# Voronoi Diagrams and Their Application in the DTFE Reconstructions of the Cosmic Web

Drozdovskaya, M.N.

Maria Nikolayevna Drozdovskaya

# Voronoi Diagrams and Their Application in the DTFE Reconstructions of the Cosmic Web

Bachelor Thesis, 19 August, 2010

Thesis Supervisors: Dr. Dion Gijswijt and Professor Dr. Vincent Icke

The Mathematical Institute and Leiden Observatory,
Leiden University

# Voronoi Diagrams and Their Application in the DTFE Reconstructions of the Cosmic Web

## Computing Voronoi diagrams with Brown's algorithm. Assessing the quality of the Delaunay Tessellation Field Estimator (DTFE) reconstructions of the Cosmic Web.

M.N. Drozdovskaya

Leiden Observatory and the Mathematical Institute, University of Leiden, 2300 RA, Leiden, The Netherlands
e-mail: mndrozd@yahoo.com

19 August, 2010

**ABSTRACT**

*Context.* To get fully acquainted with Voronoi diagrams, Delaunay triangulations and the relationship between the two. To investigate a computational side of these tessellations and an application in astronomy in the form of modeling the Cosmic Web.
*Aims.* First of all, to present and understand Brown's algorithm, its tools, benefits and drawbacks. Secondly, to familiarize ourselves with modern views on and models of the Cosmic Web, and one of the new interesting tools used, namely the Delaunay Tessellation Field Estimator (DTFE). In particular, we are interested in accessing the quality of its reconstructions quantitatively.
*Methods.* Obvious key concepts are Voronoi diagrams and Delaunay triangulations. For the computational component inversion and complexity analysis are of importance. For the astronomical component, various sampling methods and Fourier transforms come into play.
*Results.* It seems that Brown's algorithm has clear benefits in higher dimensional computations, but for two and three dimensions there may be better alternatives. Even though the DTFE reconstructions of the Cosmic Web appear to be visually satisfying, it appears that it is actually very sensitive to Poisson noise in the point distribution and in principle, minor effects may seriously distort the actual underlying continuous distribution. Greater care needs to be taken to access this further. There are many further research topics open here.

**Key words.** Voronoi diagrams – Delaunay triangulations – Brown's algorithm – Cosmic Web – Delaunay Tessellation Field Estimator (DTFE)

## 1. Introduction

The core component of this paper concerns Voronoi diagrams and Delaunay triangulations. Many structures and patterns in nature, in science and in technology possess morphological resemblance to them (Davies 1996 and Okabe 2000). Over time a large amount of literature has accumulated about these topics. This papers begins by presenting all the necessary tools that we will need in order to answer our research questions. The first thing that we investigated is computing Voronoi diagrams. We studied Brown's algorithm, which is fairly efficient, but most importantly, it can be generalized to any dimension. Furthermore, since Voronoi diagrams are so tightly related with Delaunay triangulations, we can essentially compute Delaunay triangulations with it as well. After that we investigated the application of these powerful tools in astronomy. More specifically we would like to model the matter distribution in the Cosmic Web. By means of the Delaunay Tessellation Field Estimator (DTFE) we can go from a discrete point distribution to a continuous density field. This tool is very applicable for astronomy, but not only, as this is a recurring point of interest in a number of fields. The success of the DTFE has been published in recent years, we aimed at accessing its quality in reconstructing an already known continuous field. There are some very interesting points that will require strong further research. One must always remember the great cosmological implications that can be made if one is successful in modeling the continuous distribution of matter in the Cosmic Web, such as understanding the constituents of the Cosmic Web better and being able to describe the general structure of the Universe. Furthermore, it is believed that present day structures were formed by early fluctuations and are the key to understanding the origin of the Universe.

## 2. Essential Definitions and Concepts

Let us lay down the basis of this paper immediately by stating a number of definitions and concepts. These will be strictly adhered to throughout the paper.

Let us work in the $m$-dimensional Euclidean space $\mathbb{R}^m$ from now onwards, unless explicitly stated otherwise. We will use the standard metric of $\mathbb{R}^m$, when we speak about distances.

First of all, we begin with some basic definitions.

**Definition 1.** A *closed half-space* is a set of the form $G = \left\{\boldsymbol{x} \in \mathbb{R}^m \mid \boldsymbol{a}^T\boldsymbol{x} \leq b\right\}$, where $\boldsymbol{a} \in \mathbb{R}^m$, $\boldsymbol{a} \neq 0$ and $b \in \mathbb{R}$.

**Definition 2.** A *hyperplane* is a set of the form $H = \left\{\boldsymbol{x} \in \mathbb{R}^m \mid \boldsymbol{a}^T\boldsymbol{x} = b\right\}$, where $\boldsymbol{a} \in \mathbb{R}^m$, $\boldsymbol{a} \neq 0$ and $b \in \mathbb{R}$.

*Remark 3.* If $G$ and $H$ are defined as in Definitions 1 and 2, respectively, then one can say that $G$ is generated by $H$.

*Remark 4.* A hyperplane is a closed set.

**Definition 5.** An *affine subspace* of $\mathbb{R}^m$ is any set of the form $\{\boldsymbol{x} \in \mathbb{R}^m \mid A\boldsymbol{x} = b\}$, where $A$ is a $m \times k$ matrix with coefficients in $\mathbb{R}$, $b \in \mathbb{R}$ and $k \in \{0, \ldots, m\}$.

**Definition 6.** An *m-sphere* centered at $\boldsymbol{t} \in \mathbb{R}^{m+1}$ is a set of the form $S^m = \left\{\boldsymbol{x} \in \mathbb{R}^{m+1} \mid \|\boldsymbol{x} - \boldsymbol{t}\| = \rho\right\}$, where $\rho \in \mathbb{R}_{>0}$ is the *radius* of the $m$-sphere.

*Remark 7.* Observe that this is a sphere of the highest possible dimension in $\mathbb{R}^{m+1}$.

**Definition 8.** A *k-sphere*, for $0 \leq k < m$, centered at $\boldsymbol{t} \in \mathbb{R}^{m+1}$ and of radius $\rho \in \mathbb{R}_{>0}$, denoted $S^k$, is defined as the intersection between $S^m$ and an affine subspace of $\mathbb{R}^{m+1}$ of dimension $k+1$, which contains $\boldsymbol{t}$.

Secondly, we introduce the concept of convexity.

**Definition 9.** The *straight line segment* joining two points $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{R}^m$ is a set of the form $\{\lambda\boldsymbol{p} + (1-\lambda)\,\boldsymbol{q} \mid \lambda \in [0,1]\}$.

**Definition 10.** A set of points $S$ is called a *convex set* if the straight line segment joining any two points in $S$ belongs to $S$ again, i.e.: $\lambda\boldsymbol{p} + (1-\lambda)\,\boldsymbol{q} \in S\,\forall \boldsymbol{p}, \boldsymbol{q} \in S$ and $\lambda \in [0,1]$.

*Remark 11.* An intersection of convex sets is a convex set.

*Remark 12.* A closed half-space is a convex set.

*Remark 13.* A hyperplane is a convex set.

**Definition 14.** Let $S = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subset \mathbb{R}^m$, where $1 \leq N$. A *convex combination* of points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ is:

$$\boldsymbol{x} = \lambda_1\boldsymbol{x}_1 + \ldots + \lambda_N\boldsymbol{x}_N, \tag{1}$$

where $\lambda_i \in [0,1]\,\forall i \in \{1, \ldots, N\}$ and $\sum_{i=1}^{N} \lambda_i = 1$. The set of all such convex combinations is called the *convex hull* of $S$, denoted $CH(S)$, i.e.:

$$CH(S) = \left\{\sum_{i=1}^{N} \lambda_i\boldsymbol{x}_i \mid \lambda_i \in [0,1]\,\forall i \in \{1, \ldots, N\} \text{ and } \sum_{i=1}^{N} \lambda_i = 1\right\}. \tag{2}$$

Furthermore, $CH(\emptyset) = \emptyset$.

*Remark 15.* $CH(S)$ is the unique minimal (inclusion-wise) convex set containing $S$.

Thirdly, we define more geometrical objects.

**Definition 16.** A *polyhedron* is defined as an intersection of finitely many closed half-spaces.

**Definition 17.** A *polytope* is defined as a bounded polyhedron.

*Remark 18.* Polyhedra and polytopes are closed sets.

**Definition 19.** Let $P$ be a polyhedron. A point $\boldsymbol{e} \in P$ is an *extreme point* of $P$ if $\boldsymbol{e} \in \{\lambda\boldsymbol{p} + (1-\lambda)\,\boldsymbol{q}\}$, where $\boldsymbol{p}, \boldsymbol{q} \in P$ and $\lambda \in [0,1]$, implies that $\boldsymbol{e} = \boldsymbol{p} = \boldsymbol{q}$.

The set of all extreme points of $P$ is denoted by $ext(P)$.

**Definition 20.** Let $P$ be a polyhedron. A *supporting hyperplane* to $P$ is defined as a hyperplane that has a non-empty intersection with $P$ and such that all of $P$ lies to only one side of the hyperplane.

**Definition 21.** Let $P$ be a polyhedron. A *non-trivial face* of $P$ is the intersection of $P$ with a supporting hyperplane. The *trivial faces* of $P$ are $P$ and $\emptyset$.

**Definition 22.** Let $P$ be a polyhedron. A *polyhedral complex* $\Upsilon$ is defined as a finite non-empty collection of polyhedra such that

1. if $P \in \Upsilon$, then all faces of $P$ are also in $\Upsilon$;
2. if $P, Q \in \Upsilon$, then $P \bigcap Q$ is a face of both, $P$ and $Q$.

Similarly to a polyhedral complex, we may define a *polytopal complex*.

An example of a polyhedral complex would be $\Upsilon = \{F \mid F \text{ is a face of a polyhedron } P\}$.

**Theorem 23 (Minkowski-Weyl's theorem).** The following statements are equivalent for $P \subset \mathbb{R}^m$:

1. $P$ is a polyhedron;
2. there are finitely many real vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l$ and $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n$ in $\mathbb{R}^m$, where $l, n \in \mathbb{Z}_{>0}$, such that $P = CH(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l) + nonneg(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n)$, where $nonneg(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n)$ denotes all the non-negative combinations of those vectors.

**Corollary 24.** *A set is a polytope iff it is the convex hull of a finite set of points.*

Finally, we can address Voronoi diagrams.

**Definition 25.** Let $S = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subset \mathbb{R}^m$, where $1 \leq N < \infty$, and $\boldsymbol{x}_i \neq \boldsymbol{x}_j$ for $i \neq j$ and $i, j \in \{1, \ldots, N\}$. The *dominance region* of $\boldsymbol{x}_i$ over $\boldsymbol{x}_j$ is the closed halfspace given by

$$H\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) = \left\{\boldsymbol{x} \mid \|\boldsymbol{x} - \boldsymbol{x}_i\| \leq \left\|\boldsymbol{x} - \boldsymbol{x}_j\right\|\right\} \tag{3}$$

for $j \in \{1, \ldots, N\} \setminus \{i\}$. We call the region given by

$$V(\boldsymbol{x}_i) = \bigcap_{j \in \{1,\ldots,N\}\setminus\{i\}} H\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) \tag{4}$$

the *Voronoi polyhedron* associated with $\boldsymbol{x}_i$.

