end
end
```

```
    %Sort 'complex' by dimension in decreasing order
    [~,idx]=sort([complex.dimension],'descend');
    complex = complex(idx);
end
```

A.5 DelaunayRadiusFunction

```
function complex = DelaunayRadiusFunction(X,F)
%Build a filtered Delaunay complex from a Bregman
%point cloud
%INPUTS:
%    X --> matrix, each row represents a point
%    F --> a function of Legendre type
%OUTPUTS:
%    complex --> a structure array. Each simplex is
%                represented as the row indices of
%                its vertices. The fields 'center' and
%                'radius' show the center and the radius
%                of the smallest empty dual circumball
%                of each simplex.

    %tolerance for the radii
    global rTol;

    weights = (1:size(X,1))';
    for j = 1:size(X,1)
        weights(j) =  norm(X(j,:))^2 - 2 * F(X(j,:));
    end

    complex = WeightedDelTriangulation(X,weights);

    %Compute the values of the Delaunay radius function
    for j = 1:size(complex,2)
        if complex(j).dimension == 0
            complex(j).radius = 0;
            complex(j).center = X(complex(j).vertices,:);
        end

        if isnan(complex(j).center)
            [z,rho,Front] = CircumBall(...
                X(complex(j).vertices,:),@(x) F(x));
            complex(j).center = z;
            complex(j).radius = rho;
```

```
            %MarkTheInterval
            Low = complex(j).vertices(Front);
            Upp = complex(j).vertices;
            UppwithoutLow  = setdiff(Upp, Low );

            %Mark the lower bound
            %before 'complex(j) are only simplicies with
            %higher or equal dimension
            for j2 = j:size(complex,2)
                if isequal(complex(j2).vertices , Low)
                    complex(j2).center = z;
                    complex(j2).radius = rho;
                    break;
                end
            end

            %Mark the simplicies between the lower and
            %the upper bound
            for k2 = 1:(size(UppwithoutLow,2)-1)
                    sub2 = nchoosek(UppwithoutLow,k2);
                    for j3 = 1:size(sub2,1)
                        new = sort([sub2(j3,:),Low]);

                        %Find  'new' in 'complex.vertices'
                        %before 'complex(j) are only
                        %simplicies with higher or
                        %equal dimension.
                        for j4= j:size(complex,2)
                            if isequal(...
                                 complex(j4).vertices,new)
                                complex(j4).center = z;
                                complex(j4).radius = rho;
                                break;
                            end
                        end
                    end
            end
        end
end

%To compute persistent homology we use the function
%'api.Plex4.createExplicitSimplexStream' from the
% Javaplex package. This function can not handle with
```

```
    %small computational errors. To avoid difficulties
    %later on, we make almost equal radius equal.

    %Sort the table by radius
    [~,idx]=sort([complex.radius]);
    complex = complex(idx);

    for j = 2:size(complex,2)
        if abs(complex(j).radius-complex(j-1).radius)<...
                                                      rTol
            complex(j).radius = complex(j-1).radius;
        end
    end
end
```

A.6 DelaunayPersistenceDiagram

```
function intervals = DelaunayPersistenceDiagram(X,F)
%Compute the persistence diagram of a Bregman point cloud
%using Delaunay complex construction
%INPUTS:
%   X --> matrix, each row represents a point
%   F --> function of Legendre type
%OUTPUTS:
%   intervals --> object produced by the JavaPlex package

    %the homology will be computed in all dimensions
    %strictly less than d
    global d
    %characteristic of the coefficient field
    global p;
    complex = DelaunayRadiusFunction(X,@(x) F(x));

    import edu.stanford.math.plex4.*;
    stream = api.Plex4.createExplicitSimplexStream(...
                                 max([complex.radius])+1);
    [~,m] = size(complex);

    for Q = 1:m
        if complex(Q).dimension == 0
            stream.addVertex(complex(Q).vertices,0);
        else
            stream.addElement(complex(Q).vertices,...
```

```
                                complex(Q).radius);
        end
    end
    stream.finalizeStream();

    persistence = ...
        api.Plex4.getModularSimplicialAlgorithm(d,p);
    intervals = persistence.computeIntervals(stream);
end
```

A.7 ExperimentPerformance

