A PDE approach to centroidal tessellations of domains

Fabio Camilli and Adriano Festa

June 23, 2021

Abstract

We introduce a class of systems of Hamilton-Jacobi equations that characterize critical points of functionals associated to centroidal tessellations of domains, i.e., tessellations where generators and centroids coincide, such as centroidal Voronoi tessellations and centroidal power diagrams. An appropriate version of the Lloyd algorithm, combined with a Fast Marching method on unstructured grids for the Hamilton-Jacobi equation, allows computing the solution of the system. We propose various numerical examples to illustrate the features of the technique.

AMS-Subject Classification: 65K10, 49M05, 65D99, 35F21, 49N70.

Keywords: Voronoi tessellation; K-means; power diagram; Mean Field Games; Hamilton-Jacobi equation; Fast Marching method.

1 Introduction

A partition, or tessellation, of a set Ω is a collection of mutually disjoint subsets Ωk ⊂ Ω, k = 1, . . . , K, such that ∪Kk=1Ωk = Ω. A classical model is the Voronoi tessellation and, in this case, the sets Ωk are called Voronoi diagrams. Tessellations and other similar families of geometric objects arise in several applications, ranging from graphic design, astronomy, clustering, geometric modelling, data analysis, resource optimization, quadrature formulas, and discrete integration, sensor networks, numerical methods for partial differential equations (see [2, 21]).

Partitions and tessellations are frequently associated with objective functionals, defining desired additional properties to be satisfied. A well-known example is the K-means problem in cluster analysis, which aims to subdivide a data set into K clusters such that each data point belongs to the cluster
with the nearest cluster center. Minima of the $K$-means functional result in a partitioning of the data in centroidal Voronoi diagrams, i.e., Voronoi diagrams for which generators and centroids coincide (see [12]). In other applications, the size of the cells is prescribed (capacity-constrained problem), and the partition of $\Omega$ is given by another generalization of Voronoi diagrams, called power diagrams ([1, 6]).

This work aims to introduce a PDE method for hard clustering and other problems related to partitions of a set. The starting point of our approach is a characterization of the optimal parameters of a mixture model in soft clustering analysis employing a multi-population Mean Field Games system, a PDEs system arising in differential games with a large number of players ([3]). It is well known that the hard clustering $K$-means problem can be seen as the limit of Gaussian mixture model for soft clustering analysis when the variance parameter of the mixture model goes to 0 (see [5]). Interpreting the previous limit as a vanishing viscosity limit in the multi-population Mean Field Games problem introduced in [3], we deduce the following system of first-order Hamilton-Jacobi (HJ in short) equations of Eikonal type

$$
\begin{aligned}
\left\{ 
\begin{array}{ll}
|Du_k| = 1, & x \in \mathbb{R}^d, \ k = 1, \ldots, K \\
u_k(\mu_k) = 0, & \\
S^k = \{ x \in \mathbb{R}^d : u_k(x) = \min_{j=1}^{\ldots, K} u_j(x) \}, & \\
\mu_k = \frac{\int_{S^k} x \rho(x) dx}{\int_{S^k} \rho(x) dx}, & 
\end{array}
\right.
\end{aligned}
$$

(1.1)

where $\rho$ is a density function supported in a bounded set $\Omega$, representing the distribution of the data set. The equations are coupled through the values $\mu_k$, which represent the centroid of the sets $S^k$. We show that to each critical point of the $K$-means functional corresponds a solution $u = (u_1, \ldots, u_K)$ of the previous system and, vice versa, a solution of (1.1) gives a critical point of the functional and therefore a centroidal Voronoi tessellation of $\Omega$ provided by $\{S^k\}_{k=1}^K$.

The previous PDE approach can be extended to centroidal Voronoi tessellations related to a general convex metric, corresponding to HJ equations with convex, positive homogeneous Hamiltonians. Moreover, a system of HJ equations similar to (1.1) provides a way to compute the optimal weights for the capacity-constrained problem, which aims to find a tessellation of the domain with regions of a given area. This problem arises in several applications in economy, and it is connected with the so-called semi-discrete Optimal Transport problem ([18, 20]).

Minima of functionals associated with partition problems can be computed
via the Lloyd algorithm. This simple iterative procedure alternates computation of a new Voronoi tessellation corresponding to an old set of generators with the calculation of a new set of generators as the centers of the computed Voronoi diagrams. It can be shown that the Lloyd algorithm converges to a critical point of the $K$-means functional ([13]). We consider a PDE version of Lloyd algorithm for the solution of (1.1), where the step corresponding to the computation of the Voronoi diagrams is performed solving the HJ equation via a Fast Marching technique. As we discuss later, smart management of the data and the use of acceleration techniques may considerably speed up the process.

We mention that a different approach to the hard clustering problem based on the theory of Mean Field Games has been pursued in [10]. An algorithm for the computation of the centroidal Voronoi diagram via the resolution of HJ equations is considered in [22].

The paper is organized as follows. In Section 2, we introduce the PDE approach to hard-clustering analysis in the framework of the euclidean distance and the corresponding theory of centroidal Voronoi tessellations. In Section 3, we generalize the previous approach to a general class of distance functions associated with convex HJ equations. In Section 4, we introduce a system of HJ equations which characterize centroidal power diagrams, a generalization of centroidal Voronoi tessellations where the measure of the cells is prescribed. Each section contains the corresponding numerical tests.

## 2 $K$-means, mixture models and MFG theory

This section reviews the $K$-means problem, the Gaussian mixture model, and their relation with MFG theory. Moreover, we reinterpret the convergence of the latter soft-clustering probabilistic model to the former hard-clustering one as a vanishing viscosity limit in the MFG system. In all the paper, we assume that $\Omega$ is a bounded subset of $\mathbb{R}^d$ and $\rho$ is a density function supported in $\Omega$, i.e., $\rho \geq 0$ and $\int_\Omega \rho dx = 1$, representing the distribution of the points of a given data set $\mathcal{X}$. 


2.1 The $K$-means problem and the Lloyd algorithm

The $K$-means problem aims to minimize the functional

$$
I(y_1, \ldots, y_K) = \sum_{k=1}^{K} \int_{V(y_k)} |x - y_k|^2 \rho(x) dx,
$$

(2.1)

where $V(y_k) = \{x \in \mathbb{R}^d : |x - y_k| = \min_{j=1,\ldots,K} |x - y_j|\}$. (2.2)

A minimum of the functional $I$ provides a clusterization of the data set, i.e., a partition of $\mathcal{X}$ into $K$ disjoint clusters $V(y_k)$ such that each data point belongs to the cluster with the nearest centroid $y_k$. This property can be expressed in the elegant terminology of the centroidal Voronoi tessellations (see [12, 13]).

Critical points of $I$ can be computed via the Lloyd algorithm, a simple two steps iterative procedure. Starting from an arbitrary initial set of generators, at each iteration the following two steps are performed

- construct the Voronoi tessellation $\{V(y_i)\}_{i=1}^{K}$;
- take the centroids of $\{V(y_i)\}_{i=1}^{K}$ as the new set of generators.

The procedure is repeated until an appropriate stopping criterion is met. At each iteration, the objective function $I$ decreases and the algorithm converges to a (local) minimum of (2.1) (see [13, Theorem 2.3]).
2.2 Gaussian mixture models and Mean Field Games theory

The $K$-means functional and the Lloyd algorithm can be seen as the limit of the maximum likelihood functional and the Expectation-Maximization (EM in brief) algorithm for a Gaussian mixture model when the variance parameter is sent to 0 (see [5, Chapter 7]). Consider a Gaussian mixture model

$$m(x) = \sum_{k=1}^{K} \alpha_k \mathcal{N}(x; \mu_k, \Sigma_k), \quad \text{with } \alpha_k \in (0, 1), \quad \sum_{k=1}^{K} \alpha_k = 1, \quad (2.4)$$

where $\mu_k$ and $\Sigma_k$ denote mean and covariance matrix of the Gaussian distribution $\mathcal{N}(x; \mu_k, \Sigma_k)$. The aim is to find the parameters $\alpha = (\alpha_1, \ldots, \alpha_K), \mu = (\mu_1, \ldots, \mu_K), \Sigma = (\Sigma_1, \ldots, \Sigma_K)$ in order to maximize the log-likelihood functional of the mixture model with respect to the data set $X$

$$\mathcal{L}(\alpha, \mu, \Sigma; X) = \int_{\mathbb{R}^d} \sum_{k=1}^{K} \gamma_k(x) \{\ln(\alpha_k) + \ln(\mathcal{N}(x; \mu_k, \Sigma_k))\} \rho(x) dx, \quad (2.5)$$

where

$$\gamma_k(x) = \frac{\alpha_k \mathcal{N}(x; \mu_k, \Sigma_k)}{\sum_{j=1}^{K} \alpha_j \mathcal{N}(x; \mu_j, \Sigma_j)}$$

are the responsibilities, or posterior probabilities. The previous maximization problem can be solved via the EM algorithm. Starting from an arbitrary initialization $\mu_0^k, \Sigma_0^k, \alpha_0^k, k = 1, \ldots, K$, the iterative procedure alternates two steps:

- In the Expectation step, given $\mu_{n-1}^k, \Sigma_{n-1}^k, \alpha_{n-1}^k, k = 1, \ldots, K$, the responsibilities are updated by setting

$$\gamma_n^k(x) = \frac{\alpha_{n-1}^k \mathcal{N}(x; \mu_{n-1}^k, \Sigma_{n-1}^k)}{\sum_{j=1}^{K} \alpha_{n-1}^j \mathcal{N}(x; \mu_{n-1}^j, \Sigma_{n-1}^j)}.$$

- In the Maximization step, given the responsibilities $\gamma_n^k, k = 1, \ldots, K$, the mixture parameters of (2.4) are updated by setting

$$\alpha_n^k = \int \rho(x) \gamma_n^k(x) dx,$$

$$\mu_n^k = \frac{\int x \rho(x) \gamma_n^k(x) dx}{\int \gamma_n^k(x) \rho(x) dx}, \quad \Sigma_n^k = \frac{\int \gamma_n^k(x)(x - \mu_n^k)(x - \mu_n^k)^t dx}{\int \gamma_n^k(x) \rho(x) dx}.$$

It can be shown that the previous procedure increasingly converges to a (local) maximum point of the functional $\mathcal{L}$ as $n \to \infty$. Moreover the Lloyd
algorithm for K-means problem can be interpreted as a particular case of the EM algorithm for \( \gamma_k^n(x) \) in the M-step replaced by the characteristic function of the set where \( \gamma_k^n(x) \) is maximum with respect to the indices \( k = 1, \ldots, K \).

Assume now that the covariance matrices in (2.4) are independent of \( k \) and given by \( \Sigma_k = \sigma I \), where \( I \) is the \( d \times d \) identity matrix and \( \sigma > 0 \) a constant. For \( \sigma \to 0 \), the responsibility

\[
\gamma_k(x) = \frac{\alpha_k e^{-\frac{|x-\mu_k|^2}{2\sigma}}}{\sum_{j=1}^K \alpha_j e^{-\frac{|x-\mu_j|^2}{2\sigma}}},
\]

corresponding to the index \( k \) which minimizes the distance of the mean \( \mu_k \) from \( x \), tends to 1, while, for \( j \neq k \), all the other responsibilities \( \gamma_j(x) \) tend to 0. Hence, at least formally, \( \gamma_k \to \mathbb{1}_{V(y_k)}(x) \), where \( \mathbb{1}_{V(y_k)}(x) \) is the characteristic function of the set \( V(y_k) \) defined as in (2.2). Moreover the Gaussian distributions \( \mathcal{N}(x; \mu_k, \sigma I) \) tend to Dirac functions concentrated in the centroids \( \mu_k \) and the maximum likelihood functional \( L \) tends, up to a constant, to the opposite of the K-means functional \( I \) defined in (2.1) (see [5, Sec. 9.3.2]).

We now describe the connection between the Gaussian mixture model (2.4) and a system of partial differential equations arising in the theory of Mean Field Games (we refer to [9, 17, 16] for an introduction to this theory). In [3], it has been shown that the critical points of the log-likelihood functional (2.5) can be characterized by means of the multi-population MFG system

\[
\begin{align*}
&\varepsilon \Delta u_{k,\varepsilon} + \frac{1}{2} |D u_{k,\varepsilon}|^2 + \lambda_{k,\varepsilon} = \frac{\varepsilon^2}{2} (x - \mu_{k,\varepsilon})^t (\Sigma_{k,\varepsilon}^{-1})^t \Sigma_{k,\varepsilon}^{-1} (x - \mu_{k,\varepsilon}), & x \in \mathbb{R}^d, \\
&\varepsilon \Delta m_{k,\varepsilon} + \text{div}(m_{k,\varepsilon} D u_{k,\varepsilon}) = 0, & x \in \mathbb{R}^d, \\
&\alpha_{k,\varepsilon} = \int_{\mathbb{R}^d} \gamma_{k,\varepsilon}(x) \rho(x) dx, \\
&m_{k,\varepsilon} \geq 0, \int_{\mathbb{R}^d} m_{k,\varepsilon} dx = 1, u_{k,\varepsilon}(\mu_{k,\varepsilon}) = 0,
\end{align*}
\]

(2.6)
for \( k = 1, \ldots, K \), where

\[
\gamma_{k,\varepsilon}(x) = \frac{\alpha_{k,\varepsilon} m_{k,\varepsilon}(x)}{\sum_{j=1}^{K} \alpha_{j,\varepsilon} m_{j,\varepsilon}(x)},
\]

\[
\mu_{k,\varepsilon} = \frac{\int_{\mathbb{R}^d} x \gamma_{k,\varepsilon}(x) \rho(x) dx}{\int_{\mathbb{R}^d} \gamma_{k,\varepsilon}(x) \rho(x) dx},
\]

\[
\Sigma_{k,\varepsilon} = \frac{\int_{\mathbb{R}^d} (x - \mu_{k,\varepsilon}) (x - \mu_{k,\varepsilon})^t \gamma_{k,\varepsilon}(x) \rho(x) dx}{\int_{\mathbb{R}^d} \gamma_{k,\varepsilon}(x) \rho(x) dx}.
\]

More precisely, a solution of (2.6) is given by a family of quads (\( u_{k,\varepsilon}, \lambda_{k,\varepsilon}, m_{k,\varepsilon}, \alpha_{k,\varepsilon} \)), \( k = 1, \ldots, K \), with

\[
u_{k,\varepsilon}(x) = \frac{\varepsilon}{2} (x - \mu_{k,\varepsilon})^{t} \Sigma_{k,\varepsilon}^{-1} (x - \mu_{k,\varepsilon}), \quad \lambda_{k,\varepsilon} = \varepsilon^2 \text{Tr}(\Sigma_{k,\varepsilon}^{-1}),
\]

\[
m_{k,\varepsilon}(x) = \mathcal{N}(x; \mu_{k,\varepsilon}, \Sigma_{k,\varepsilon}) = C_k e^{-\frac{u_{k,\varepsilon}(x)}{\varepsilon^2}},
\]

\[
\alpha_{k,\varepsilon} = \int_{\mathbb{R}^d} \gamma_{k,\varepsilon}(x) \rho(x) dx,
\]

and the corresponding parameters (\( \alpha_{k,\varepsilon}, \mu_{k,\varepsilon}, \Sigma_{k,\varepsilon} \)), \( k = 1, \ldots, K \), are a critical point of the log-likelihood functional (2.5). Note that in general the solution of (2.6) is not unique.

In cluster analysis, the responsibilities can be used to assign a point to the class with the highest \( \gamma_{k,\varepsilon} \), i.e. the data set is divided into the disjoint subsets

\[
S_{u,\varepsilon}^k = \{ x \in \mathbb{R}^d : \gamma_{k,\varepsilon}(x) = \max_{j=1,\ldots,K} \gamma_{j,\varepsilon}(x) \}.
\]

Taking into account (2.7) and the definition of \( m_{k,\varepsilon} \) in (2.10), we see that the clusters \( S_{u,\varepsilon}^k \) can be equivalently defined as

\[
S_{u,\varepsilon}^k = \{ x \in \mathbb{R}^d : u_{k,\varepsilon}(x) = \min_{j=1,\ldots,K} u_{j,\varepsilon}(x) \}.
\]