The set given by

$$\mathcal{V}(S) = \{V(\boldsymbol{x}_1), \ldots, V(\boldsymbol{x}_N)\} \tag{5}$$

is the *Voronoi diagram* generated by $S$. All of the faces of $V(\boldsymbol{x}_i)\,\forall i \in \{1, \ldots, N\}$ and $\mathcal{V}(S)$ form a polyhedral complex.

An $\boldsymbol{x}_i \in S$, where $i \in \{1, \ldots, N\}$, is called a *Voronoi point*.

An extreme point of $V(\boldsymbol{x}_i) \subset \mathcal{V}(S)$, where $i \in \{1, \ldots, N\}$, is called a *Voronoi vertex*.

We see that all of the faces of $V(\boldsymbol{x}_i)\,\forall i \in \{1, \ldots, N\}$ and $\mathcal{V}(S)$ form a polyhedral complex from the following. If we pick an element from this polyhedral complex, then all of its faces are in the polyhedral complex just from the way we defined it. If we pick two elements from this polyhedral complex, then their intersection can be the empty set, which is a face by definition and in the polyhedral complex by definition. Or it can be a non-empty set, which will always be a mutual face of some number of Voronoi polyhedra. The reason being is that any face is an intersection of some dominace regions. Therefore by Definition 22 the specified set is indeed a polyhedral complex.

*Remark 26.* An $(m-1)$-sphere centered at a Voronoi vertex, passing through a Voronoi point and not containing any other Voronoi points in its interior, will pass through at least another $m$ Voronoi points.

We can also define a so-called farthest-point Voronoi diagram. To distinguish it from the one defined above, we can term the diagram in Definition 25 to be the nearest-point Voronoi diagram.

**Definition 27.** Let $S = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subset \mathbb{R}^m$, where $1 \leq N < \infty$, and $\boldsymbol{x}_i \neq \boldsymbol{x}_j$ for $i \neq j$ and $i, j \in \{1, \ldots, N\}$. We call the region given by

$$V_{fp}(\boldsymbol{x}_i) = \bigcap_{j \in \{1,\ldots,N\} \setminus \{i\}} H\left(\boldsymbol{x}_j, \boldsymbol{x}_i\right) \tag{6}$$

the *farthest-point Voronoi polyhedron* associated with $\boldsymbol{x}_i$.

The set given by

$$\mathcal{V}_{fp}(S) = \left\{V_{fp}(\boldsymbol{x}_1), \ldots, V_{fp}(\boldsymbol{x}_N)\right\} \tag{7}$$

is the *farthest-point Voronoi diagram* generated by $S$. All of the faces of $V_{fp}(\boldsymbol{x}_i)\, \forall i \in \{1, \ldots, N\}$ and $\mathcal{V}_{fp}(S)$ form a polyhedral complex.

An extreme point of $V_{fp}(\boldsymbol{x}_i) \subset \mathcal{V}_{fp}(S)$, where $i \in \{1, \ldots, N\}$, is called a *farthest-point Voronoi vertex*.

*Remark 28.* An $(m-1)$-sphere centered at a farthest-point Voronoi vertex, passing through a Voronoi point, outside of which there are no Voronoi points, will pass through at least another $m$ Voronoi points. (I.e. such an $(m-1)$-sphere will contain all other Voronoi points in its interior, besides the ones it passes through.)

## 3. Computing Voronoi Diagrams (Brown's Algorithm)

Now a question arises: how can we compute Voronoi diagrams? Moreover, how can we compute them efficiently? There are a number of algorithms available nowadays for such purposes of different complexities, usually depending on the dimension worked with. In this paper we will study Brown's algorithm (Brown 1979). The main reason for focusing on this specific algorithm is that the algorithm can be easily generalized to $m$ dimensions. In fact, we will present it here for $m$ dimensions directly. Furthermore, we can compute nearest-point and farthest-point Voronoi diagrams simultaneously with it.

We begin by introducing the main tool of Brown's algorithm, namely inversion.

**Definition 29.** Let $\boldsymbol{c} \in \mathbb{R}^m$ and $r \in \mathbb{R}_{>0}$. The *inversion* with *center of inversion* $\boldsymbol{c}$ and *radius of inversion* $r$ is the following map:

$$\iota : \mathbb{R}^m \setminus \{\boldsymbol{c}\} \longrightarrow \mathbb{R}^m \setminus \{\boldsymbol{c}\} \tag{8}$$

$$\boldsymbol{p} \mapsto \boldsymbol{p}', \tag{9}$$

where $\boldsymbol{p} \in \mathbb{R}^m \setminus \{\boldsymbol{c}\}$, $\boldsymbol{p}' \in \overrightarrow{\boldsymbol{cp}}$ and $\|\boldsymbol{p} - \boldsymbol{c}\| \cdot \|\boldsymbol{p}' - \boldsymbol{c}\| = r^2$.

A note on notation, above $\overrightarrow{\boldsymbol{cp}}$ denotes a ray from $\boldsymbol{c}$ in the direction of $\boldsymbol{p}$.

*Remark 30.* $\iota$ is a bijection.

There are some key properties that need to be mentioned immediately and which will form the basis behind Brown's algorithm:

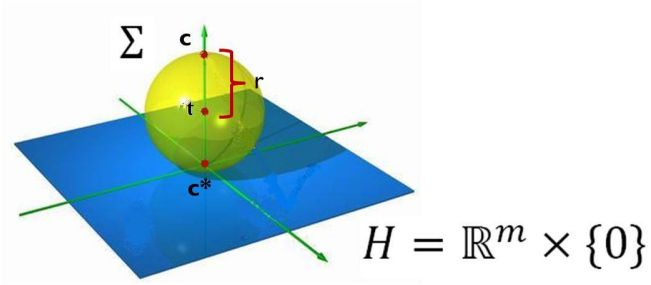


**Fig. 1.** Step II of Brown's algorithm: illustration of Σ.

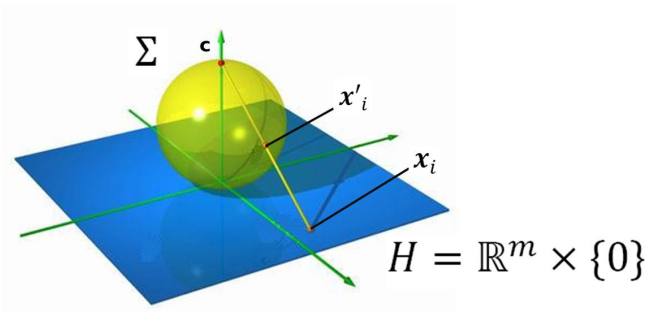


**Fig. 2.** Step II of Brown's algorithm: illustration of inversion.

1. Inversion is involutory, i.e.: $\iota(\iota(\boldsymbol{p})) = \boldsymbol{p}$.
2. A $k$-sphere that contains $\boldsymbol{c}$ maps onto a hyperplane that does not contain $\boldsymbol{c}$, where $k \in \{0, \ldots, m-1\}$.
3. A hyperplane that does not contain $\boldsymbol{c}$ maps onto a $k$-sphere that contains $\boldsymbol{c}$, where $k \in \{0, \ldots, m-1\}$.
4. A hyperplane that contains $\boldsymbol{c}$ is mapped onto itself.
5. A $k$-sphere that does not contain $\boldsymbol{c}$ maps onto another $k$-sphere that does not contain $\boldsymbol{c}$, where $k \in \{0, \ldots, m-1\}$.

With these tools in mind, we will now go through Brown's algorithm step by step.

I. Let $S = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subset \mathbb{R}^m$, where $m+1 \leq N < \infty$, and $\boldsymbol{x}_i \neq \boldsymbol{x}_j$ for $i \neq j$ and $i, j \in \{1, \ldots, N\}$.
Assume that no $m+1$ points of $S$ are contained in a hyperplane.

Observe that the minimal value of $N$ is $m+1$, because otherwise we could reduce to a lower dimension. The need for the assumption stated above will become clear at a later stage.

II. Identify $\mathbb{R}^m$ with $H = \mathbb{R}^m \times \{0\}$.
Pick a point $\boldsymbol{c} \in \mathbb{R}^{m+1}$ such that $\boldsymbol{c} \notin H$.
Let $\boldsymbol{c}^*$ be the orthogonal projection of $\boldsymbol{c}$ onto $H$.
Set

$$r = \frac{1}{2}\left\|\boldsymbol{c} - \boldsymbol{c}^*\right\|. \tag{10}$$

Set

$$\boldsymbol{t} = \frac{1}{2}\left(\boldsymbol{c} + \boldsymbol{c}^*\right). \tag{11}$$

Let $\Sigma$ be the $m$-sphere centered at $\boldsymbol{t}$ and with radius $r$. View Fig. 1 for an illustration.
Perform inversion on $S$ with the center of inversion $\boldsymbol{c}$ and the radius of inversion $r$ to attain the new set $S' = \{\boldsymbol{x'}_1, \ldots, \boldsymbol{x'}_N\} \subset \mathbb{R}^{m+1}$. View Fig. 2 for an illustration.

Essentially what we did in this step, upon picking such a point $\boldsymbol{c}$, is define an $m$-sphere $\Sigma$ tangent to $H$ at $\boldsymbol{c}^*$ and centered at the midpoint between $\boldsymbol{c}$ and $\boldsymbol{c}^*$. Then upon performing inversion we

mapped the points from the set that we started with onto new points belonging to $\Sigma$.

*Claim 31.* $S' \subset \Sigma$.

*Proof.* This follows immediately from the third property of inversion listed earlier in this section. ■

III. Construct $CH(S')$ with Preparata and Hong's algorithm (Preparata and Hong 1977).

*Claim 32.* $CH(S')$ is a polytope inscribed in $\Sigma$, hence its faces form a polytopal complex.

*Claim 33.* $S' = ext(CH(S'))$.

*Proof.* From Claim 32 and the definition of an extreme point of a polytope (Definition 19) it follows that $S' \supset ext(CH(S'))$.
Observe that for a finite set $T = \{\boldsymbol{x} \mid \boldsymbol{x} \in \Sigma\}$, no $\boldsymbol{x} \in T$ belongs to $CH(T \setminus \{\boldsymbol{x}\})$. This follows from the fact that any (strict) convex combination of elements of $T$ is in the interior of $\Sigma$, i.e. closer to the center by Pythagoras' theorem. Therefore it holds that $S' \subset ext(CH(S'))$.
Thus $S' = ext(CH(S'))$. ■

IV. Let $F$ be a face of $CH(S')$ of dimension $m$.
Perform inversion on the supporting hyperplane of $F$ with the center of inversion $\boldsymbol{c}$ and the radius of inversion $r$ to obtain an $m$-sphere $S^m$, which intersects with $H$ in an $(m-1)$-sphere $S^{m-1}$. Do this for all such faces $F$.

All of this follows directly from the inversion properties listed earlier in this section.

