# Convergent geometric estimators with digital volume and surface integrals [*]

Jacques-Olivier Lachaud

Laboratoire de Mathématiques (LAMA), UMR CNRS 5127
Université Savoie Mont Blanc, Chambry, France,
`jacques-olivier.lachaud@univ-savoie.fr`

**Abstract.** This paper presents several methods to estimate geometric quantities on subsets of the digital space $\mathbb{Z}^d$. We take an interest both on global geometric quantities like volume and area, and on local geometric quantities like normal and curvatures. All presented methods have the common property to be *multigrid convergent*, i.e. the estimated quantities tend to their Euclidean counterpart on finer and finer digitizations of (smooth enough) Euclidean shapes. Furthermore, all methods rely on digital integrals, which approach either volume integrals or surface integrals along shape boundary. With such tools, we achieve multigrid convergent estimators of volume, moments and area in $\mathbb{Z}^d$, of normals, curvature and curvature tensor in $\mathbb{Z}^2$ and $\mathbb{Z}^3$, and of covariance measure and normals in $\mathbb{Z}^d$ even with Hausdorff noise.

**Keywords:** Digital geometry; Volume estimation; Moments estimation; Normal estimation; Curvatures estimation; Area estimation; Multigrid convergence; Digital integration; Integral invariants; Digital moments; Voronoi covariance measure; Stability.

## 1 Introduction

*Objectives.* We are interested in the geometry of subsets of the digital space $\mathbb{Z}^d$, where $\mathbb{Z}$ is the set of integer numbers. More precisely, when seeing these subsets as a sampling of a Euclidean shape, say $X$, we would like to recover an approximation of the geometry of $X$ with solely the information of its sampling. It is clear that this task cannot be done without further hypothesis on $X$ and on the sampling method. First, at a fixed sampling resolution, there are infinitely many shapes having the same sampling. Second, subsets of $\mathbb{Z}^d$ have no canonic tangent plane or differential geometry. To address the first issue, we will take a look at specific families of Euclidean shapes, generally by requiring smoothness properties. We will then show that we can achieve *multigrid convergence* properties for some estimators on such shapes, i.e. when the sampling gets finer and finer, the estimation gets better. The second issue is addressed by using digital integrals, i.e. well-chosen sums.

This paper presents the main ingredients and results of three methods that provide multigrid convergent estimators of the most common geometric quantities: volume, area, tangent, normal, curvatures, etc. Their common denominator is to use digital integrals, i.e. sums that approach integrals defined on the Euclidean shape. The stability of these integrals in turns induces the multigrid convergence of these estimators.

This topic is in fact rather old, since Gauss and Dirichlet already knew that the volume of a convex set can be approximated by counting digital points within (reported in [KR04]). Furthermore, it is related to numerical integration. The purpose of this paper is not to provide an exhaustive lists of multigrid convergent digital estimators. We may point several sources and surveys in the literature that provides many references or comparisons: [KŽ00,CK04,dVL09,CLR12]. This work compiles methods and results developed in several papers [CLL13,CLL14,LCLng,LT15,CLT14,CLMT15]. Note that the topic of digital geometric estimation through integrals is very active at the present time. Among the very recent works, we may quote the varifold approach of [Bue14,BLM15] for normal and mean curvature estimation, the estimation of intrinsic volumes of [EP16] with persistent homology, or the estimation of Minkowski tensors with an extension of the Voronoi covariance measure [HKS15].

*Main definitions and notations.* A *digitization process* is a family of maps mapping subsets of $\mathbb{R}^d$ towards subsets of $\mathbb{Z}^d$, parameterized by a positive real number $h$. The parameter $h$ defines the *gridstep* of the digitization, a kind of sampling distance. The digitization process $\mathtt{D}_h$ is *local* whenever $\mathbf{z} \in \mathtt{D}_h(X)$ depends solely on $X \cap N(h\mathbf{z})$, where $N(h\mathbf{z})$ is a neighborhood of radius $O(h)$ around point $h\mathbf{z} \in \mathbb{R}^d$.