### 2.3 A system of Hamilton-Jacobi equations for the K-means problem

In order to deduce a PDE characterization for centroidal Voronoi tessellations, we interpret the convergence of maximum likelihood functional to the \( K \) means functional in term of the limit in the MFG system (2.6). Assuming that \( \Sigma_k = \sigma I \) and passing to the limit in (2.6) for \( \varepsilon, \sigma \to 0^+ \) in such a way that \( \varepsilon/\sigma^2 \to 1 \), we observe that the responsibility \( \gamma_{k,\varepsilon} \) in (2.7) converges to
the characteristic function of the set where $\alpha_k m_k$ is maximum with respect to $\alpha_j m_j$, $j = 1, \ldots, K$ or, equivalently, where $u_k$ is minimum with respect to $u_j$. Hence, we formally obtain that (2.6) converges to the first order multi-population MFG system

\[
\begin{cases}
\frac{1}{2}|Du_k|^2 + \lambda_k = \frac{1}{2}|x - \mu_k|^2, & x \in \mathbb{R}^d, \\
\text{div}(m_k Du_k(x)) = 0, & x \in \mathbb{R}^d, \\
\alpha_k = \int_{\mathbb{R}^d} 1_{S^k_u}(x) \rho(x) dx, \\
m_k \geq 0, \int_{\mathbb{R}^d} m_k(x) dx = 1, u_k(\mu_k) = 0,
\end{cases}
\tag{2.12}
\]

for $k = 1, \ldots, K$, with

\[
S^k_u = \{ x \in \mathbb{R}^d : u_k(x) = \min_{j=1,\ldots,K} u_j(x) \},
\tag{2.13}
\]

\[
\mu_k = \frac{\int_{\mathbb{R}^d} x 1_{S^k_u}(x) \rho(x) dx}{\int_{\mathbb{R}^d} 1_{S^k_u}(x) \rho(x) dx}.
\tag{2.14}
\]

In the following, solution of HJ equations will be always intended in viscosity solution sense [4]. The coupling among the $K$ systems in (2.12) is in the definition of the subsets $S^k_u$. The coefficient $\alpha_k$ represents the fraction of the data set contained in the cluster $S^k_u$. The ergodic constant $\lambda_k$, which can be characterized as the supremum of the real number $\lambda$ for which the HJ equation admits a subsolution (see [4]), is equal to 0. Hence, the solution of the first equation with the condition $u_k(\mu_k) = 0$ is given by $u_k(x) = |x - \mu_k|^2/2$. The solution, in the sense of distribution, of the Fokker-Planck equation is given by $m_k = \delta_{\mu_k}(\cdot)$, where $\delta_{\mu_k}$ denotes the Dirac function in $\mu_k$, and the HJ equations are independent of $m_k$ and $\alpha_k$.

Recalling that the unique viscosity solution of the problem

\[
\begin{cases}
|Du| = 1 & x \in \mathbb{R}^d, \\
u(\mu) = 0
\end{cases}
\]

is given by $u(x) = |x - \mu|$, we can write the following simplified version of (2.12)

\[
\begin{cases}
|Du_k| = 1 & x \in \mathbb{R}^d, \\
u_k(\mu_k) = 0, \\
S^k_u = \{ x \in \mathbb{R}^d : u_k(x) = \min_{j=1,\ldots,K} u_j(x) \}, \\
\mu_k = \frac{\int_{\mathbb{R}^d} x 1_{S^k_u}(x) \rho(x) dx}{\int_{\mathbb{R}^d} 1_{S^k_u}(x) \rho(x) dx}
\end{cases}
\tag{2.15}
\]
for \( k = 1, \ldots, K \), which is a system of \( K \) HJ equations coupled through the sets \( S_u^k \). For each \( k \), the first two conditions in (2.15) imply that \( u_k \) is the Euclidean distance from \( \mu_k \), the third condition determine the Voronoi diagram corresponding to the generator \( \mu_k \), while the last condition entails \( \mu_k \) being the centroid of \( S_u^k \).

We now show that the multi-population MFG system (2.15) characterizes critical points of the functional (2.1) or, equivalently, CVTs of the set \( \Omega \).

**Proposition 2.2.** We have:

(i) Let \((y_1, \ldots, y_K)\) be a critical point of the functional \( I \) in (2.1) with Voronoi diagrams \( V(y_k) \). Then, there exists a solution of (2.15) such that \( \mu_k = y_k \) and \( S_u^k = V(y_k) \).

(ii) Given a solution \( u = (u_1, \ldots, u_K) \) of (2.15), then \((\mu_1, \ldots, \mu_K)\) is a critical point of \( I \) with Voronoi diagrams \( V(y_k) = S_u^k \).

**Proof.** Assume that \((y_1, \ldots, y_K)\) is a critical point of the functional \( I \), hence

\[
y_k = \frac{\int_{\mathbb{R}^d} x \rho(x) \mathbbm{1}_{V(y_k)}(x) dx}{\int_{\mathbb{R}^d} \rho(x) \mathbbm{1}_{V(y_k)}(x) dx}, \quad \forall k = 1, \ldots, K, \tag{2.16}
\]

with \( V(y_k) \) defined as in (2.2). Consider the function \( u = (u_1, \ldots, u_K) \) defined by

\[
u_k(x) = |x - y_k|, \tag{2.17}\]

Then \( u = (u_1, \ldots, u_K) \) is a solution of the HJ equations in (2.15) with \( \mu_k = y_k \). Moreover, by (2.2) and (2.13), we have

\[
S_u^k = \{ x \in \mathbb{R}^d : |x - y_k| = \min_{j=1,\ldots,K} |x - y_j| \} = V(y_k)
\]

and therefore

\[
y_k = \frac{\int_{V(y_k)} x \rho(x) dx}{\int_{V(y_k)} \rho(x) dx} = \frac{\int_{\mathbb{R}^d} x \rho(x) \mathbbm{1}_{S_u^k}(x) dx}{\int_{\mathbb{R}^d} \rho(x) \mathbbm{1}_{S_u^k}(x) dx}.
\]

We conclude that \( u = (u_1, \ldots, u_K) \) in (2.17) gives a solution of (2.15).

Now assume that \( u = (u_1, \ldots, u_K) \) is a solution of (2.15). Then \( u_k(x) = |x - \mu_k| \) with

\[
\mu_k = \frac{\int_{\mathbb{R}^d} x \mathbbm{1}_{S_u^k}(x) \rho(x) dx}{\int_{\mathbb{R}^d} \mathbbm{1}_{S_u^k}(x) \rho(x) dx}.
\]

Since \( S_u^k = \{ x \in \mathbb{R}^d : |x - y_k| = \min_{j=1,\ldots,K} |x - y_j| \} \), we conclude that \( \mu_k \) satisfies (2.16) with \( V(y_k) = S_u^k \). \(\Box\)
The previous result can be restated in the terminology of the Voronoi tessellation, saying that a solution of the system (2.15) determines a CVT and vice versa, see (2.3). We have the following existence result for (2.15)

**Theorem 2.3.** Let $\rho$ be a positive and smooth density function defined on a smooth bounded set $\Omega$. Then, there exists a solution $u$ to (2.15). Moreover, any limit point of the Lloyd algorithm determine a solution of (2.15).

**Proof.** The first assertion is consequence of the existence of critical points of $I$ and the equivalence result provided by Prop. 2.2. The second part of the statement follows from the convergence of the Lloyd algorithm and standard stability results in viscosity solutions theory. $\blacksquare$

### 2.4 A PDE algorithm for the K-means problem

We describe and test a method for the $K$-means problem obtained by combining the numerical approximation of the system (2.15) with the Lloyd algorithm.

We introduce a regular triangulation of $\Omega$, the support of $\rho$, given by a collection of $N$ disjoint triangles $T := \{T_i\}_{i=1}^N$. We denote with $\Delta x$ the maximal area of the triangles, i.e. $\max_{i=1,...,N} |T_i| < \Delta x$, and we assume that $\Omega \subseteq \bigcup_1^N T_i \approx \Omega$. We denotes with $G := \{X_i\}_{i=1,...,N}$ the set of the centroids of the triangles $T_i$ and, for a piecewise linear function $U : G \to \mathbb{R}$, we set $U_i := U(X_i)$.

For the approximation of the HJ equation in (2.15), we consider the semi-Lagrangian monotone scheme

$$G_i(U) = \min_{a \in B(0,1)} \{ \|U\|(X_i - ha) + h \},$$

where $h$ is a fictitious-time parameter (generally taken of order $O(\sqrt{\Delta x})$, see [15] for details), and $\|U\|$ a standard linear interpolation operator on the simplices of the triangulation.