```
function ExperimentPerformance(NumOfPoints,...
    DimOfContSpace,Divergence,F,DF)
%Compare the Cech and Delaunay radius function algorithms
%INPUTS:
%   NumOfPoints --> a vector containing the cardinality of
%                   the Bregman point clouds
%   DimOfContSpace --> an integer, the dimension of
%                      the containing space
%   Divergence --> a string array (for plotting)
%   F --> a function of Legendre type
%   DF --> the corresponding Bregman divergence
%OUTPUTS:
%   A plot that shows the average running time for
%   several cardinality

    global optTol rTol d p;

    assert(DimOfContSpace + 1 < min(NumOfPoints),...
        'To get the DelaunayPersistenceDiagram function
           work we need DimOfContSpace + 1 < min(
           NumOfpoints)')

    %Measure the running time
    %of 'CechPersistenceDiagram' and
    %'DelaunayPersistenceDiagram'. The functions
    %are called 'k' times for the same data, then
    %the mean of the running times will be taken.
    k = 10;
    tCech = zeros(k,1)';
    tDel = zeros(k,1)';
    AvgtCech = zeros(size(NumOfPoints,2),1)';
    AvgtDel = zeros(size(NumOfPoints,2),1)';
```

```
    for s = 1:size(NumOfPoints,2)
        n = NumOfPoints(s);
        X = 1 + (100-1).*rand(n,DimOfContSpace);

        for j  = 1:k
            tic;
            CechPersistenceDiagram(X,...
                           @(x) F(x),@(x,y) DF(x,y));
            tCech(j) = toc;

            tic;
            DelaunayPersistenceDiagram(X,@(x) F(x));
            tDel(j) = toc;
        end
        AvgtCech(s) = mean(tCech);
        AvgtDel(s) = mean(tDel);
    end

    %Plot the average running times
    plot(NumOfPoints,AvgtCech,'r-o');
    hold on;
    plot(NumOfPoints,AvgtDel,'g-o');
    xlabel(['number of points in a ',...
        num2str(DimOfContSpace),' dimensional space']);
    ylabel('running time in seconds');
    title(Divergence);
    legend({'Cech complex','Delaunay complex'},...
            'Location','northwest');
    hold off;
    saveas(gcf,['ExpPerf-',...
           'Dim',num2str(DimOfContSpace),...
           '-',Divergence],'jpg');
end
```

A.8 BottleneckDistance

```
function  D = BottleneckDistance(intervalsA,intervalsB)
%Compute the bottleneck distance between
%persistence diagrams
%INPUTS:
%   intervalsA,intervalsB --> objects created by the
%                 function 'persistence.computeIntervals()
```

```
%                'from JavaPlex package
%OUTPUTS
%    D --> array containing the bottleneck distances in
%          each dimension

    import edu.stanford.math.plex4.*;

    MaxDeg = max(numel(intervalsA.getBettiSequence),...
                    numel(intervalsB.getBettiSequence));
    for k = 1:MaxDeg
        intervalsA_dimk =...
                 intervalsA.getIntervalsAtDimension(k-1);
        intervalsB_dimk =...
                 intervalsB.getIntervalsAtDimension(k-1);
        D(k) = edu.stanford.math.plex4.bottleneck. ...
          BottleneckDistance.computeBottleneckDistance(...
          intervalsA_dimk,intervalsB_dimk);
    end
end
```

A.9 ExperimentStability

```
function ExperimentStability(NumOfPoints,...
            DimOfContSpace,ItSteps,Divergence,F,DF)
%Test the stability of persistence diagrams
% using synthetic datasets
%INPUT:
%    NumOfPoints --> cardinality of the Bregman point cloud
%    DimOfContSpace --> dimension of the space where the
%                       Bregman point cloud come from
%    ItSteps --> number it iteration steps
%    Divergence --> a string array (for plotting)
%    F --> a function of Legendre type
%    DF --> to F corresponding Bregman divergence

    global optTol rTol d p;

    %Generate a synthetic dataset
    X = 1 + (70-1).*rand(NumOfPoints,DimOfContSpace);

    %arrays for the bottleneck distances
    bd0 = zeros(1,ItSteps);
    bd1 = zeros(1,ItSteps);
    bd2 = zeros(1,ItSteps);
```