V. Associate with each $F$ the closed half-space $T$ such that $T$ is generated by the supporting hyperplane of $F$ and contains $CH(S')$.
If $T$ also contains $\boldsymbol{c}$, then the center of $S^{m-1}$ is a nearest-point Voronoi vertex $\boldsymbol{v}_F$. Otherwise, the center of $S^{m-1}$ is a farthest-point Voronoi vertex $\boldsymbol{v}_{fp_F}$.

It is not obvious why this holds. We will now prove why $\boldsymbol{v}_F$ is indeed a nearest-point Voronoi vertex.

*Proof.* We must prove that $\boldsymbol{v}_F$ satisfies the defining property of a nearest-point Voronoi vertex stated in Remark 26.

1. Consider some $\boldsymbol{v}_F$ generated by Brown's algorithm, where $F$ is a $m$-dimensional face of $CH(S')$.
In order for $F$ to be an $m$-dimensional face of $CH(S')$ in $\mathbb{R}^{m+1}$, it must be that:
$$|ext(F)| \geq m+1. \tag{12}$$
Since $F$ is a face of $CH(S')$, by Claim 33 it must be that:
$$ext(F) \subset S'. \tag{13}$$
We have established that $ext(F)$ consists of at least $m+1$ points of $S'$.
Taking inversion of both sides of equation (13) yields:
$$\iota(ext(F)) \subset S. \tag{14}$$
Then according to equation (12) and the fact that $\iota$ is a bijection (Remark 30):
$$|\iota(ext(F))| \geq m+1. \tag{15}$$
We have established that $\iota(ext(F))$ consists of at least $m+1$ points of $S$.
Since $\iota(ext(F))$ is contained in $H$ and in $\Sigma$, it must be that:
$$\iota(ext(F)) \subset S^{m-1}. \tag{16}$$
We have established that $S^{m-1}$ contains at least $m+1$ points of $S$.
2. Consider some $\boldsymbol{v}_F$ generated by Brown's algorithm, where $F$ is an $m$-dimensional face of $CH(S')$.
From the steps of the algorithm, we know that there is a supporting hyperplane to this $F$, such that $CH(S')$ and $\boldsymbol{c}$ lie to one side of it. Let us denote this supporting hyperplane $H_F$.
If $\boldsymbol{c} \in H_F$, then by the fourth property of $\iota$:
$$\iota(H_F) = H_F. \tag{17}$$
Since $ext(F) \subset F$ and $F \subset H_F$, by definition of a supporting hyperplane (Definition 20), it must be that:
$$ext(F) \subset H_F. \tag{18}$$
Taking inversion of both sides of the above inclusion:
$$\iota(ext(F)) \subset \iota(H_F) = H_F. \tag{19}$$
Since $\iota(ext(F))$ is simultaneously contained in $H$, it must be that:
$$\iota(ext(F)) \subset H \cap H_F. \tag{20}$$
Since $\iota(ext(F))$ consists of at least $m+1$ points of $S$, $H \cap H_F$ will contain at least $m+1$ points of $S$. Since $H \cap H_F$ is a hyperplane in $\mathbb{R}^m$, this contradicts the assumption of step I. ↯
(Now it is clear why that assumption had to be made.)
Therefore it must be that $\boldsymbol{c} \notin H_F$.
In step IV of the algorithm we inverted $H_F$ into $S^m$. We observe that $\boldsymbol{c} \in S^m$, by the third property of $\iota$. Since $S' \setminus ext(F)$ lie closer to $\boldsymbol{c}$ than the points on the other side of $H_F$, $S' \setminus ext(F)$ will be mapped further away from $\boldsymbol{c}$, i.e. outside of $S^m$. So in the interior of $S^m$ there will be no points of $S$, and therefore there will be no points of $S$ in the interior of $S^{m-1}$.

Both statements of Remark 26 are met. We observe here that in fact such a sphere is unique, because $F$ is an $m$-dimensional face and hence, prior to inversion, these Voronoi points that the sphere passes through were affinely independent. It must be that $\boldsymbol{v}_F$ is indeed a nearest-point Voronoi vertex. ■

Now we will prove why $\boldsymbol{v}_{fp_F}$ is indeed a farthest-point Voronoi vertex.

*Proof.* We must prove that $\boldsymbol{v}_{fp_F}$ satisfies the defining property of a farthest-point Voronoi vertex stated in Remark 28.

1. Exactly the same as for the nearest-point Voronoi vertex (see above).
2. Consider some $\boldsymbol{v}_{fp_F}$ generated by Brown's algorithm, where $F$ is a $m$-dimensional face of $CH(S')$.
From the steps of the algorithm, we know that there is a supporting hyperplane to this $F$, such that $CH(S')$ and $\boldsymbol{c}$ lie to opposite sides of it and $\boldsymbol{c}$ does not belong to the supporting hyperplane itself. Let us denote this supporting hyperplane $H_F$.
In step IV of the algorithm we inverted $H_F$ into $S^m$. We observe that $\boldsymbol{c} \in S^m$, by the third property of $\iota$. Since $S' \setminus ext(F)$ lie further from $\boldsymbol{c}$ than the points on the other side of $H_F$,

$S' \setminus ext(F)$ will be mapped closer to $\boldsymbol{c}$, i.e. inside of $S^m$. So in the interior of $S^m$ there will be all the remaining points of $S$, and therefore there will be all the remaining points of $S$ in the interior of $S^{m-1}$.

Since both statements of Remark 28 are met and the same observation can be made about the uniqueness of this sphere as in the nearest-point proof, it must be that $\boldsymbol{v}_{fp_F}$ is indeed a farthest-point Voronoi vertex. ■

VI. Let $E$ be a $(m-h)$-dimensional face of $CH(S')$ bounded by $l$ $(m-h+1)$-dimensional faces of $CH(S')$, $F_1, \ldots, F_l$, where $l \in \mathbb{Z}_{>0}$.
Then there will be a $h$-dimensional face generated by the Voronoi vertexes acquired from $F_1, \ldots, F_l$ in the Voronoi diagram. To determine to which Voronoi diagram these $h$-dimensional faces belong we have the following procedure:
   - If all the Voronoi vertexes acquired from $F_1, \ldots, F_l$ are nearest-point Voronoi vertexes, then the $h$-dimensional face belongs to the nearest-point Voronoi diagram.
   - If all the Voronoi vertexes acquired from $F_1, \ldots, F_l$ are farthest-point Voronoi vertexes, then the $h$-dimensional face belongs to the farthest-point Voronoi diagram.
   - If some of the Voronoi vertexes acquired from $F_1, \ldots, F_l$ are nearest-point Voronoi vertexes and some are farthest-point Voronoi vertexes, then the $h$-dimensional face is an unbounded face of the Voronoi diagram. It will begin at the nearest-point vertexes and go in the direction of the farthest-point vertexes, if it is to belong to the nearest-point Voronoi diagram and vice versa.

From the above step we combinatorially know which Voronoi vertexes are connected and how. It is quite easy now to get the Voronoi polytopes. Pick a Voronoi point, then its Voronoi polytope will be the convex hull of its nearest Voronoi vertexes. This is the same for nearest-point and farthest-point diagrams. However if the Voronoi point we picked gives rise to a Voronoi polyhedron, then the situation is no longer trivial. Recall the Minkowski-Weyl's theorem (Theorem 23). Compute the convex hull of the nearest nearest-point (farthest-point) Voronoi vertexes to the selected Voronoi point (depending on the type of diagram we are computing), call it $C$. Consider pairs of nearest-point and farthest-point vertexes. If the faces that we acquired them from were "adjacent", then call the direction vector between those two $\boldsymbol{d}_j$, where $j$ belongs to some index set. By "adjacent" we mean those that intersect in an $(m-1)$-dimensional face. Take all the non-negative combinations of all the vectors $\boldsymbol{d}_j$, call it $N$. Then the Voronoi polyhedron associated with the Voronoi point is $C + N$.

*Proof.* The initial combinatorial connections presented follow trivially by construction.
The reason why the convex hull of the nearest Voronoi vertexes to a certain Voronoi point is indeed the Voronoi polytope associated with it follows directly, since we took the convex hull of the nearest Voronoi vertexes and there will be no other Voronoi vertexes inside the convex hull by construction and by the definition of the convex hull (Definition 14). Thus there will indeed be all the point closest to that Voronoi point and no other inside that convex hull, and by definition it is then the Voronoi polytope (Definition 25).
As already mentioned earlier the fact that $C + N$ is a polyhedron follows from the Minkowski-Weyl theorem (Theorem 23) and the fact that it is the Voronoi polyhedron has the same reasoning as for the Voronoi polytope just stated. ■

There is still one final claim that remains to be verified.

*Claim 34.* Brown's algorithm generates all Voronoi vertexes.

*Proof.* Consider an arbitrary nearest-point (farthest-point) Voronoi vertex $\boldsymbol{v}_F$ ($\boldsymbol{v}_{fp_F}$). Then by Remark 26 (28) there will exist an $(m-1)$-sphere centered at $\boldsymbol{v}_F$ ($\boldsymbol{v}_{fp_F}$) such that it passes through at least $m+1$ Voronoi points and does not contain any (contains all) other Voronoi points in its interior.
By the same method as in the proof of step V, this will mean that in $\mathbb{R}^{m+1}$ there will be a hyperplane $H_F$ containing the inversions of the $m+1$ Voronoi points belonging to the $(m-1)$-sphere. Furthermore, the inversions of all other Voronoi points will be to one side of $H_F$, otherwise the $(m-1)$-sphere would not be empty (contain all other Voronoi points). That will also be the side (opposite to the one) containing $\boldsymbol{c}$.
Therefore we observe that such a $H_F$ would be the supporting hyperplane of a m-dimensional face of a polytope $P$, with $P$ and $\boldsymbol{c}$ lying to the same (opposite) side(s) of it.
Hence Brown's algorithm finds all nearest-point (farthest-point) Voronoi vertexes. ■

## 4. Complexity Analysis

Computational complexity is of importance to any algorithm. Let us address the time complexity of Brown's algorithm.

The complexity of step I is constant. Step II is of complexity $O(N)$, because inversion is performed on the $N$ points one starts with. The complexity of step III depends upon the algorithm used to calculate the convex hull. If we were to use Preparata and Hong's algorithm, then in two dimensions the complexity of the computation will be $O(Nlog(N))$. It may appear that step IV will be very complex, but actually we already know through which points the $(m-1)$-spheres will pass. This means that there is only a limited number of locations the centers of these $(m-1)$-spheres will be, therefore the complexity of this step is $O(N)$. The complexities of both, step V and step VI, are constant. Adding the complexity of all the six steps together, the complexity of Brown's algorithm turns out to be $O(Nlog(N))$ in two dimensions. In $m$ dimensions, the complexity will need to be determined after knowing the complexity of computing the convex hull. What is clear is the fact that step III appears to be the most complex and plays the determining role in complexity analysis.

## 5. Computing the Convex Hull (Preparata and Hong's Algorithm)

It was just established that computing the convex hull in $m$ dimensions is the key to quickly computing Voronoi diagrams by means of Brown's algorithm. There are a number of algorithms available for this purpose, however not all work in $m$ dimensions and not all are applicable to this case, because they may require additional knowledge about the points, which is not available here.