Our algorithm is a two steps iterative procedure. Starting from an arbitrary assignment $\mu^{(0)} = (\mu^{(0),1}, \ldots, \mu^{(0),K})$ for the centroids, we iterate

1. For $k = 1, \ldots, K$ and $j_k = \arg\min_{i=1,...,N} |X_i - \mu^{(n),k}|$, solve the problem

$$\begin{cases}
G_i(U^{(n),k}) = 1, & i = 1, \ldots, N, \\
U^{(n),k}_{j_k} = 0,
\end{cases}$$

(2.18)
(recall that $U_{j_k}^{(n),k}$ denotes the value at the $n$-th iteration of the approximate solution of the $k$-th equation at point $X_{j_k}$), and define

$$S^{(n+1),k} = \bigcup \left\{ T_i : i \text{ is s.t. } U_i^{(n),k} = \min_{j=1,...,K} U_{j_i}^{(n),j} \right\}.$$  

(ii) Compute the new centroids points

$$\mu^{(n+1),k} = \frac{\sum_{T_i \in S^{(n+1),k}} X_i |T_i| \rho(X_i)}{\sum_{T_i \in S^{(n+1),k}} |T_i| \rho(X_i)}.$$  

(2.19)

We iterate these two steps till meeting a stopping criterion as

$$\max_k \{||\mu^{(n+1),k} - \mu^{(n),k}||\} < \varepsilon.$$  

Test 1. The first test is a simple problem to check the basic features of the technique. We consider a circular domain $\Omega := B(0, 1)$ and we consider a CVT composed of 6 cells. The density function $\rho$ is chosen uniformly distributed on $\Omega$, i.e. $\rho(x) = 1/|\Omega|$, where $|\Omega| = \pi$. We set the approximation parameter $\Delta x = 0.004$ and the stopping criterion $\varepsilon = \Delta x/10$. Figure 1 shows tessellations computed by the algorithm starting from different sets of initial centroids. The evolution of the centroids is marked in red with a sequential number related to the iteration number. We can observe that in all the cases, the centroids move from the initial guess toward an optimal tessellation of the domain, where the optimality is intended referred to the functional (2.1). The convergence toward optimality is highlighted in the last picture in Figure 1, where the value of the K-means functional is evaluated at the end of every iteration for the previous three cases.

Test 2. We consider a bounded domain $\Omega$ given by the union of two squares $[0, 1] \times [0, 1], [-1, 0] \times [-1, 0]$ and a section of a circle $B(0, 1) \cap [-1, 0] \times [0, 1]$ and we remove by the domain the circle $B([-0.4, 0.4], 0.2)$, as displayed in Figure 2. Then, a CVT of $\Omega$ given by three cells, i.e. $K = 3$, is computed.

At first, the density function $\rho$ is given by a uniform distribution on $\Omega$, i.e. $\rho(x) = 1/|\Omega|$, where $|\Omega| = (2 + \pi/4) - \pi(1/5)^2 \approx 2.66$.

In Figure 2 we see the evolution of the centroids $\mu^{(n)}$ starting from the initial position

$$\mu^{(0)} = ([-0.6, -0.6], [-0.4, -0.6], [-0.4, 0]).$$
Figure 1: Three Voronoi tessellations with $K = 6$ computed starting from different initial centroids and $\Delta x = 0.004$, above/lef: $\mu^{(0)} = ([0.4, 0.6], [0.6, 0.4], [0.6, -0.4], [-0.4, -0.6], [-0.6, -0.4], [-0.6, 0.4])$; above/right: $\mu^{(0)} = ([0.4, 0.6], [0.6, 0.4], [0.6, -0.4], [-0.6, -0.4], [0.1, 0.6], [0.1, 0.1])$; bottom/lef: $\mu^{(0)} = ([0.4, 0.6], [0.6, 0.4], [-0.4, 0.6], [-0.4, -0.6], [-0.6, -0.4], [0.1, 0.1])$; bottom/right: evolution of the $K$–means functional for iteration step of the algorithm.

The two images in the top panels of Figure 2 are relative to different discretization parameters $\Delta x := \max |T_i|$ and $\varepsilon = \Delta x/10$. We underline how the number of iterations does not increase much for a smaller stopping parameter, e.g., setting $\varepsilon$ to $10^{-6}$ we obtained numerical convergence for $n = 11$. Moreover, the approximation of the position of the centroids $\mu^{(n)}$,
Figure 2: Above: uniformly distributed $\rho$, left: $\Delta x = 0.01$, right: $\Delta x = 0.001$. Bottom: $\rho$ is a multivariate normal distribution around $[0.5, 0.5]$, left: $\Delta x = 0.01$; right: $\Delta x = 0.001$.

once reached convergence, is sufficiently accurate even in the presence of a discretization parameter $\Delta x$ relatively coarse. This suggests, at least in this example, avoiding excessive refinement of $\Delta x$ to prevent increasing computational cost for the algorithm.
Even in this easy case, we can observe an additional feature of the
method: the approximation of the critical points is monotone on the func-
tional $I$ while a point may have a non-monotone migration toward the cor-
rect approximation. This is because the evolution of $I$ in the algorithm is
monotone (cf. Fig 1 of the previous test) at any iteration, but not for a
single centroid.

We complete this test with a case where $\rho$ is not constant. Consider a
multivariate normal distribution around the point $[0.5, 0.5]$ and covariance
matrix $I$, i.e.,

$$\rho(x) = \frac{1}{2\pi|\Omega|} e^{-\frac{(x_1 - 0.5)^2 - (x_2 - 0.5)^2}{2}}.$$

The results are shown in the bottom panels of Figure 2, with the same
choice of the parameters as in the previous test. We observe, as expected,
a reduction of the dimension of the sets $S_k^{(n)}$ in correspondence to higher
values of the density function $\rho$. Even if we need few more steps to reach
the numerical convergence, the algorithm shows similar performances and
stops for $n = 13$.

**Remark 2.4.** The previous numerical procedure may be computationally
expansive, with the bottleneck given by the resolution of $K$-eikonal equations
on the whole domain of interest, see (2.18). In some cases, the first step
of Lloyd algorithm may turn to be very expansive, in particular if we use,
to solve (2.18), a value iteration method, i.e. a fixed point iteration on the
whole computational domain (see for details [14]).

This aspect may be considerably mitigated with the use of a more rational
way to process the various parts of the domain, as in the case of Fast March-
ing methods (see [23]). In those methods, the nodes of the discrete grid are
processed ideally only once, thanks to the information about the character-
istics of the problem that may be derived by the same updating procedure.
The case of unstructured grids is slightly more complicated than the stan-
dard one, and it requires an updating procedure that consider the geometry
of the triangles of the grid. We refer to [24] for a precise description of the
algorithm in this case.

### 3 Geodesic Voronoi tessellations and HJ equations

In this section, motivated by the relation between (2.15) and CVT in the
Euclidean case, we introduce a system of HJ equations which characterize
CVT for a general class of convex metrics.
Consider a set-valued map \( x \mapsto C(x) \subset \mathbb{R}^d \) and assume that

(H1) for each \( x \in \mathbb{R}^d \), \( C(x) \) is a compact, convex set and \( 0 \in C(x) \);

(H2) there exists \( L > 0 \) such that \( d_H(C(x), C(y)) \leq L|x - y| \), for all \( x, y \in \mathbb{R}^d \);

(H3) there exists \( \delta > 0 \) such \( B(0, \delta) \subset C(x) \) for any \( x \in \mathbb{R}^d \).

(here \( d_H \) denotes the Hausdorff distance). For \( x, y \in \mathbb{R}^d \), let \( \mathcal{F}_{x,y} \) be the set of all the trajectories \( X(\cdot) \) defined by the differential inclusion

\[
\dot{X}(t) \in C(X(t)), \; X(0) = x, \; X(T) = y,
\]

for some \( T = T(X(\cdot)) > 0 \). Note that, because of the assumptions on the map \( C(x) \), \( \mathcal{F}_{x,y} \) is not empty. The function \( d_C : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \), defined by

\[
d_C(x,y) = \inf_{\mathcal{F}_{x,y}} T(X(\cdot)), \tag{3.1}
\]

is a distance function, equivalent to the Euclidean distance (see [8]). Define a Hamiltonian \( H : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \) as the support function of the convex set \( C \), i.e.

\[
H(x,p) = \sup_{q \in C(x)} p \cdot q. \tag{3.2}
\]

Then, \( H : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \) is a continuous function and satisfies the following properties

\[
H(x,0) = 0, \; H(x,p) \geq 0;
\]

\( H(x,p) \) is convex and positive homogeneous in \( p \),

i.e. for \( \lambda > 0 \), \( H(x, \lambda p) = \lambda H(x, p) \);

\[
|H(x,p) - H(x,q)| \leq L|x - y|(1 + |p|).
\]

Moreover, for any \( y \in \mathbb{R}^d \), the function \( u : \mathbb{R}^d \to \mathbb{R} \), defined by \( u(x) = d_C(y,x) \), is the unique viscosity solution of the problem

\[
\begin{aligned}
\left\{ \\
H(x,Du) = 1, \quad x \in \mathbb{R}^d, \\
u(y) = 0.
\end{aligned} \tag{3.3}
\]

**Remark 3.1.** Some examples of distance defined by a convex map \( C \) are

1. if \( C(x) = \{ p \in \mathbb{R}^d : \|p\|_s = (\sum_{i=1}^d |p_i|^s)^{1/s} \leq 1 \} \) for \( s > 1 \), then \( d_C \) is the Minkowski distance \( d_C(x,y) = \|x - y\|_s \) and \( H(x,p) = \|p\|^2/\|p\|_s; \)
2. if \( C(x) = a(x)B(0,1) \), where \( a(x) \geq \delta > 0 \), then \( H(x,p) = a(x)|p| \);

3. if \( C(x) = A(x)^{1/2}B(0,1) \), where \( A \) is a positive definite matrix such that \( A(x)\xi \cdot \xi \geq \delta > 0 \) for \( \xi \in \mathbb{R}^d \), then \( d_C \) is the Riemannian distance induced by the matrix \( A \) on \( \mathbb{R}^d \) and \( H(x,p) = \sqrt{A(x)p \cdot p} \).