```
%array for Hausdorff distances
hd = [];

for k = 1:ItSteps
    %Create a Bregman point cloud that converges to X
    Y = X + 0.01*(1/k)*rand(NumOfPoints,...
                            DimOfContSpace);
    hd(k) = HausdorffDist(X,Y);

    intervalsX = CechPersistenceDiagram(X,...
                    @(x) F(x),@(x,y) DF(x,y));
    intervalsY = CechPersistenceDiagram(Y,...
                    @(x) F(x),@(x,y) DF(x,y));

    botDist = BottleneckDistance(intervalsX,...
                                 intervalsY);
    bd0(k) = botDist(1);

    if size(botDist,2) > 2
        bd1(k) = botDist(2);
        bd2(k) = botDist(3);
    elseif size(botDist,2) > 1
        bd1(k) = botDist(2);
        bd2(k) = 0;
    else
        bd1(k) = 0;
        bd2(k) = 0;
    end

end

%Plot the results
yyaxis left
hdplot = plot(hd,"b--s");
set(gca,'Ycolor','b');
xlabel([num2str(NumOfPoints),' points in a ',...
   num2str(DimOfContSpace), ' dimensional space']);
ylabel('Hausdorff distance')

yyaxis right
bd0plot = plot(bd0,"r-o");
set(gca,'Ycolor','r');
```

```
    if any(bd1)
        hold on;
        bd1plot = plot(bd1);
        bd1plot.Color = [0.9290 0.6940 0.1250];
        bd1plot.Marker = "o";
        bd1plot.LineStyle = "-";
    end
    if any(bd2)
        hold on;
        bd2plot = plot(bd2,"m-o");
    end
    hold off;
    ylabel('bottleneck distance')

    if any(bd2)
        legend([bd0plot,bd1plot,bd2plot],...
                        {'dim 0','dim 1','dim 2'});
    elseif any(bd1)
        legend([bd0plot,bd1plot],{'dim 0','dim 1'});
    else
        legend([bd0plot],{'dim 0'});
    end