Let us quickly mention Preparata and Hong's algorithm (Preparata and Hong 1977) here, which was referenced in Brown's original paper (Brown 1979). The input for their algorithm is a set of $N$ points in $\mathbb{R}^m$, $A = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N\} \subset \mathbb{R}^m$, which are sorted with a preliminary step. Namely they are sorted according to their first coordinate and relabeled if necessary, so that $\boldsymbol{a}_{i_1} < \boldsymbol{a}_{j_1} \Leftrightarrow i < j$ for $i, j \in \{1, \ldots, N\}$. Then we can summarize the procedure in three key steps.

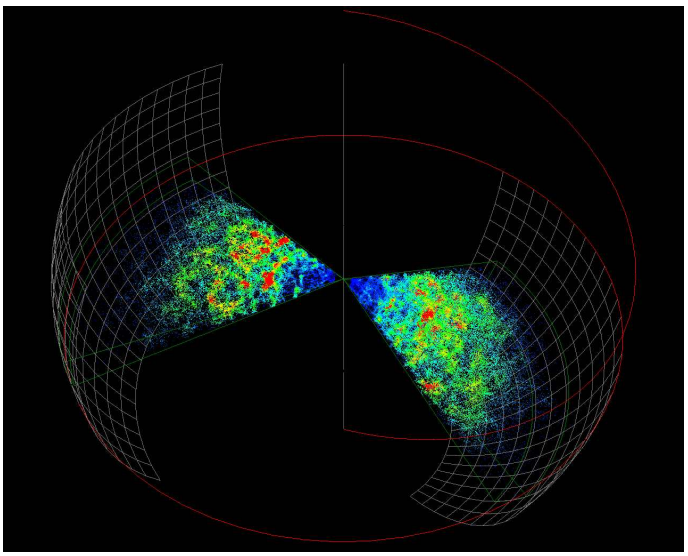

**Fig. 3.** 2dF Galaxy Redshift Survey (http://magnum.anu.edu.au/ TDFgg/ 2010).

1. Subdivide $A$ into $A_1 = \left\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{\frac{N}{2}}\right\}$ and $A_2 = \left\{\boldsymbol{a}_{\frac{N}{2}+1}, \ldots, \boldsymbol{a}_N\right\}$.
2. Apply this algorithm recursively to $A_1$ and $A_2$ to obtain $CH(A_1)$ and $CH(A_2)$.
3. Apply a merge algorithm to $CH(A_1)$ and $CH(A_2)$ to obtain $CH(A)$.

The output of the algorithm is in the form of a list of vertexes.

The authors of this algorithm have shown that the time complexity is $O(Nlog(N))$ in two and three dimensions, while for higher dimensions $O(Nlog(N))$ is the lower bound. Therefore for higher dimensions it may be that it is wiser to use a different algorithm. For example Chand and Kapur's algorithm (Chand and Kapur 1970), which is based on the so-called "gift wrapping" principle. Otherwise, one can apply the Quickhull algorithm of variable complexity depending on conditions (Barber, Dobkin and Huhdanpaa 1966).

## 6. The Cosmic Web

Let us now discuss the distribution of baryonic matter in the Universe on scales of a few up to more than a hundred megaparsecs. Generally it is assumed in cosmology that most prominent locators of baryonic matter are galaxies. It is known that galaxies are not randomly distributed in space. There are regions in space that have a very high galaxy density, but also regions where nearly no galaxies are seen at all. In fact, the distribution is visually similar to a cellular or a foam-like pattern, which has been termed as the *Cosmic Web*. For a vivid visualization, see Fig. 3. The Cosmic Web appears to have four distinctive morphological components, listed in order of decreasing galaxy density:

1. *clusters* - spherical gravitationally bound systems of galaxies,
2. *filaments* - thread-like gravitationally bound systems of galaxies,
3. *walls* - two dimensional sheets of gravitationally bound galaxies,
4. *voids* - convex under-dense regions.

It is possible to make a more quantitative distinction between the above four constituents by means of the deformation eigenvalues of the Zel'dovich approximation (Platen 2009), but for our purposes the above qualitative descriptions are sufficient. It is important to note that the term void, which is an abbreviation for a "region devoid of galaxies", is not that trivial to define. For example, it is not clear how to trace the borders of a void. Currently, there is no consensus on a definition; however for this paper the above mentioned one is quite satisfactory. The richest clusters contain many thousands of galaxies within a relatively small volume with a diameter of only a few megaparsecs and can be thought of as cosmic nodes (Weygaert 2009). Meanwhile, the approximate diameters of voids lie in the range of $20-50h^{-1}Mpc$ and can be thought of as cosmic depressions (Weygaert 2009). Furthermore, hierarchy in the structures of the Cosmic Web has been observed, suggesting that large components have formed as a result of small structures merging.

The Cosmic Web has two main characteristics:

1. it is highly anisotropic;
2. clusters weave the Cosmic Web.

The first point is intuitively clear from the above description. The second point is deduced from the fact that clusters are the most massive components of the Cosmic Web and the theory about gravity being the dominant interaction. It is a possibility to study the Cosmic Web from the point of view of clusters; however it also makes sense to start with locating the voids, because they are the largest in size and they are very striking features. The *Watershed Void Finder (WVF)* developed by Erwin Platen, Rien van de Weygaert and Bernard J.T. Jones in 2007 appears to be one of the best void finders up to date (Platen 2007 and Platen 2009).

The WVF is a multistep process and in this paper the focus will be on the very first, yet essential, step. The input of the WVF, in this cosmological context, will either be a sample of galaxy positions obtained from redshift surveys or the positions of a large number of particles produced by $N$-body simulations of cosmic structure formation. The first step of the WVF is then to construct a continuous density field from the given discrete data by using the *Delaunay Tessellation Field Estimator (DTFE)*.

## 7. Further Definitions and Concepts

Before we begin discussing the DTFE, we must understand its tools. Delaunay triangulations are the key to its procedure. From now on we will focus only on nearest-point Voronoi diagrams and will omit "nearest-point" when speaking about them.

**Definition 35.** The subdivision of a space into simplexes is referred to as a *triangulation* (a *simplex* is a generalization of the notion of a two dimensional triangle and a three dimensional tetrahedron to higher dimensions).

**Definition 36.** Let $S = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subset \mathbb{R}^m$, where $1 \leq N < \infty$, and $\boldsymbol{x}_i \neq \boldsymbol{x}_j$ for $i \neq j$ and $i, j \in \{1, \ldots, N\}$. The *Delaunay triangulation* of $S$, denoted by $DT(S)$, is the triangulation of $S$ such that no point in $S$ is in the interior of the circumsphere of any simplex in $DT(S)$.

An $\boldsymbol{x}_i \in S$, where $i \in \{1, \ldots, N\}$, is called a *Delaunay point*.

*Remark 37.* $DT(S)$ always exists and is unique.

There are a number of interesting connections between Voronoi diagrams and Delaunay triangulations.

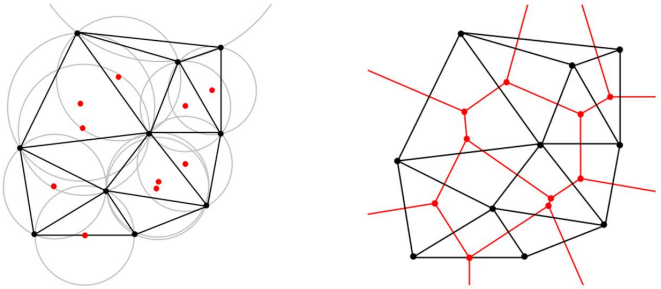

**Fig. 4.** Voronoi diagram and Delaunay triangulation relation: The left figure is the Delaunay triangulation (black lines) of the set of black points and the circumcircles (grey circles) of the triangles belonging to it. The centers of these circumcircles are the red points. They in fact correspond to the set of Voronoi points. They are also shown in the right figure. If we connect these red points we will acquire the Voronoi diagram. For a detailed discussion on the connection rules refer to Okabe 2000.

*Remark 38.* The set of Voronoi points is equal to the set of Delaunay points.

*Remark 39.* The set of Voronoi vertexes is equal to the set of the centers of the circumspheres of the simplexes in $DT(S)$. In fact this can be used to construct one from the other, see Fig. 4 for the construction steps and illustration in $\mathbb{R}^2$.

*Remark 40.* In literature it is often loosely spoken about duality between Voronoi diagrams and Delaunay triangulations. In fact this is not a trivial matter. One can define duality between the two in terms of Legendre dual functions (Chynoweth 1996). Yet another way to define duality between them is via the duality of polytopes (see Grünbaum 1967 for the definition). However in $\mathbb{R}^2$, one can even view the duality in terms of the non-unique dual planar graphs (see Grimaldi 2004).

*Remark 41.* $\mathcal{V}(S)$ is a tessellation of $\mathbb{R}^m$, while $DT(S)$ is a tessellation of $CH(S)$.

Observe that one can also define the farthest-point Delaunay triangulations too and relate to them the farthest-point Voronoi diagrams, but it will not be useful for our purposes (Okabe 2000).

## 8. The DTFE

From now onwards we will discuss Voronoi diagrams and Delaunay triangulations as they are applied in the DTFE for cosmological purposes, so we will limit ourselves to $\mathbb{R}^3$, at most, and nearest-point tessellations (as mentioned earlier).

The DTFE is essentially a three step process. We will present the basics here, for further details refer to Schaap 2007. As mentioned earlier, the input into the DTFE is a discrete set of points in $\mathbb{R}^3$, let us call it $S$. From cosmological reasons, it is logical to suppose that $S$ consists of finitely many distinct points, i.e. $S = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subset \mathbb{R}^3$, where $4 \leq N < \infty$, and $\boldsymbol{x}_i \neq \boldsymbol{x}_j$ for $i \neq j$ and $i, j \in \{1, \ldots, N\}$. We require to have at least 4 points, for the reasons explained in the section on Brown's algorithm. Furthermore, for reasons stated in that section, we also assume that $S$ is not contained in a hyperplane and that no 4 points of $S$ are contained in a hyperplane. Another assumption that we cannot avoid when applying the DTFE to cosmological data is that $S$ is an unbiased sample (see Definition 42 for clarity) of the underlying continuous density field.

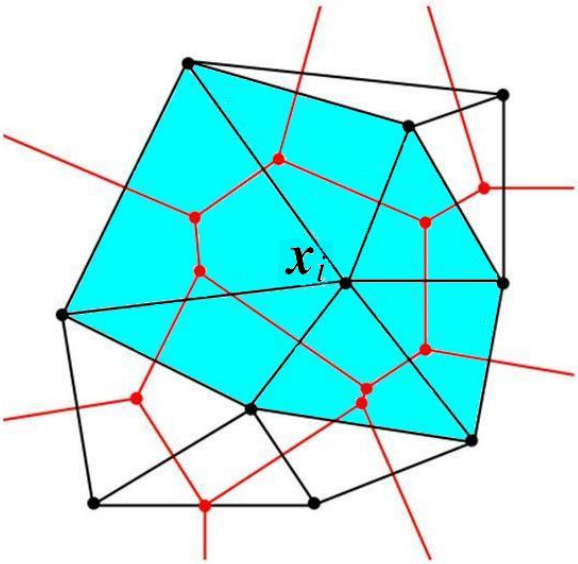


**Fig. 5.** Contiguous cell associated with $\boldsymbol{x}_i$: Black lines form the Delaunay triangulation and the red lines form the Voronoid diagram.