In this work, we consider three simple local digitization processes, defined below after a few notations. If $\mathbf{z} \in \mathbb{Z}^d$, then $Q(\mathbf{z})$ denotes the (closed) unit $d$-dimensional cube of $\mathbb{R}^d$ centered on $\mathbf{z}$ and aligned with the axes of $\mathbb{Z}^d$. We further define $Q_h(\mathbf{z}) := h \cdot Q(\mathbf{z})$, so-called $h-cube$, as the scaled version of $Q(\mathbf{z})$ by factor $h$, *i.e.* $Q_h(\mathbf{z})$ is a $d$-dimensional cube centered at $h\mathbf{z}$ with edge length $h$). Then we define the *Gauss digitization* $\mathtt{G}_h(X)$, the *inner Jordan digitization* $\mathtt{J}_h^-(X)$ and *outer Jordan digitization* $\mathtt{J}_h^+(X)$ at step $h$ of a shape $X \in \mathbb{X}$ as (see Fig. 1):

$$\mathtt{G}_h(X) := \left\{ \mathbf{z} \in \mathbb{Z}^d, h\mathbf{z} \in X \right\} , \tag{1}$$

$$\mathtt{J}_h^-(X) := \left\{ \mathbf{z} \in \mathbb{Z}^d, Q_h(\mathbf{z}) \subset X \right\} , \tag{2}$$

$$\mathtt{J}_h^+(X) := \left\{ \mathbf{z} \in \mathbb{Z}^d, Q_h(\mathbf{z}) \cap X \neq \emptyset \right\} . \tag{3}$$

*First relations between shape and its digitization.* We will need to compare the geometry of the Euclidean shape $X$ and its topological boundary $\partial X$ with the "geometry" of their digitizations. So, for $Z \subset \mathbb{Z}^d$, we define the *body* of $Z$ at step $h$ as $[Z]_h := \bigcup_{\mathbf{z} \in Z} Q_h(\mathbf{z})$. We call *Jordan strip* the digitization $\mathtt{J}_h^0(X) := \mathtt{J}_h^+(X) \setminus \mathtt{J}_h^-(X)$, which is a kind of digitization of $\partial X$. The following properties
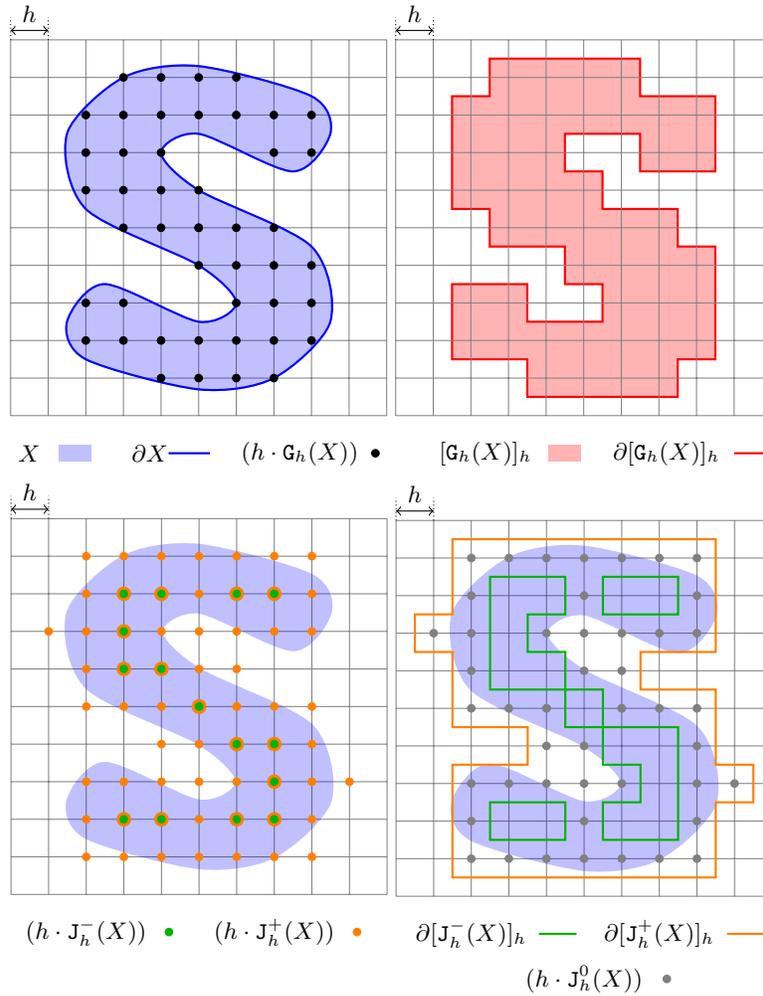
**Fig. 1.** Illustration of digitization processes and notations.

clarify the relations between $X$, $\partial X$ and their digitizations and are easy to derive [LCLng].

**Lemma 1** $\mathtt{J}_h^-(X) \subset \mathtt{G}_h(X) \subset \mathtt{J}_h^+(X)$ *and* $[\mathtt{J}_h^-(X)]_h \subset X \subset [\mathtt{J}_h^+(X)]_h$.

**Lemma 2** $[\mathtt{J}_h^0(X)]_h = [\mathtt{J}_h^+(X)]_h \setminus \mathrm{Int}([\mathtt{J}_h^-(X)]_h)$ *and* $\partial X \subset [\mathtt{J}_h^0(X)]_h$.