Moreover it is possible to consider the distance function corresponding to a Hamiltonian \( H \) defined by

\[
H(x,p) = \max\{H_1(x,p), \ldots, H_N(x,p)\},
\]

where \( H_n, n = 1, \ldots, N \) are Hamiltonians of the types above.

Given a set of generators \( \{y_k\}_{k=1}^{K}, y_k \in \overline{\Omega} \), we define a geodesic Voronoi tessellation of \( \Omega \) as the union of the geodesic Voronoi diagrams

\[
V(y_k) = \{ x \in \Omega : d_C(x,y_k) = \min_{j=1,\ldots,K} d_C(x,y_j) \} \quad (3.4)
\]

(the point of \( V(y_k) \cap V(y_j) \) are assigned to the diagram with the smaller index).

**Definition 3.2.** A geodesic Voronoi tessellation \( \{V(y_k)\}_{k=1}^{K} \) of \( \Omega \) is said to be a geodesic centroidal Voronoi tessellation (GCVT in short) if, for each \( k = 1, \ldots, K \), the generator \( y_k \) of \( V(y_k) \) coincides with the centroid of \( V(y_k) \), i.e.

\[
\int_{V(y_k)} \rho(x)d_C(y_k,x)dx = \min_{z \in V(y_k)} \int_{V(y_k)} \rho(x)d_C(z,x)dx. \quad (3.5)
\]

Consider the functional (see \[12, Eq. (1.9)\], \[19\], \[22\])

\[
\mathcal{I}_{dc}(y_1, \ldots, y_K) = \sum_{k=1}^{K} \int_{V(y_k)} d_C(y_k,x)^2 \rho(x)dx, \quad (3.6)
\]

where \( V(y_k) \) as in (3.4). In \[19, Thereom 1\], it is proved that the previous functional is continuous and its critical points determine GCVTs of the domain \( \Omega \). Moreover, the corresponding Lloyd algorithm converges to a (local) minimum of \( \mathcal{I}_{dc} \).

We characterize GCVTs of \( \Omega \) via the following system of HJ equations

\[
\begin{align*}
H(x,Du_k) &= 1, \quad x \in \Omega, \\
u_k(\mu_k) &= 0, \\
S^k_u &= \{ x \in \mathbb{R}^d : u_k(x) = \min_{j=1,\ldots,K} u_j(x) \}, \\
\int_{S^k_u} \rho(x)u_k(x)dx &= \min \left\{ \int_{S^k_u} \rho(x)u_y(x)dx : u_y \text{ solution of (3.3) with } y \in S^k_u \right\} \quad (3.7)
\end{align*}
\]
for $k = 1, \ldots, K$. Recall that the unique solution of (3.3) is given by $u(x) = d_C(y, x)$, hence $u_k(x) = d_C(\mu_k, x)$. Furthermore, the last condition in (3.7), see also (3.5), implies that the points $\mu_k$ are the centroids of the sets $S_u^k$ with respect to the metric $d_C$. On the other hand, the HJ equations are coupled via the points $\mu_1, \ldots, \mu_k$ which are the centroids of the sets $S_u^k$, $k = 1, \ldots, K$ and therefore they are unknown. Indeed, the true unknowns in system (3.7) are the points $\mu_k$, $k = 1, \ldots, K$, since they determine the functions $u_k$ as viscosity solutions of the corresponding HJ equations and consequently the diagrams $S_u^k$.

**Proposition 3.3.** The following conditions are equivalent:

(i) Let $(y_1, \ldots, y_K)$ be a critical point of the functional $I_{d_C}$ in (3.6) with geodesic Voronoi diagrams $V(y_k)$. Then, there exists a solution of (3.7) such that $\mu_k = y_k$ and $S_u^k = V(y_k)$.

(ii) Given a solution $u = (u_1, \ldots, u_K)$ of (3.7), then $(\mu_1, \ldots, \mu_K)$ is a critical point of $I_{d_C}$ with geodesic Voronoi diagrams $V(y_k) = S_u^k$.

The proof is similar to the one of Prop. 2.2 and we omit it. As consequence of the results in [19], we have an existence result for (3.7).

**Theorem 3.4.** Let $\rho$ be a positive and smooth density function defined on a smooth bounded set $\Omega$. Then, there exists a solution to (3.7). Moreover, any limit point of the Lloyd algorithm correspond to a solution of (3.7).

We observe that, in terms of (3.7), the Lloyd algorithm can be rewritten as follows: given an initial guess $(\mu^{(0)}_1, \ldots, \mu^{(0)}_K)$, at the $(n)^{th}$-step,

- Solve the $K$ (uncoupled) HJ equations

$$
\begin{cases}
H(x, Du_k^{(n)}) = 1, \\
u_k^{(n)}(\mu^{(n)}, k) = 0,
\end{cases}
$$

for $k = 1, \ldots, K$ and compute the Voronoi diagrams

$$S_u^{k, (n)} = \{x \in \Omega : u_k^{(n)}(x) = \min_{j=1,\ldots,K} u_j^{(n)}(x)\}, \quad k = 1, \ldots, K.$$

- Compute the new centroids $\mu^{(n+1)}_k$ via the optimization problem

$$
\int_{S_u^{k, (n)}} \rho(x) u_{\mu^{(n+1)}, k}^{x} dx = \min \left\{ \int_{S_u^{k, (n)}} \rho(x) u_y(x) dx : u_y \text{ solution of (3.3) with } y \in S_u^{k, (n)} \right\}.
$$
In the first step of the iterative procedure, it is sufficient to solve problem (3.8) the set \( \Omega \), the support of the density \( \rho \).

### 3.1 A PDE algorithm for Geodesic centroidal Voronoi tessellations

We implement the Lloyd algorithm for GCVTs, alternating the numerical-resolution of \( K \) HJ equations and updating of the centroids points. The set \( \Omega \) is approximated as in Section 2.4. To compute the new centroids, since \( u_y(x) = d_C(y, x) \), the optimization problem (3.9) has its discrete version as

\[
\sum_{X_j \in \mathcal{S}^{(n+1),k}} \rho(X_j)d_C^2(\mu_k^{(n+1)}, X_j) = \min \left\{ \sum_{X_j \in \mathcal{S}^{(n+1),k}} \rho(X_j)d_C^2(Y, X_j) : Y \in \mathcal{S}^{(n+1),k} \right\}.
\]

and, called \( \mathcal{H}(Y) = \sum_{X_j \in \mathcal{S}^{(n+1),k}} \rho(X_j)d_C(Y, X_j) \), the maximal growth direction is (see [22])

\[
\delta_k := D\mathcal{H}(Y) = \frac{1}{2} \sum_{X_j \in \mathcal{S}^{(n+1),k}} \rho(X_j)Dd_C(Y, X_j)n_Y(X_j),
\]

where \( n_Y(x) \) is the unit vector tangent at \( Y \) to the geodesic path joining \( X_i \) to \( Y \).

To approximate the HJ equation (3.8) we consider the semi-Lagrangian scheme

\[
G_i(U) = \min_{\alpha \in C(X_i)} \{ [U](X_i - ha) + hL(X_i, \alpha) \},
\]

where \( L(x, \alpha) = \sup_{p \in \mathbb{R}^d} \{ pa - H(x, p) \} \) is the Legendre transform of \( H \).