    title(Divergence);
    saveas(gcf,['ExpStab-',num2str(NumOfPoints),...
        'points-',num2str(DimOfContSpace),...
        'dimSpace-',Divergence],'jpg')
end
```

References

[1] Herbert Edelsbrunner and Hubert Wagner. Topological data analysis with Bregman divergences. *ArXiv*, abs/1607.06274, July 2016.

[2] Vanessa Robins. Towards computing homology from approximations. *Topology Proceedings*, 24, January 1999.

[3] Herbert Edelsbrunner, David Letscher, and Afra Zomorodian. Topological persistence and simplification. *Discrete and Computational Geometry*, 28:511–533, January 2002.

[4] Gunnar Carlsson and Afra Zomorodian. Computing persistent homology. *Discrete and Computational Geometry*, 33:249–274, February 2005.

[5] Joseph Chan, Gunnar Carlsson, and Raul Rabadan. Topology of viral evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 110, October 2013.

[6] Vin de Silva and Robert Ghrist. Coverage in sensor networks via persistent homology. *Algebraic & Geometric Topology*, 7, April 2007.

[7] Jean-Baptiste Bardin, Gard Spreemann, and Kathryn Hess. Topological exploration of artificial neuronal network dynamics. *Network Neuroscience*, 3(3):725–743, January 2019.

[8] Anna Huang. Similarity measures for text document clustering. *Proceedings of the 6th New Zealand Computer Science Research Student Conference*, January 2008.

[9] Minh Do and Martin Vetterli. Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, 11:146–58, February 2002.

[10] Jau-Yuen Chen, Charles A. Bouman, and Jan P. Allebach. Multiscale branch-and-bound image database search. In *Storage and Retrieval for Image and Video Databases V*, volume 3022, pages 133–144. International Society for Optics and Photonics, 1997.

[11] Cédric Févotte, Nancy Bertin, and Jean-Louis Durrieu. Nonnegative matrix factorization with the Itakura-Saito divergence: With application to music analysis. *Neural computation*, 21:793–830, October 2008.

[12] Herbert Edelsbrunner and John Harer. *Computational Topology: An Introduction.* Applied Mathematics. American Mathematical Society, 2010.

[13] Steve Y. Oudot. *Persistence Theory: From Quiver Representations to Data Analysis.* American Mathematical Society, 2015.

[14] Ulrich Bauer. Ripser: efficient computation of Vietoris-Rips persistence barcodes. *ArXiv*, abs/1908.02518, August 2019.

[15] Andrew Tausz, Mikael Vejdemo-Johansson, and Henry Adams. JavaPlex: A research software package for persistent (co)homology. In Hong H. and Yap C., editors, *Proceedings of ICMS 2014*, Lecture Notes in Computer Science 8592, pages 129–136, 2014.

[16] Allen Hatcher. *Algebraic Topology.* Algebraic Topology. Cambridge University Press, 2002.

[17] Wojciech Chachólski, Barbara Giunti, and Claudia Landi. Invariants for tame parametrised chain complexes. *ArXiv*, abs/2003.03969, March 2020.

[18] Lev M. Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR Computational Mathematics and Mathematical Physics*, 7(3):200 – 217, 1967.

[19] Jean-Daniel Boissonnat, Frank Nielsen, and Richard Nock. Bregman Voronoi diagrams. *Discrete and Computational Geometry, Springer Verlag*, hal-00488441:pp.200, June 2010.

[20] Heinz H. Bauschke and Jonathan M. Borwein. Legendre functions and the method of random Bregman projections. *Journal of Convex Analysis*, 4(1):27–67, 1997.

[21] James R. Munkres. *Topology.* Topology. Prentice-Hall, 2000.

[22] Claude E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, July 1948.

[23] R. Tyrrell Rockafellar. *Convex analysis.* Princeton University Press, 1972.

[24] Herbert Edelsbrunner, Ziga Virk, and Hubert Wagner. Smallest Enclosing Spheres and Chernoff Points in Bregman Geometry. In Bettina Speckmann and Csaba D. Tóth, editors, *34th International Symposium on Computational Geometry (SoCG 2018)*, volume 99 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 35:1–35:13. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2018.

[25] Richard Nock and Frank Nielsen. Fitting the smallest enclosing bregman ball. In Gama J., Camacho R., Brazdil P.B., Jorge A.M., and Torgo L., editors, *Machine Learning: ECML 2005*, volume 3720 of *Lecture Notes in Computer Science*, pages 649–656, October 2005.

[26] Frank Nielsen and Richard Nock. On the smallest enclosing information disk. *Information Processing Letters*, 105:93–97, January 2006.

[27] Jacob E. Goodman and Joseph O'Rourke. *Handbook of discrete and computational geometry.* Discrete mathematics and its applications. Chapman & Hall/CRC, 2004.

[28] Frederick McCollum. Power diagrams. MATLAB Central File Exchange, July 9, 2020. `https://www.mathworks.com/matlabcentral/fileexchange/44385-power-diagrams`.

[29] Ulrich Bauer and Herbert Edelsbrunner. The Morse theory of Čech and Delaunay complexes. *Transactions of the American Mathematical Society*, 369(5):3741–3762, December 2016.

[30] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization.* Cambridge University Press, USA, 2004.

[31] Frédéric Chazal, Vin de Silva, and Steve Oudot. Persistence stability for geometric complexes. *Geometriae Dedicata*, 173(1):193–214, December 2013.

[32] Frédéric Chazal, Vin de Silva, Marc Glisse, and Steve Oudot. The structure and stability of persistence modules. *SpringerBriefs in Mathematics*, 2016.

[33] Zachary Danziger. Hausdorff distance. MATLAB Central File Exchange, July 8, 2020. `https://www.mathworks.com/matlabcentral/fileexchange/26738-hausdorff-distance`.

[34] James R. Munkres. *Elements of Algebraic Topology.* Addison Wesley Publishing Company, 1984.

[35] Hitesh Gakhar and Jose A. Perea. Künneth formulae in persistent homology. *ArXiv*, abs/1910.05656, October 2019.

[36] Sunhyuk Lim, Facundo Memoli, and Osman Berat Okutan. Vietoris-Rips persistent homology, injective metric spaces, and the filling radius. *ArXiv*, abs/2001.07588, January 2020.

[37] Gunnar Carlsson and Benjamin Filippenko. Persistent homology of the sum metric. *Journal of Pure and Applied Algebra*, 224(5):106244, May 2020.

[38] Samuel Eilenberg and Norman E. Steenrod. *Foundations of Algebraic Topology.* Princeton University Press, 1952.

[39] Herbert Edelsbrunner, Ziga Virk, and Hubert Wagner. Topological data analysis in information space. *ArXiv*, abs/1903.08510, April 2019.