In fact, we can be more precise and relate these sets with the Hausdorff distance. Recall that the $\epsilon$-*offset* of a shape $X$, denoted by $X^\epsilon$, is the set of points of $\mathbb{R}^d$ at distance lower or equal to $\epsilon$ from $X$. We can state with some elementary arguments that boundaries of Jordan digitizations are close to the boundary of the shape in the Hausdorff sense:

**Lemma 3 ([LCLng], Lemma 3)** *Let $X$ be a compact domain of $\mathbb{R}^d$. Then* $[\mathsf{J}_h^0(X)]_h \subset (\partial X)^{\sqrt{d}h}$, $\partial[\mathsf{J}_h^-(X)]_h \subset (\partial X)^{\sqrt{d}h}$ *and* $\partial[\mathsf{J}_h^+(X)]_h \subset (\partial X)^{\sqrt{d}h}$.

The remarkable point here is that the sole requirement on $X$ is compactness ! If the shape $X$ has a smoother boundary, we can get tighter bounds for the Gauss digitization. Therefore, let the *medial axis* $\mathrm{MA}(\partial X)$ of $\partial X$ be the subset of $\mathbb{R}^d$ whose points have more than one closest point to $\partial X$. The *reach* $\mathrm{reach}(X)$ of $X$ is the infimum of the distance between $\partial X$ and its medial axis. Shapes with positive reach have a $C^2$ smooth boundary almost everywhere and have principal curvatures bounded by $\pm 1/\mathrm{reach}(X)$. We then have:

**Theorem 1 ([LT15])** *Let $X$ be a compact domain of $\mathbb{R}^d$ such that the reach of $\partial X$ is greater than $\rho$. Then, for any digitization step $0 < h < 2\rho/\sqrt{d}$, the Hausdorff distance between sets $\partial X$ and $\partial[\mathsf{G}_h(X)]_h$ is less than $\sqrt{d}h/2$. In particular,* $\partial[\mathsf{G}_h(X)]_h \subset (\partial X)^{\frac{\sqrt{d}}{2}h}$.

The *projection* $\pi^X$ of $\mathbb{R}^d \setminus \mathrm{MA}(\partial X)$ onto $\partial X$ is the map which associates to any point its closest point on $\partial X$. From the properties of the medial axis, the projection is defined almost everywhere in $\mathbb{R}^d$. We may thus associate to any point $\hat{\mathbf{x}} \in \partial[\mathsf{G}_h(X)]_h$ the point $\mathbf{x} := \pi^X(\hat{\mathbf{x}}) \in \partial X$, such that the distance between $\hat{\mathbf{x}}$ and its projection $\mathbf{x}$ is smaller than $\sqrt{d}h/2$. We have just constructed a mapping between a shape boundary and its digitization, which will help us for defining local geometric estimators.

*Multigrid convergence.* Let $V$ be any vector space (generally $\mathbb{R}$ or $\mathbb{R}^d$). A *geometric quantity* is an application that associates a value in $V$ to any subset of $\mathbb{R}^d$, with the property that it is invariant to some group operations, most often the group of rigid transformations. Notable examples are the volume and the area. A *local geometric quantity* is an application that associates a value in $V$ to a subset $X$ of $\mathbb{R}^d$ and a point $\mathbf{x}$ on $\partial X$. Common examples are the normal vector, the mean curvature or principal curvatures and directions. A *discrete geometric estimator* is an application that associates a value in $V$ to a subset of $\mathbb{Z}^d$ and a gridstep $h \in \mathbb{R}^+$. A *local discrete geometric estimator* is an application that associates a value in $V$ to a subset $Z$ of $\mathbb{Z}^d$, a point in $\partial[Z]_h$ and a gridstep $h \in \mathbb{R}^+$.

**Definition 1 (Multigrid convergence)** *A discrete geometric estimator $\hat{E}$ (resp. a local discrete geometric estimator $\hat{F}$) is said to be* multigrid convergent *to some geometric quantity $E$ (resp. to some local geometric quantity $F$) for the family of shapes $\mathbb{X}$ and digitization process $\mathsf{D}$, if and only if, for any $X \in \mathbb{X}$, there exists a gridstep $h_X > 0$, such that $\forall h, 0 < h < h_X$, we have:*

$$|E(X) - \hat{E}(\mathsf{D}_h(X), h)| \leq \tau_X(h), \qquad (4)$$

*respectively* $\forall \mathbf{x} \in \partial X, \forall \hat{\mathbf{x}} \in \partial[\mathsf{D}_h(X)]_h, \|\hat{\mathbf{x}} - \mathbf{x}\|_\infty \leq h$,

$$|F(X, \mathbf{x}) - \hat{F}(\mathsf{D}_h(X), \hat{\mathbf{x}}, h)| \leq \tau_X(h), \qquad (5)$$

*where the function $\tau_X : \mathbb{R}^+ \setminus \{0\} \to \mathbb{R}^+$ has null limit at $0$ and defines the* speed of convergence *of the estimator.*

In both definitions, the multigrid convergence property characterizes estimators that give better and better geometric estimates as the grid sampling gets finer and finer. We have now all the notions to study the multigrid convergence of several discrete geometric estimators.