Starting from an arbitrary assignment \( \mu^{(0)} = (\mu^{(0),1}, \ldots, \mu^{(0),K}) \) for the centroids, we iterate

(i) For \( k = 1, \ldots, K \) and \( j_k = \arg\min_{i=1, \ldots, N} |X_i - \mu^{(n),k}| \), solve the problem

\[
\begin{cases}
G_i(U^{(n),k}) = 1, & i = 1, \ldots, N, \\
U^{(n),j_k}_i = 0,
\end{cases}
\]

and define

\[
\mathcal{S}^{(n+1),k} = \bigcup \left\{ T_i : i \text{ is s.t. } U^{(n),k}_i = \min_{j=1, \ldots, K} U^{(n),j}_i \right\}.
\]
(ii) For $k = 1, \ldots, K$ compute the new centroids iterating a gradient descent search. More precisely, fixed a tolerance $\varepsilon_c > 0$, initialize $z_k = \mu^{(n),k}$ and iterate

(a) Find the value $\alpha_k$ defined as

$$
\alpha_k := \arg \min \mathcal{H}(z_k + \alpha_k \delta_k)
= \arg \min \sum_{X_j \in S^{(n+1),k}} \rho(X_j) d_C^2(z_k - \alpha_k \delta_k, X_j).
$$

(b) If $|\alpha_k| < \varepsilon_c$, set $\mu^{(n+1),k} = z_k - \alpha_k \delta_k$, otherwise set $z_k = z_k - \alpha_k \delta_k$ and go back to step (a).

We iterate these (i)-(ii) until meeting a stopping criterion as

$$
\max_k \{|\mu^{(n+1),k} - \mu^{(n),k}|\} < \varepsilon.
$$

For the numerical tests, we consider the case of the Minkowski distance on $\mathbb{R}^2$, see Remark 3.1.1 We remind that such a distance generalizes the Manhattan distance (case $s = 1$) and the Chebyshev distance (case $s \to \infty$, $d_C(x, y) = \max_i(|x_i - y_i|)$). In Figure 3 are shown the balls $B(0, 1)$ in the Minkowski distance for various values of $s$. 

Figure 3: Unitary balls $B(0, 1)$ in the Minkowski distance for various values of $s$. 

(ii) For $k = 1, \ldots, K$ compute the new centroids iterating a gradient descent search. More precisely, fixed a tolerance $\varepsilon_c > 0$, initialize $z_k = \mu^{(n),k}$ and iterate

(a) Find the value $\alpha_k$ defined as

$$
\alpha_k := \arg \min \mathcal{H}(z_k + \alpha_k \delta_k)
= \arg \min \sum_{X_j \in S^{(n+1),k}} \rho(X_j) d_C^2(z_k - \alpha_k \delta_k, X_j).
$$

(b) If $|\alpha_k| < \varepsilon_c$, set $\mu^{(n+1),k} = z_k - \alpha_k \delta_k$, otherwise set $z_k = z_k - \alpha_k \delta_k$ and go back to step (a).

We iterate these (i)-(ii) until meeting a stopping criterion as

$$
\max_k \{|\mu^{(n+1),k} - \mu^{(n),k}|\} < \varepsilon.
$$

For the numerical tests, we consider the case of the Minkowski distance on $\mathbb{R}^2$, see Remark 3.1.1 We remind that such a distance generalizes the Manhattan distance (case $s = 1$) and the Chebyshev distance (case $s \to \infty$, $d_C(x, y) = \max_i(|x_i - y_i|)$). In Figure 3 are shown the balls $B(0, 1)$ in the Minkowski distance for various values of $s$. 

19
**Test 3.** We first consider the problem on a simple L-shaped bounded domain \( \Omega = [0, 1] \times [0, 1] \cup [-1, 0] \times [-1, 1] \) as displayed in Figure 4. In this case, the Chebyshev distance (i.e. \( s \to \infty \)) provides an optimal tessellation which is trivially guessed: due to the geometrical characteristics of the domain and the distance (the contour lines of the distance from a point are squares, see Fig. 3) the solution, for \( K = 3 \) and uniform density function (\( \rho(x) = 1/|\Omega| \), where \(|\Omega| = 3\)), is simply composed by the three squares \( \{[0, 1]^2, [-1, 0] \times [-1, 0], [-1, 0] \times [0, 1]\} \), with the centroids \( \bar{\mu} = ([0.5, 0.5], [-0.5, -0.5], [-0.5, 0.5]) \).

In Figure 4, top panels, we see the evolution of the centroids \( \mu^{(n)} \) starting from the initial position \( \mu^{(0)} = ([0.5, 0.5], [-0.5, -0.5], [-0.5, 0.5]) \) for two different values of \( \Delta x \). Also in this case we can observe as the position of the centroids and the rough structure of the tessellation is correctly reconstructed even in the presence of a larger grid. This is also highlighted by the evolution of the euclidean norm of the error \( \mu^{(n)} - \bar{\mu} \) reported in Fig. 4 (bottom right). The evolution of the error and the total number of iterations necessary to converge to the correct approximation are barely affected by \( \Delta x \). On the other hand, the final approximation apparently converges to \( \bar{\mu} \) with order \( \Delta x \).

We perform the same test with a different initial position \( \mu^{(0)} \) equal to \( ([0.5, 0.5], [-0.5, -0.5], [-0.5, 0.5]) \), see Fig. 4, bottom left panel. Since in this case the optimal tessellation is unique, the algorithm converges to the same configuration. The number of iterations necessary is clearly affected by the initial guess of \( \mu \).

**Test 4.** We consider the Minkowski distance with \( s = 1 \). Since contour lines of the distance assumes a rhombus shape, we expect to be able to see this in the tessellation that we obtain. In addition, we want to show as our technique, with the help of an acceleration method, can successfully address the GCVT problem with a larger \( K \). This is not intended to be an accurate performance evaluation (which is not the main goal of this paper), but only a display of the possibilities give by the techniques proposed. We consider tesselations of \( \Omega = B(0, 1) \) with \( K = 20 \) and of \( \Omega = [-1, 1]^2 \) with \( K = 30 \). In the first case (the circle), the function \( \rho \) is given by a uniform distribution, while, in the second case, by multivariate normal distribution around the point \([0, 0]\) and covariance matrix \( I \), i.e.,

\[
\rho(x) = \rho([x_1, x_2]) = \frac{1}{2\pi|\Omega|} e^{-\frac{(x_1)^2 + (x_2)^2}{2}},
\]
Figure 4: Above: left panel $\Delta x = 0.01$; right panel: $\Delta x = 0.001$ starting from $\mu^{(0)} = ([-0.6, -0.6], [-0.4, -0.6], [-0.4, 0])$. Bottom: left panel $\mu^{(0)} = ([-0.6, 0.6], [-0.4, 0.6], [-0.5, 0.4]), \Delta x = 0.001$, right panel, convergence of the error on the centroids for iterations on the euclidean norm.

where $|\Omega| = 4$. The resulting tessellations are shown in Figure 5.

We see that our technique can address without too much troubles a problem with an higher $K$: indeed, this parameter enters in the first step
of the Lloyd algorithm linearly. Since we did not observe a substantial change of the number of iterations of the algorithm for a larger $K$, the technique remains computationally feasible, even performed on a standard laptop computer.

4 A system of HJ equations for centroidal power diagrams

In this section, we consider a generalization of centroidal Voronoi diagrams, called centroidal power diagrams. We first introduce the definition of power diagrams, or weighted Voronoi diagrams, and then describe centroidal power diagrams and a system of HJ equations that can be used to compute them. Given a set of $K$ distinct points $\{y_i\}_{i=1}^K$ in $\Omega$ and $K$ real numbers $\{w_i\}_{i=1}^K$, the power diagrams generated by the couples $(y_i, w_i)$ are defined by

$$V(y_i, w_i) = \{ x \in \Omega : |x - y_i|^2 - w_i = \min_{j=1,\ldots,K} (|x - y_j|^2 - w_j) \}. \quad (4.1)$$

As Voronoi diagrams, power diagrams provide a tessellation of the domain $\Omega$, i.e. $V(y_i, w_i) \cap V(y_j, w_j) = \emptyset$ for $i \neq j$ and $\cup_{i=1}^K V(y_i, w_i) = \overline{\Omega}$. Note that, whereas Voronoi diagrams are always non-empty, some of the power diagrams may be empty and the corresponding generators belong to another
diagram. Power diagrams reduce to Voronoi diagrams if the weights \( w_i \) coincide, but they have an additional tuning parameter, the weights vector \( w = (w_1, \ldots, w_k) \), which allows to impose additional constraints on the resulting tessellation.

A typical application of power diagrams is the problem of partitioning a given set in a capacity constrained manner (see [1]). Consider the measure \( \pi(dx) = \rho(x)dx \) supported in the set \( \Omega \), \( K \) distinct points \( \{y_i\}_{i=1}^K \) in \( \Omega \) and, to each point \( y_i \), associate a cost \( c_i > 0 \) with the property that \( \sum_{i=1}^K c_i = \pi(\Omega) \).