**Definition 42.** An *unbiased sample* is defined as a sample that was obtained by a method free from bias.

Now let us go through the DTFE step by step.

I. Construct $DT(S)$ using some algorithm.

The result of this would be that $CH(S)$ is tessellated into tetrahedra, which is deduced from the definition of a Delaunay triangulation (Definition 36). Furthermore observe from the construction that the set of all the extreme points of all the tetrahedra just obtained is $S$ itself. Before we proceed an additional term needs to be introduced.

**Definition 43.** The *contiguous cell* of a Delaunay point is the union of all the simplexes that have that Delaunay point as one of their extreme points (see Fig. 5 for an example in $\mathbb{R}^2$).

II. Estimate the density field at each point of $S$ by

$$\widehat{\rho}(\boldsymbol{x}_i) = \frac{4}{Vol(W_i)}, \tag{21}$$

where $W_i$ is the contiguous cell associated with $\boldsymbol{x}_i$ and $Vol(W_i)$ is the volume of it.

Note that we could also have included a relative measure of influence of $\boldsymbol{x}_i$ on the other points in the numerator of the equation, but since in the cosmological applications the mass of data points is always assumed to be the same, for now, we have omitted it. The normalization factor of 4 is necessary, because every tetrahedron has four vertexes. (If we were working in $\mathbb{R}^m$, this normalization factor would have to be $m + 1$.) Furthermore, we assume that the density field is constant inside each tetrahedron.

III. Interpolate the density field by means of
- linear interpolation, or
- spline interpolation, or
- Natural Neighbor Interpolation, or
- other.

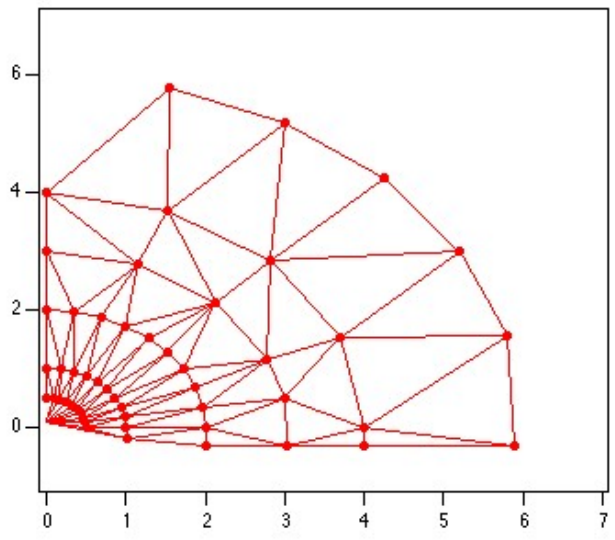


**Fig. 6.** Sensitivity of the Delaunay triangulation: Solid red lines form the Delaunay triangulation of the set of red points. Obseve how the volume of the triangles gets smaller as the local point density increases towards the origin. Furthermore, observe how the shape of the triangles, hence their volume too, changes as the local geometry of the point distribution becomes more regular towards the origin.

As a result of this step the density field is no longer discontinuous at the boundaries of the tetrahedra.

Finally one attains the desired output, namely a continuous map $\rho : \mathbb{R}^3 \longrightarrow \mathbb{R}$ representing the underlying continuous density field of the given discrete point distribution.

It is not immediately clear why the estimate in step II of the DTFE should make any sense. In order to understand it we must observe the following qualitative property of the Delaunay triangulation.

*Remark 44.* The Delaunay triangulation is very sensitive to the local point density and is very sensitive to the local geometry of the point distribution. This is quite clear from Fig. 6.

Another point of interest here is the type of boundary conditions imposed, since the set of points that we are given is always finite. It has been decided that vacuum boundary conditions would correspond to the optimal model. In other words, there are no other data points outside of the convex hull of the given discrete set.

The claim by the founders of the DTFE is that it guarantees a continuous density field which retains the morphological character of the underlying point distribution and that it is capable of delineating three fundamental characteristics of the megaparsec cosmic matter distribution, namely anisotropy, the presence and shape of voids and even the hierarchy of substructures. Indeed the results presented in Schaap 2000, 2007, Platen 2007, 2009 and Weygaert 2009 suggest that this might be the case.

## 9. Evaluating the Quality of the DTFE

A number of methods have been applied to evaluate the similarities between the discrete data that one starts with and the continuous field one acquires after applying the DTFE. The results have been quite satisfactory (Schaap 2000, 2007, Platen 2007, 2009 and Weygaert 2009). As well as visually, the results seem pleasant. However a control test has not been conducted. This is one of the research questions investigated in this paper.

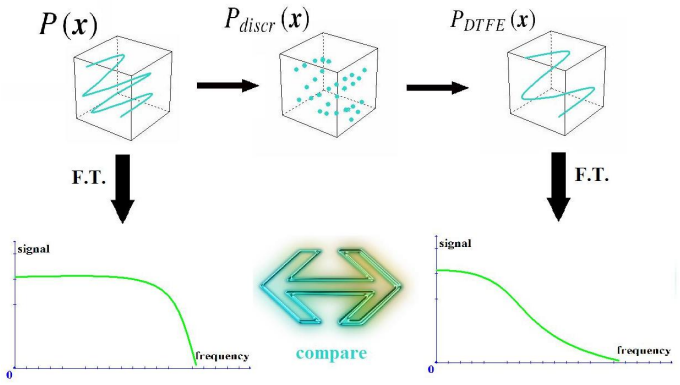


**Fig. 7.** The method illustrated.

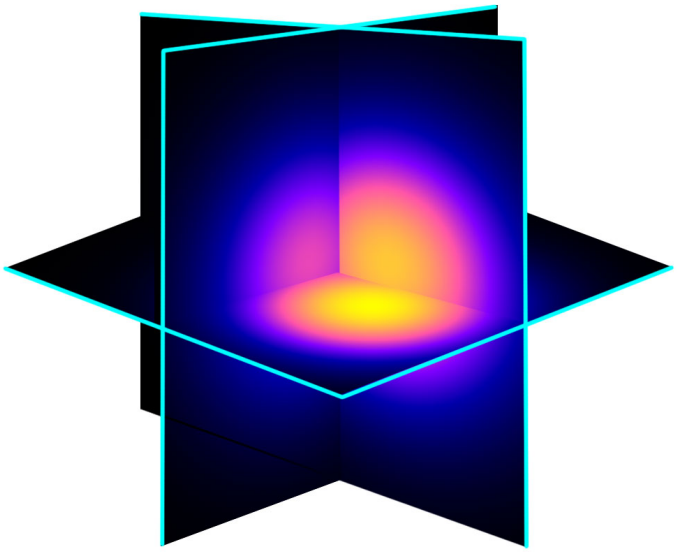

**Fig. 8.** The Gaussian cloud: three slices together.

A control test is always a good way to see whether the method employed is working well. In this paper we did exactly that with the DTFE by comparing a known continuous distribution with its DTFE reconstruction. Let us present the method step by step and also illustrate the steps in Fig. 7.

1. Take a known continuous point distribution $P(\boldsymbol{x})$.
2. Generate a discrete point distribution $P_{discr}(\boldsymbol{x})$ from $P(\boldsymbol{x})$.
3. Make a DTFE reconstruction $P_{DTFE}(\boldsymbol{x})$ from $P_{discr}(\boldsymbol{x})$.
4. Compute the Fourier spectrum of $P(\boldsymbol{x})$, denoted $FT(P(\boldsymbol{x}))$.
5. Compute the Fourier spectrum of $P_{DTFE}(\boldsymbol{x})$ denoted $FT(P_{DTFE}(\boldsymbol{x}))$.
6. Compare the Fourier spectra $FT(P(\boldsymbol{x}))$ and $FT(P_{DTFE}(\boldsymbol{x}))$.

Fourier transforms give us insight into the frequencies that are present in our continuous distributions. By comparing $FT(P(\boldsymbol{x}))$ and $FT(P_{DTFE}(\boldsymbol{x}))$ we will be able to see how specific frequencies are reconstructed by the DTFE.

## 10. Results

We have decided to take the Gaussian cloud as our known continuous distribution, because of its easy, single-peak structure and smoothness. The Gaussian cloud follows the normal distribution in all directions from its center. Using Vincent Icke's code

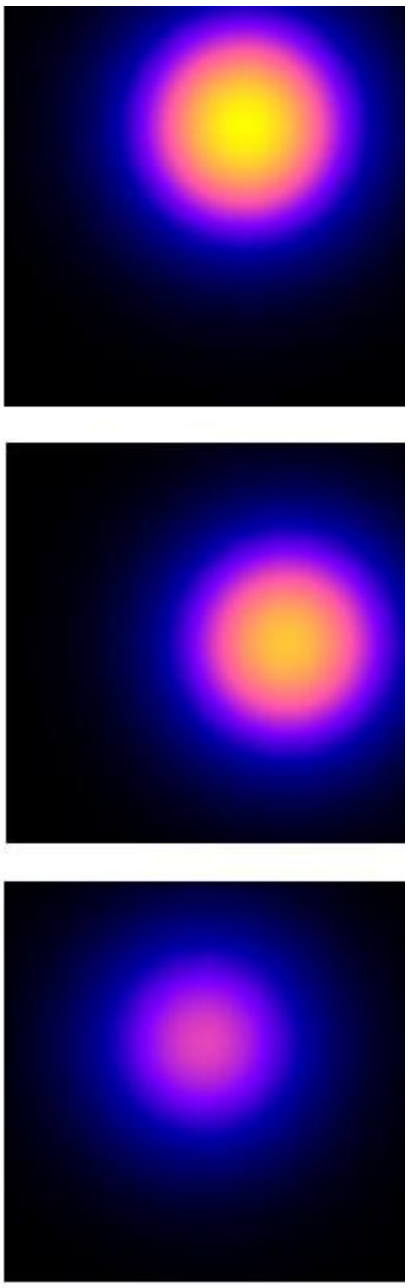

**Fig. 9.** The Gaussian cloud: from top to bottom $x = y = 0.5$ plane, $y = z = 0.5$ plane and $x = z = 0.5$ plane.

we were able to generate such a Gaussian cloud in a unit cube (a cube that has $x, y, z \in [0, 1]$). The distribution is visualized in Fig. 9 by means of slices. The relation of the slices to each other is displayed in Fig. 8. The unit cube was set to have a grid of $200^3$ cells. The center of the Gaussian cloud was set to be $(x, y, z) = (0.3, 0.4, 0.5)$. The peak of the Gaussian cloud is normalized to unity in the code.

Now we had to generate a discrete point distribution from this Gaussian cloud. Using Chael Kruip and Vincent Icke's codes we were able to do that by means of a specific sampling function (Kruip 2010). The first test was asked to sample the distribution randomly on the same grid and cube. An important factor here was the number of sampling points. In Fig. 10 we see a three dimensional image of the discrete point distribution obtained from the continuous Gaussian cloud by sampling it with 5000 points, in Fig. 11 - with 20000 points and in Fig. 12 - with 80000 points.

Finally, we have applied Erwin Platen's code to generate the DTFE reconstructions of the three discrete point distribution, see Fig. 13, 14 and 15 respectively. The same unit cube was used, only the number of the grid cells is now $128^3$.

As we see from these three DTFE reconstructions, the reconstructions are not anywhere nearly as smooth as the original data. We observe a strong piece-wise construction of an image. The general shape and the rough location of the Gaussian cloud are indeed represented, however the smoothness is gone and the single-peak is absent. Instead the peaks are now multiple. As we compare the three reconstructions, we see that the piece-wise structure increases with the number of points, but on the other hand the reconstruction appears to be more spherical. Furthermore, the peaks get lower as the number of points increases.

So it seems that the DTFE did not do a good job, however there may be another reason for our observations. Recall that the known continuous distribution was sampled with a random process and as we know any random process is accompanied by Poisson noise. It may very well be that the DTFE picks up the Poisson noise present in the discrete point distribution, which is not present in our starting continuous Gaussian cloud and it may be the reason our DTFE reconstructions are no good. Furthermore, Poisson noise increases with the number of points we sample with, which corresponds with the observation of the reconstructions worsening in some sense with the number of points too.

At this point we decided to not proceed with the Fourier spectra analysis, because the results could be already observed visually. An observation such as the one we just made calls for another test. What we could do is try a different sampling method, rather than random, we could do a more regular sampling. The initial random sampling worked as follows:

1. Pick a grid cell.
2. Set a random constant $k \in [0, 1]$.
3. Determine the density value in the grid cell. If this value is higher than $k$, then place a point.

This cycle keeps repeating itself until the desired number of sample points is acquired. The new regular sampling has the following design:

1. Start at the very first grid cell.
2. Set a random constant $k \in [0, 1]$.
3. Determine the density value in the very first grid cell. If this value is higher than $k$, then place a point.
4. Move on to the next grid cell, in order.

This cycle keeps going over the whole grid multiple times until the desired number of points is attained. In fact there will still be Poisson noise within each cell. However towards the center of the Gaussian cloud this Poisson noise will be suppressed by the regularity of the grid, and so we will only feel this Poisson noise far away from the center of the Gaussian cloud.

Using Vincent Icke's code we were able to generate such a sampling procedure. In Fig. 16 we see a three dimensional image of the discrete point distribution obtained from the continuous Gaussian cloud by this regular sampling with 5000 points, in Fig. 17 - with 20000 points and in Fig. 18 - with 80000 points. The DTFE reconstructions are following in Fig. 19, 20 and 21, respectively. Furthermore, we do indeed observe the crystalline structure in the point distribution as a result of the regularity of the grid. For comparison, see the zoomed in (almost the same zoom in all) images of the center of the Gaussian cloud with the two sampling method for all three point numbers in Fig. 22 - 27.

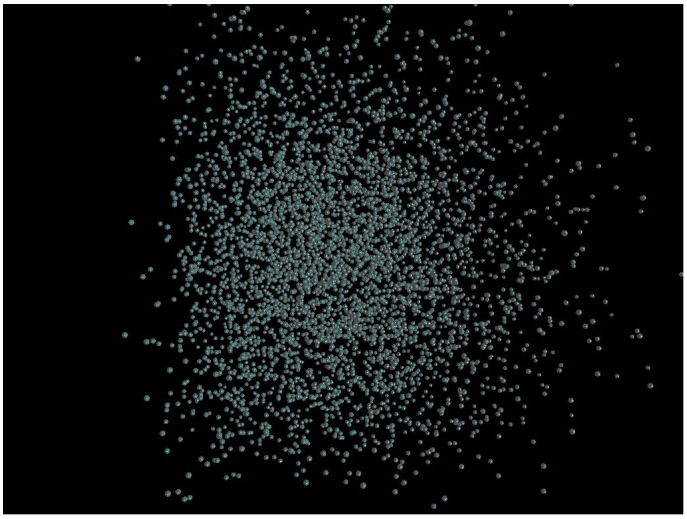

**Fig. 10.** The Gaussian cloud sampled randomly by 5000 points.

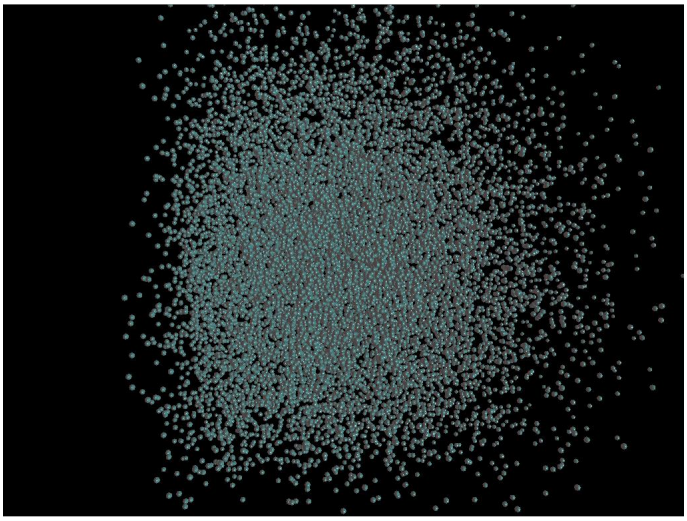

**Fig. 11.** The Gaussian cloud sampled randomly by 20000 points.

When we look at these new DTFE reconstructions we observe that the results are actually not much better and we still observe the same dependence on the number of points sampled with. There another possible explanation for this. Perhaps when the code runs over the grid cells multiple times, it actually happens quite often that multiple points need to be placed in a grid cell. The code is designed in such a way that it displaces the points from the center of the grid cell by a very small random amount. This may result in recurrent conglomeration of points close to centers of the grid cells. In turn the DTFE will pick that up and misinterpret it as a peak.

## 11. Outlook

At this stage the Poisson noise, if that is indeed the reason for the observed fluctuations in the reconstruction, has not been solved. It is very likely that it is the reason, but we were not able to conclude that from these results. Therefore, other sampling methods could definitely be investigated. It seems that the regular grid has not been as successful as hoped, because it appears that it may have an additional flaw of its own. A potential solution could lie with the centroidal distributon.

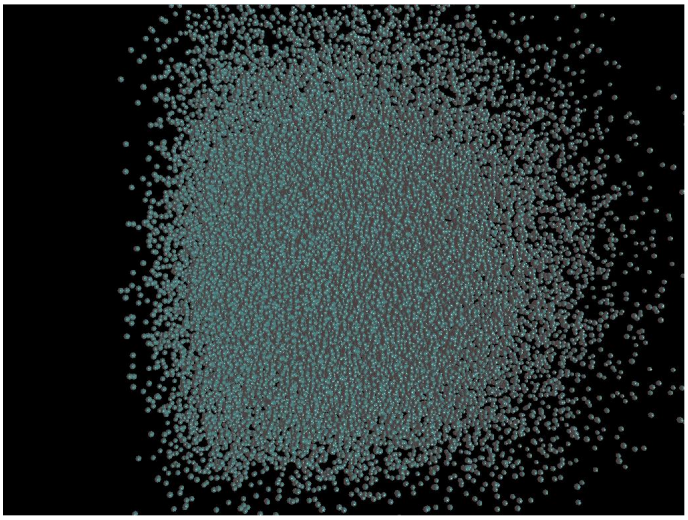

**Fig. 12.** The Gaussian cloud sampled randomly by 80000 points.

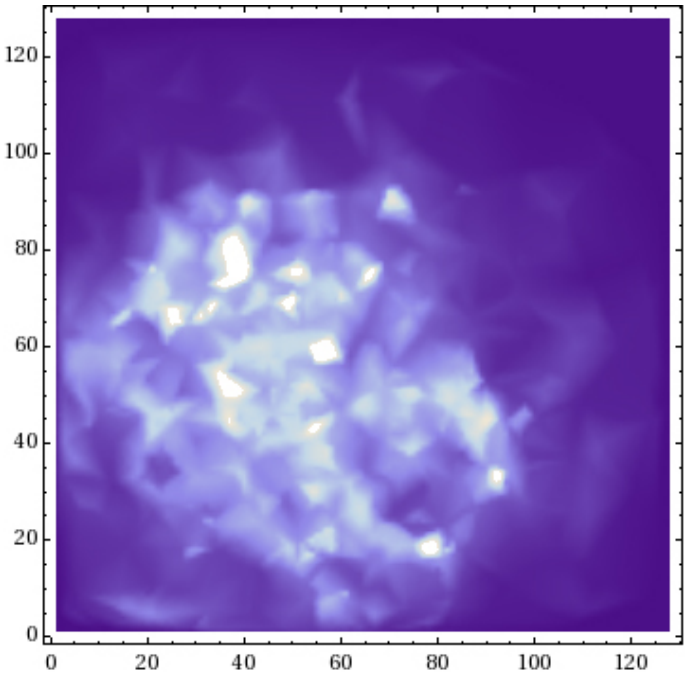


**Fig. 13.** The DTFE reconstruction of the Gaussian cloud sampled randomly by 5000 points.

Furthermore, here we have taken a Gaussian cloud - a structure that does not require high resolution to view. However if this is applied to cosmological studies a resolution research needs to be carried out, because one may not be able to view certain features if the grid cells are too large. Furthermore the size of the grid cells limit the part of the Fourier spectrum that can be spoke of meaningfully. Therefore if the reconstructions can be made visually satisfying and the analysis via Fourier spectra was to be carried out, then the resolution study is essential. Also it is important to note that it would be more representative of the Cosmic Web to analyze wave superposition rather than "artificial" distributions such as the Gaussian cloud.

Finally, there are a number of other open interesting questions here. Perhaps the DTFE can be improved by not assuming constant density in each tetrahedron.