Given a partition of \( \Omega \) in a family of \( K \) subsets \( R_i \), define the cost of each subset as \( \int_{R_i} |x - y_i|^2 \pi(dx) \). The aim is to find a partition \( \{R_i\}_{i=1}^K \) of \( \Omega \) such that the total cost

\[
Q(R_1, \ldots, R_K) = \sum_{i=1}^K \int_{R_i} |x - y_i|^2 \pi(dx)
\]

is minimized under the constraint \( \pi(R_i) = \int_{R_i} \rho(x)dx = c_i \). It is shown in [1] that the minimum of the previous functional exists and it is reached by a power diagram. The weights of the corresponding tessellation can be found as maximizer of the concave functional

\[
\mathcal{F}(w_1, \ldots, w_k) = \sum_{i=1}^K \int_{V(y_i, w_i)} |x - y_i|^2 \rho(x)dx - \sum_{i=1}^K w_i (\pi(V(y_i, w_i)) - c_i).
\]

Moreover the gradient of \( \mathcal{F} \) is given by

\[
\frac{\partial \mathcal{F}}{\partial w_i} = c_i - \pi(V(y_i, w_i))
\]

and, if \( (w_1, \ldots, w_k) \) is a critical point of \( \mathcal{F} \), then the power diagram generated by the couples \( (y_i, w_i) \) satisfies the capacity constraint \( \pi(V(y_i, w_i)) = c_i \).

The previous optimization problem is also connected with the semi-discrete optimal mass transport problem, i.e. optimal transport of a continuous measure \( \pi \) on a discrete measure \( \nu = \sum_{i=1}^K c_i \delta_{y_i} \) (see [18, 20]). Algorithms to compute the optimal solutions are described in [11, 18, 20].

We consider centroidal power diagrams, i.e. power diagram for which generators coincide with the corresponding centroids. Indeed, it has been observed that the use of centroidal power diagrams in the capacity constrained partitioning problem avoid generating irregular or elongated cells (see [6, 25]).

**Definition 4.1.** A power diagram tessellation \( \{V(y_i, w_i)\}_{i=1}^K \) of \( \Omega \) is said to be a centroidal power diagram tessellation if, for each \( i = 1, \ldots, K \), the
generator \( y_i \) of \( V(y_i, w_i) \) coincides with the centroid of \( V(y_i, w_i) \), i.e.

\[
y_i = \frac{\int_{V(y_i, w_i)} x \rho(x) dx}{\int_{V(y_i, w_i)} \rho(x) dx}.
\]

In [25], a centroidal power diagram which satisfies the capacity constraints \( \pi(V(y_i, w_i)) = c_i \), with \( \sum_{k=1}^{K} c_k = \pi(\Omega) \), is characterized as a saddle point of the functional

\[
G(y_1, \ldots, y_{k}, w_1, \ldots, w_{k}) = \sum_{i=1}^{K} \int_{V(y_i, w_i)} |x-y_i|^2 \rho(x) dx - \sum_{i=1}^{K} w_i(\pi(V(y_i, w_i)) - c_i).
\]

Note that the previous functional is similar to one defined in (4.3), but it depends also on the generators \((y_1, \ldots, y_k)\). For \((y_1, \ldots, y_k)\) fixed, \(G\) is concave with respect to \(w = (w_1, \ldots, w_K)\) and therefore it admits a maximizer which determine a power diagram \( \{V(y_i, w_i)\}_{i=1}^{K} \). For \((w_1, \ldots, w_k)\) realizing the capacity constraints \( \pi(V(y_i, w_i)) = c_i \), \(G\) coincides with the functional \(I\) in (2.1), hence it is minimized by the centroids of sets \( \{V(y_i, w_i)\}_{i=1}^{K} \).

We propose the following HJ system for the characterization of the saddle points of \(G\)

\[
\begin{aligned}
|Du_k| &= 2|x - \mu_k|^2, \quad x \in \Omega, \\
u_k(\mu_k) &= -\omega_k, \\
S^k_u &= \{ x \in \mathbb{R}^d : u_k(x) = \min_{j=1, \ldots, K} u_j(x) \} \\
\mu_k &= \frac{\int_{S^k_u} x \rho(x) dx}{\int_{S^k_u} \rho(x) dx}, \\
\pi(S^k_u) &= c_k.
\end{aligned}
\]

(4.5)

For a given \( y \in \mathbb{R}^d \), a solution of

\[
\begin{aligned}
|Du| &= 2|x - y|^2, \\
u(y) &= -\omega,
\end{aligned}
\]

(4.6)

is given by \( u_y(x) = -\omega + |x - y|^2 \). The system (4.5) depends on the \( 2K \) parameters \((\mu_k, \omega_k)\). If there exists a solution \( u = (u_1, \ldots, u_K) \) to the previous system, then \( u_k = -\omega_k + |x - \mu_k|^2 \), \( \mu_k \) is the centroid of \( S^k_u \). Moreover

\[
S^k_u = \left\{ x \in \mathbb{R}^d : -w_k + |x - y_k|^2 = \min_{j=1, \ldots, K} \{-w_j + |x - y_j|^2\} \right\} = V(y_k, w_k),
\]

24
where $V(y_k, w_k)$ is defined as in (4.1), and $\pi(S^k_u) = c_k$. Hence $\{S^k_u\}_{k=1}^K$ gives a centroidal power diagram of $\Omega$ realizing the capacity constraint.

Following [6, 7, 25], we consider the following Lloyd algorithm for (4.5). Given an initial guess $(\mu(0), \omega(0)) = (\mu(0), 1, \ldots, \mu(0), k, \omega(0), 1, \ldots, \omega(0), k)$, at $(n)^{th}$-step,

- Solve the $K$ (uncoupled) HJ equations
  \[
  \begin{aligned}
  |Du^{(n)}_k| &= 2|x - \mu^{(n-1), k}|, \quad x \in \Omega, \\
  u^{(n)}_k(\mu^{(n-1), k}) &= -\omega^{(n-1), k}, \quad k = 1, \ldots, K
  \end{aligned}
  \]
  and compute the power diagram of generators $(\mu^{(n-1), k}, \omega^{(n-1), k})$
  \[
  S^{k, (n)}_u = \{x \in \Omega : u^{(n)}_k(x) = \min_{j=1,\ldots,K} u^{(n)}_j(x)\}.
  \]

- Compute the new centroids
  \[
  \mu^{(n), k} = \frac{\int_{S^{k, (n)}_u} x \rho(x) dx}{\int_{S^{k, (n)}_u} \rho(x) dx}, \quad k = 1, \ldots, K
  \]

- Compute the new weights $\omega^{(n), k}$, $k = 1, \ldots, K$, by maximizing the functional
  \[
  \sum_{k=1}^K \int_{S^{k, (n)}_u} |x - y_k|^2 \rho(x) dx - \sum_{i=1}^k w_k(\pi(S^{k, (n)}_u) - c_k)
  \]

Convergence of the Lloyd algorithm for centroidal power diagram tessellations is discussed in [7].

The extension to general convex metrics can be done in a similar way (see [25]). In this case, the corresponding system of HJ equations is given by

\[
\begin{aligned}
H(x, Du_k) &= 1, \quad x \in \Omega, \\
\mu_k &= -\omega_k, \\
S^k_u &= \{x \in \mathbb{R}^d : u_k(x) = \min_{j=1,\ldots,K} u_j(x)\} \\
\int_{S^k_u} \rho(x) u_k(x) dx &= \min \{\int_{S^k_u} \rho(x) u_y(x) dx : u_y \text{ solution of (3.3) with } y \in S^k_u\}, \\
\pi(S^k_u) &= c_k.
\end{aligned}
\]

### 4.1 A PDE algorithm for centroidal power diagrams

The procedure to obtain an approximation of centroidal power diagrams contains all the tools already described in the previous sections and it includes a three steps procedure: resolution of $K$ HJ equations, update of the
centroids points and optimization step for the weights. Starting from an arbitrary assignment \((\mu^{(0)}, w^{(0)}) = (\mu^{(0),1}, \ldots, \mu^{(0),K}, w^{(0),1}, \ldots, w^{(0),K})\) for the centroids and the weights, we iterate

(i) For \(k = 1, \ldots, K\) and \(j_k = \arg\min_{i=1,\ldots,N} \|X_i - \mu^{(n),k}\|\), solve the problem
\[
\begin{align*}
G_i(U^{(n),k}) &= 2\|X_k - \mu^{(n),k}\|, \quad i = 1, \ldots, N, \\
U_{j_k}^{(n),k} &= -w_k^{(n-1)},
\end{align*}
\] (4.7)
and define
\[
S^{(n+1),k} = \bigcup \left\{ T_i : \quad i \text{ is s.t. } U_{i}^{(n),k} = \min_{j=1,\ldots,K} U_{i}^{(n),j} \right\}.
\]

(ii) Compute the new centroids points
\[
\mu^{(n+1),k} = \frac{\sum_{T_i \in S^{(n+1),k}} X_i |T_i| \rho(X_i)}{\sum_{T_i \in S^{(n+1),k}} |T_i| \rho(X_i)}.
\] (4.8)

(iii) Compute the new weights \(w^{(n+1),k}\) as local maximum of the Lagrangian function
\[
L(Y, w) = \sum_{i=1}^{k} \sum_{X_j \in S^{(n+1),k}} \rho(X_j) d_C^2(Y, X_j) - \sum_{i=1}^{k} w_k (\pi(S^{(n+1),k}) - c_k)
\]
We iterate these three steps till meeting a stopping criterion as
\[
\max\{ |\mu^{(n+1),k} - \mu^{(n),k}|, |\omega^{(n+1),k} - \omega^{(n),k}| \} < \varepsilon.
\]

**Test 4.** We test the centroidal power diagram procedure in a simple case given by the unitary square \(\Omega = [0,1]^2\) for \(K = 6, 8\) and capacity constraint given respectively by
\[
c = (0.3, 0.25, 0.18, 0.12, 0.1, 0.05),
\]
\[
c = (0.3, 0.24, 0.15, 0.1, 0.08, 0.06, 0.05, 0.02).
\]
Clearly we have \(\bigcup_k S^{(n),k} = \Omega\), for any \(n\) and therefore we \(\sum_k c_k = |\Omega| = 1\).

The same technique is used to generate some power diagrams of more complex domains: in Figure 7 we show the optimal tessellation of a text and a rabbit-shaped domain. In the first case, the algorithm parameters are set to \(K = 8, c = (0.33, 0.22, 0.1, 0.1, 0.1, 0.05, 0.05, 0.05), \Delta x = 0.002\). In the second one, \(K = 6, c = (0.3, 0.15, 0.15, 0.15, 0.15, 0.10), \Delta x = 0.002\).
Figure 6: left panel: $K = 6$, $\Delta x = 0.015$, $\mu(0) = \{0.4, 0.5, 0.6\} \times \{0.4, 0.6\}$; right panel: $\Delta x = 0.001$, $\mu(0) = \{0.4, 0.45, 0.55, 0.6\} \times \{0.4, 0.6\}$, $c = (0.3, 0.24, 0.15, 0.1, 0.08, 0.06, 0.05, 0.02)$ (in this image some points of the evolution of the centroids are omitted for a better clarity).

5 Conclusions

As shown along with the paper, PDE theory is a robust framework to solve classic (and less traditional) tessellation problems. The main advantage of this approach is the high adaptability of the framework to specific variations of the problem (presence of constraints, non-conventional distance functions, etc.). This increased adaptability comes with a precise cost. A PDE approach is more computationally demanding than other methods available in the literature. However, the recent developments of numerical methods for nonlinear PDEs, and the increment of the accessibility to more powerful computational resources at any level, make these techniques progressively more appealing in many applicative contests.

Acknowledgment: The present research was partially supported by MIUR grant “Dipartimenti Eccellenza 2018-2022” CUP: E11G18000350001, DISMA, Politecnico di Torino.
Figure 7: Optimal power diagrams of a text and a rabbit shaped domain. The parameters are set $K = 8$, $c = (0.33, 0.22, 0.1, 0.1, 0.1, 0.05, 0.05, 0.05)$, $\Delta x = 0.002$ (left) $K = 6$, $c = (0.25, 0.15, 0.15, 0.15, 0.15, 0.10)$, $\Delta x = 0.002$ (right).

References

[1] Aurenhammer, F.; Hoffmann, F.; Aronov, B. Minkowski-type theorems and least-squares clustering, Algorithmica 20 (1998), 61-76.

[2] Aurenhammer, F.; Klein, R.; Lee, D. T. Voronoi Diagrams and Delaunay Triangulations. World Scientific Publishing Company, Singapore, 2013.
[3] Aquilanti, L.; Cacace, S.; Camilli, F.; De Maio, R. A mean field games approach to cluster analysis. Applied Math. Optim., DOI: 10.1007/s00245-019-09646-2.

[4] Bardi, M.; Capuzzo Dolcetta, I. Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman equations. Birkhäuser, Boston, 1997.

[5] Bishop, C.M. Pattern recognition and Machine Learning. Information Science and Statistics, Springer, New York, 2006.

[6] Bourne, D. P.; Kok, P. J. J.; Roper, S. M.; Spanjer, W. D. T. Laguerre tessellations and polycrystalline microstructures: a fast algorithm for generating grains of given volumes, arXiv:1912.07188.

[7] Bourne, D. P.; Roper, S. M. Centroidal power diagrams, Lloyd’s algorithm, and applications to optimal location problems. SIAM J. Numer. Anal. 53 (2015), no. 6, 2545-2569.

[8] Capuzzo Dolcetta, I. A generalized Hopf-Lax formula: analytical and approximations aspects. Geometric control and nonsmooth analysis, 136-150, Ser. Adv. Math. Appl. Sci., 76, World Sci. Publ., Hackensack, NJ, 2008.

[9] Carmona, R.; Delarue, F. Probabilistic theory of mean field games with applications. I. Mean field FBSDEs, control, and games. Probability Theory and Stochastic Modelling, 83. Springer, Cham, 2018.

[10] Coron, J.L. Quelques exemples de jeux à champ moyen. Ph.D. thesis, Université Paris-Dauphine, 2018.

[11] De Goes, F.; Breeden, K.; Ostromoukhov, V.; Desbrun, M. Blue noise through optimal transport, ACM Transactions on Graphics 31 (2012), art. no. 171.

[12] Du, Q.; Faber, V.; Gunzburger, M. Centroidal Voronoi tessellations: applications and algorithms. SIAM Rev. 41 (1999), no. 4, 637-676.

[13] Du, Q.; Emelianenko, M.; Ju, L. Convergence of the Lloyd algorithm for computing centroidal Voronoi tessellations. SIAM J. Numer. Anal. 44 (2006), no. 1, 102-119.

[14] Falcone, M.; Ferretti, R. Semi-Lagrangian approximation schemes for linear and Hamilton-Jacobi equations. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2014.
[15] Festa, A., Guglielmi, R., Hermosilla, R., Picarelli, A., Sahu, S., Sassi, A. Silva, F. J.: Hamilton-Jacobi-Bellman equations. In Optimal control: novel directions and applications, 127-261, Lecture Notes in Math., 2180, Springer, Cham, 2017.

[16] Gomes, D.; Saude, J. Mean field games - A brief survey. Dyn. Games Appl. 4 (2014), no. 2, 110-154.

[17] Lasry, J. M.; Lions, P.L. Mean Field Games, Jpn. J. Math., 2 (2007), 229-260.

[18] Lévy, B, A numerical algorithm for L2 semi-discrete optimal transport in 3D. ESAIM Math. Model. Numer. Anal. 49 (2015), no. 6, 1693–1715.

[19] Liu, Y.-J.; Yu, M.; Li, B.-J.; He, Y. Intrinsic Manifold SLIC: A Simple and Efficient Method for Computing Content-Sensitive Superpixels. IEEE Transactions on Pattern Analysis and Machine Intelligence 40 (2017), 653-666.

[20] Mérigot, Q. A multiscale approach to optimal transport, Comput. Graph. Forum 30 (2011), 1583-1592.

[21] Okabe, A.; Boots, B.; Sugihara, K.; Chiu, S. N. Spatial tessellations: concepts and applications of Voronoi diagrams. With a foreword by D. G. Kendall. Second edition. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 2000.

[22] Peyré, G.; Cohen, L. Surface Segmentation Using Geodesic Centroidal Tessellation. In Proceedings - 2nd International Symposium on 3D Data Processing, Visualization, and Transmission, (2004), 995-1002.

[23] Sethian, J. A. Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science. Cambridge Monographs on Applied and Computational Mathematics, 3. Cambridge University Press, Cambridge, 1999.

[24] Sethian, J. A.; Vladimirsky, A. Fast methods for the Eikonal and related Hamilton–Jacobi equations on unstructured meshes. Proceedings of the National Academy of Sciences 97 (2000), 5699-5703.

[25] Xin, S.-Q.; Lévy, B.; Chen, Z.; Chu, L.; Yu, Y.; Tu, C.; Wang, W. Centroidal power diagrams with capacity constraints: Computation,
applications, and extension, ACM Transactions on Graphics 35 (2016),
art. no. 244.