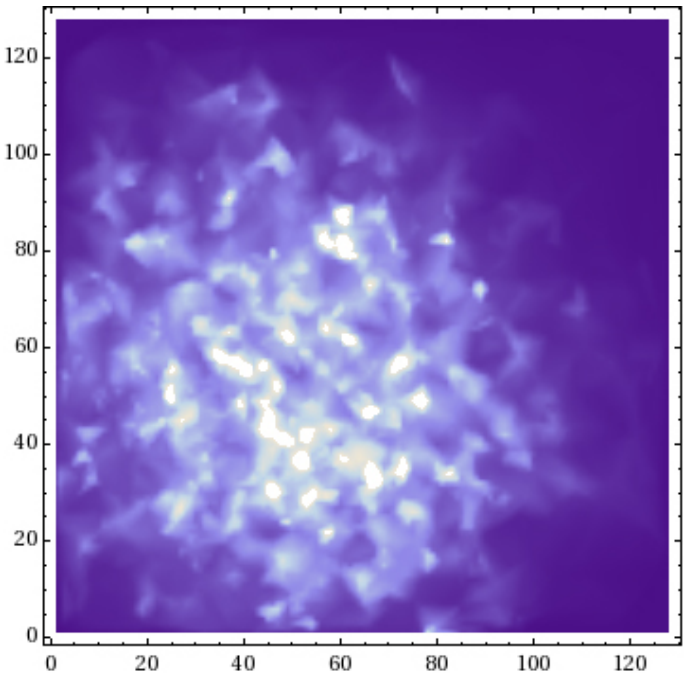


**Fig. 14.** The DTFE reconstruction of the Gaussian cloud sampled randomly by 20000 points.

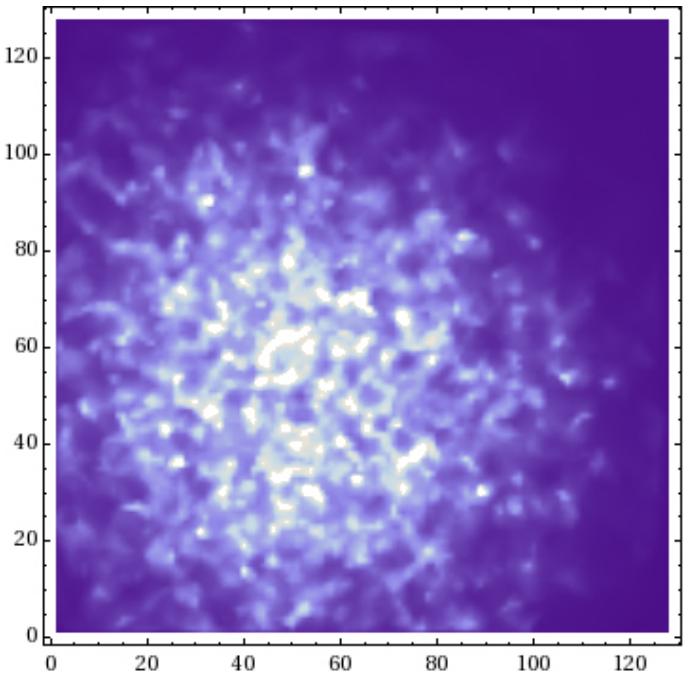


**Fig. 15.** The DTFE reconstruction of the Gaussian cloud sampled randomly by 80000 points.

## 12. Conclusions

In this paper we acquainted ourselves with Voronoi diagrams, Delaunay triangulations and their relationship to each other. We studied Brown's algorithm, which allows to compute Voronoi diagrams in any dimension via its main tool of inversion. We then looked at how Delaunay triangulations were applied in astronomy to analyze the Cosmic Web in the WVF, whose first step was to use them in the DTFE. We studied how the DTFE transforms a discrete point distribution into a continuous density field.

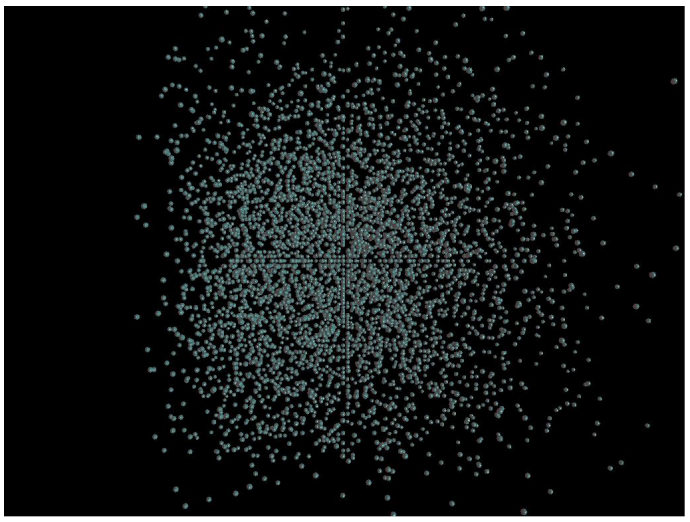

**Fig. 16.** The Gaussian cloud sampled regularly by 5000 points.

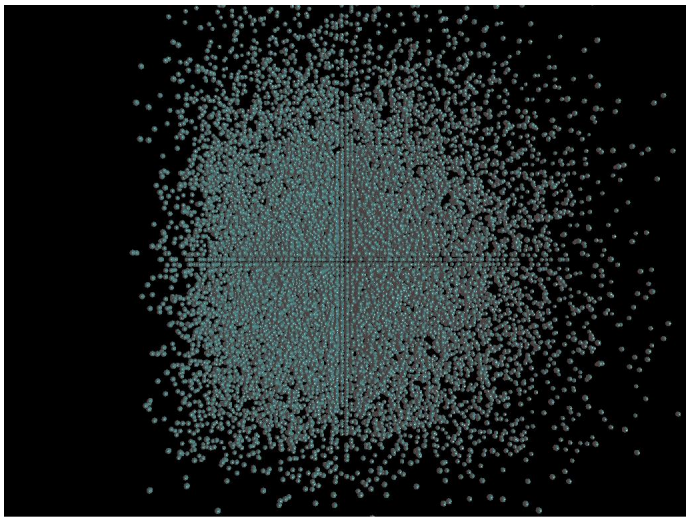

**Fig. 17.** The Gaussian cloud sampled regularly by 20000 points.

With high hopes for the method's success we went and tried to compare a known continuous distribution with its DTFE reconstruction. The results were quite surprising. We discussed the possible reasons for that and suggested further tests that can be done. Taking into account the importance of Voronoi diagrams, Delaunay triangulations and the popular desire of going from a discrete distribution to a continuous distribution, not only in astronomy, but in numerous other field too, it is of great interest to science that further research is to be carried out in this field.

*Acknowledgements.* I would like to thank both my supervisors, Prof. Dr. Vincent Icke and Dr. Dion Gijswijt, for all their help, time and advice with this project. Also I would like to thank Chael Kruip for all his input.

## References

Barber, C. Bradford and David P. Dobkin, Hannu Huhdanpaa, 1966, The Quickhull Algorithm for Convex Hulls, ACM Transactions on Mathematical Software, Volume 22, Number 4, 469-483

Berg, Mark de and Otfried Cheong, Marc van Kreveld, Mark Overmans, 2008, Computational Geometry: Algorithms and Applications, 3rd ed., Springer

Boissonnat, Jean-Daniel and Mariette Yvinec, 1997, Algorithmic geometry, Chapter 17

Brown, Kevin Q., 1979, Voronoi Diagrams from Convex Hulls, Information Processing Letters, Volume 9, number 5, 223-228

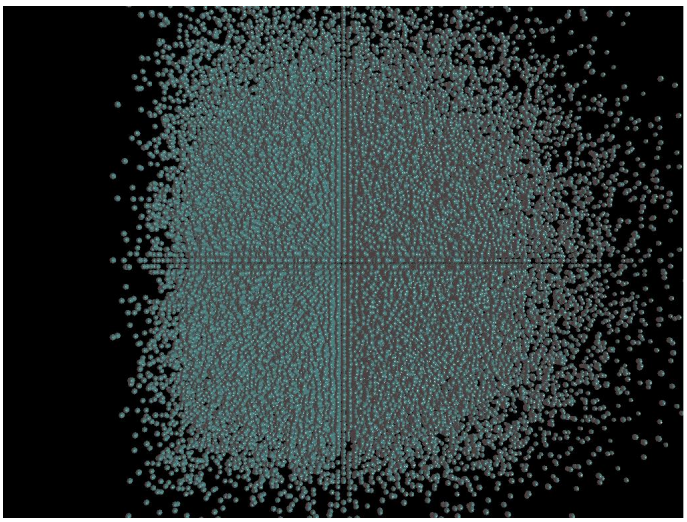

**Fig. 18.** The Gaussian cloud sampled regularly by 80000 points.

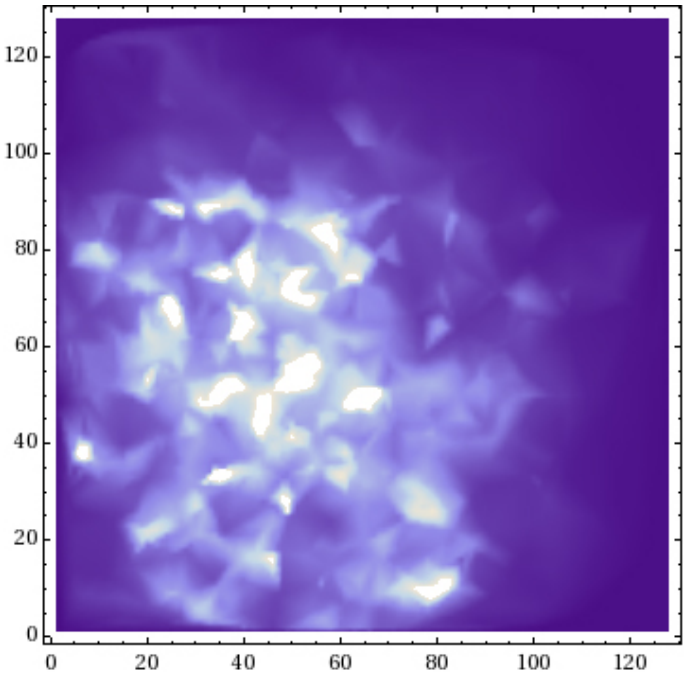


**Fig. 19.** The DTFE reconstruction of the Gaussian cloud sampled regularly by 5000 points.

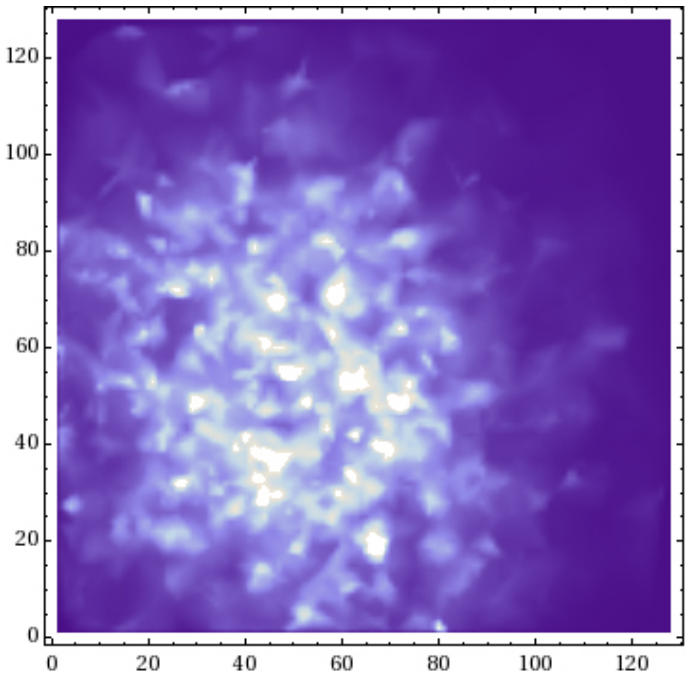


**Fig. 20.** The DTFE reconstruction of the Gaussian cloud sampled regularly by 20000 points.

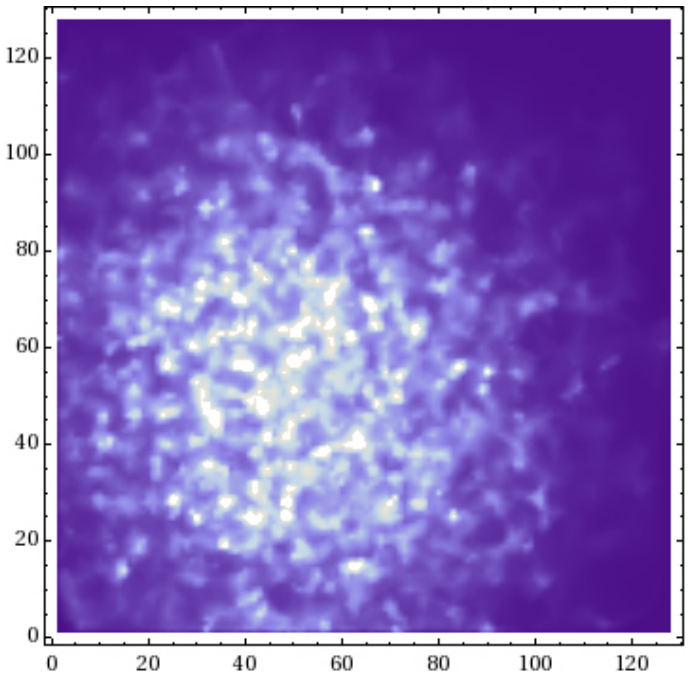


**Fig. 21.** The DTFE reconstruction of the Gaussian cloud sampled regularly by 80000 points.

Chand, Donald R. and Sham S. Kapur, 1970, Journal of the Association for Computing Machinery, Volume 17, Number 1, 78-86
Chynoweth, Simon, 1996, Z. Angew. Math. Mech., 76, 339-342
Davies, G.A. and D.J. Bell, 1996, Z. Angew. Math. Mech., 76, 343-346
Diestel, Reinhard, 2005, Graph Theory
Fukuda, Komei, 1998, Frequently Asked Questions in Polyhedral Computation
Grimaldi, Ralph P., 2004, Discrete and Combinatorial Mathematics: An Applied Introduction, Fifth Edition, Pearson Education, Inc.
Grünbaum, Branko, 1967, Convex Polytopes, John Wiley & Sons, Ltd.
Harary, F., 1973, Graph Theory
Kedlaya, Kiran S., 2006, Geometry Unbound
Kruip, C.J.H. and J.-P.Paardekooper, B.J.F. Clauwens, V. Icke, 2010, Mathematical Properties of the SimpleX Algorithm, Astronomy & Astrophysics
Okabe, Atsuyuki and Barry Boots, Kokichi Sugihara, Sung Nok Chiu, 2000, Spatial Tessellations: Concepts and Applications of Voronoi Diagrams, 2nd ed., John Wiley & Sons, ltd.
Platen, Erwin, 2009, A Void Perspective of the Cosmic Web, Ipskamp Drukkers
Platen, Erwin and Rien van de Weygaert, Bernard J.T. Jones, 2007, A Cosmic Watershed: the WVF Void Detection Technique, Mon. Not. R. Astron. Soc. 380, 551-570
Preparata, F.P. and S.J. Hong, 1977, Convex Hulls of Finite Sets of Points in Two and Three Dimensions, Communications of the ACM, Volume 20, Number 2, 87-93
Rekha, Thomas R., 2004, Lectures in Geometric Combinatorics
Schaap, Willem Egbert, 2007, DTFE: The Delaunay Tesselation Field Estimator
Schaap, W.E. and R. van de Weygaert, 2000, Continuous Fields and Discrete Samples: Reconstruction Through Delaunay Tessellations, Astronomy and Astrophysics, 363, L29-L32
The 2dF Galaxy Redshift Survey, http://magnum.anu.edu.au/ TDFgg/, as seen on 23 June, 2010

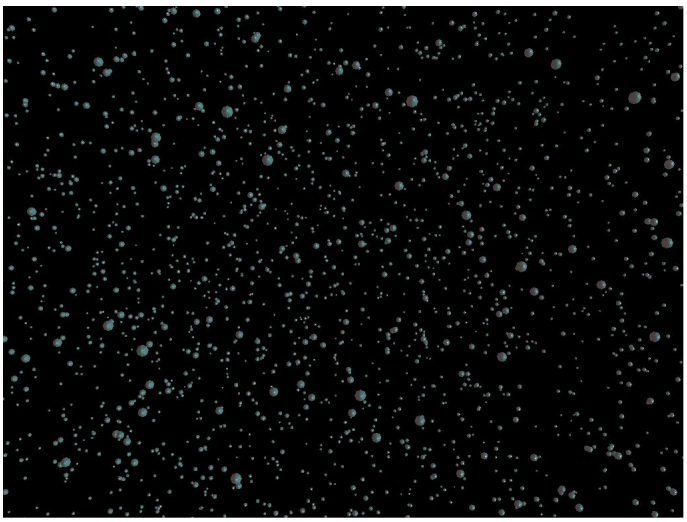

**Fig. 22.** The Gaussian cloud sampled randomly by 5000 points: zooming in on the center.

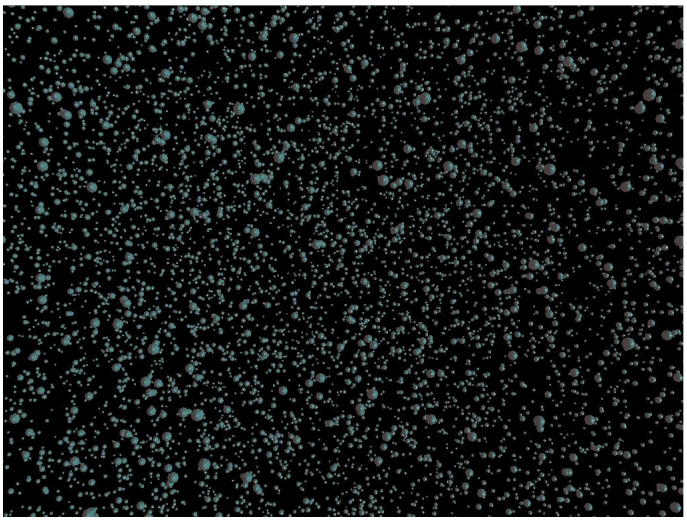

**Fig. 23.** The Gaussian cloud sampled randomly by 20000 points: zooming in on the center.

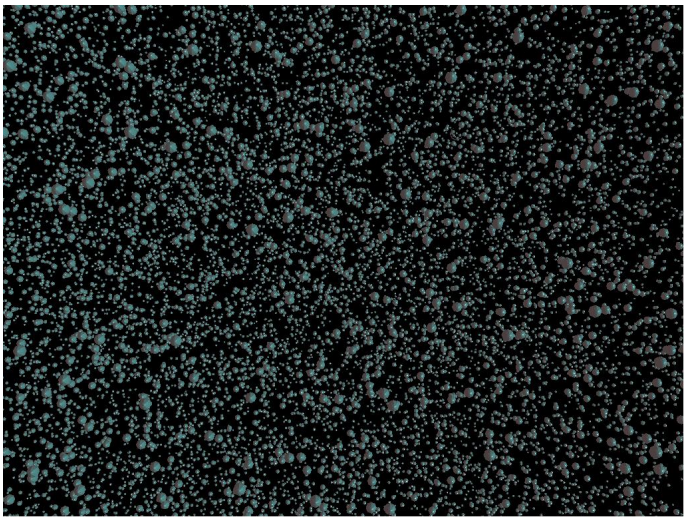

**Fig. 24.** The Gaussian cloud sampled randomly by 80000 points: zooming in on the center.

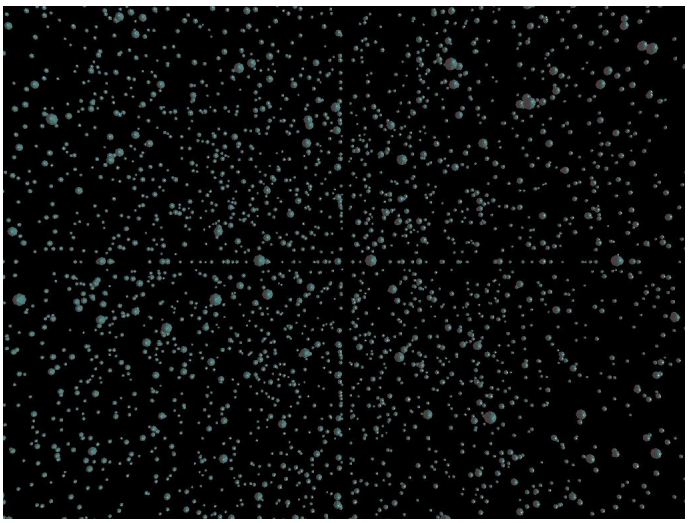

**Fig. 25.** The Gaussian cloud sampled regularly by 5000 points: zooming in on the center.

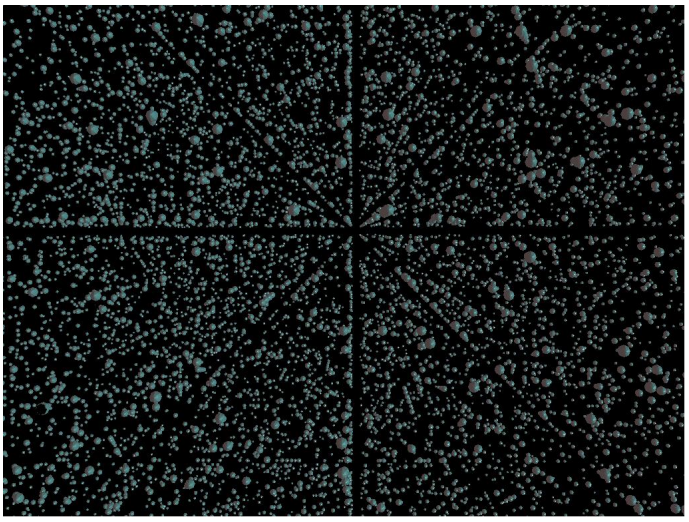

**Fig. 26.** The Gaussian cloud sampled regularly by 20000 points: zooming in on the center.

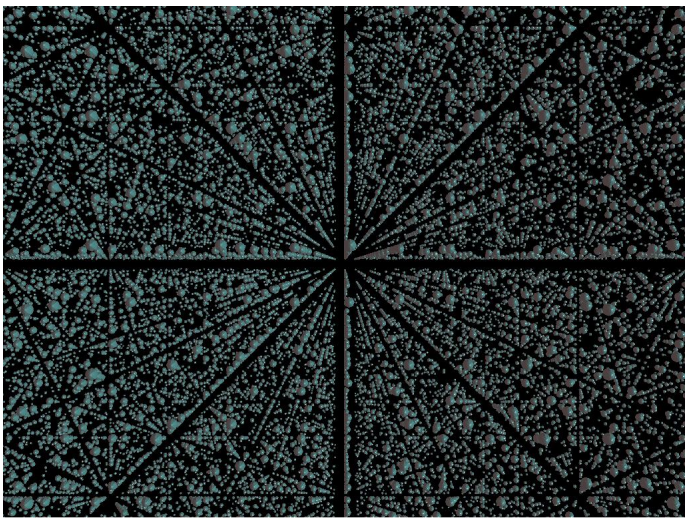

**Fig. 27.** The Gaussian cloud sampled regularly by 80000 points: zooming in on the center.

Weygaert, Rien van de and Willem Schaap, 2009, Lecture Notes in Physics, Data Analysis in Cosmology, The Cosmic Web: Geometric Analysis, 665