## 2 Volume and moments estimators

In this section, $X$ is some compact domain of $\mathbb{R}^d$ and $Z$ is a subset of $\mathbb{Z}^d$. We take here an interest in estimating volume and moments from digital sets. These results will be used to define digital integral invariant estimators of curvatures in the following section. In the whole section, let $(p_i)_{i=1\ldots d}$ be the integers defining the moment exponents, with $0 \leq p_i \leq 2$, and let $\sigma := p_1 + \cdots + p_d$, with $\sigma \leq 2$.

*Moments and digital moments.* The $p_1 \cdots p_d$-moment of $X$ is defined as

$$m^{p_1 \cdots p_d}(X) := \int \cdots \int_X x_1^{p_1} \cdots x_d^{p_d} \, dx_1 \ldots dx_d.$$

The $0 \cdots 0$-moment of $X$ is the volume of $X$ (denoted $\mathrm{Vol}^d(X)$). The $p_1 \cdots p_d$-*digital moment of $Z$ at step $h$* is defined as

$$\hat{m}_h^{p_1 \cdots p_d}(Z) := h^{d+p_1+\cdots+p_d} \sum_{(z_1,\ldots,z_d) \in Z} z_1^{p_1} \cdots z_d^{p_d}.$$

The $0 \cdots 0$-digital moment of $Z$ is the digital volume of $Z$ (denoted by $\widehat{\mathrm{Vol}}^d(Z, h)$).

It is well known that $\widehat{\mathrm{Vol}}^d$ is multigrid convergent toward $\mathrm{Vol}^d$ for the family of convex shapes and the Gauss digitization, with a convergence speed of $O(h)$, and even faster for smoother shapes [Hux90,KN91,Mö99,Guo10]. We wish to go further on multigrid convergence of moments, so we take a special interest in (digital) moments of $h$-cubes. The following equalities, obtained by simple integration, show that discrepancies between digital and continuous moments begin with order two, and only when one $p_i = 2$.

**Lemma 4** *Let $\mathbf{z} \in \mathbb{Z}^d$. Point $\mathbf{z}$ is the Gauss digitization of $h$-cube $Q_h(\mathbf{z})$, but also its inner or outer Jordan digitization. Moments and digital moments of $h$-cubes satisfy $\hat{m}_h^{p_1 \cdots p_d}(\{\mathbf{z}\}) = m^{p_1 \cdots p_d}(Q_h(\mathbf{z})) + E(p_1, \ldots, p_d)$, where $E = \frac{h^{d+4}}{12}$ when one $p_i$ equals 2 and otherwise $E = 0$.*

*Errors in volume estimation.* The following volume "convergence" theorem is remarkable since it requires only the compactness of $X$. Its proof requires Lemma 4, a volume relation on symmetric difference of sets, the definition of Jordan strip, and Lemma 3.

**Theorem 2 ([LCLng])** *Let $X$ be a compact domain of $\mathbb{R}^d$. Let $\mathsf{D}$ be any digitization process such that $\mathsf{J}_h^-(X) \subset \mathsf{D}_h(X) \subset \mathsf{J}_h^+(X)$. Digital and continuous volumes are related as follows:*

$$\left| \mathrm{Vol}^d(X) - \widehat{\mathrm{Vol}}^d(\mathsf{D}_h(X), h) \right| \leq \mathrm{Vol}^d(\partial X^{\sqrt{d}h}). \tag{6}$$

Another proof, written independently, is in [HKS15]. This theorem states a multi-grid convergence property whenever $\partial X$ is $d-1$-rectifiable, but not in the general case: consider for instance the set $X$ of rational numbers in the unit cube. A more useful — but more restricted — convergence theorem relates this error bound to the area of $\partial X$. It uses Theorem 2 but also the fact that, for sets $X$ with positive reach, the volume of some $\epsilon$-offset of $\partial X$ is upper-bounded by a constant times the area of $\partial X$ (Proof of Lemma 10 [LT15].

**Theorem 3** *With the same hypotheses as Theorem 2 with the further requirement that the reach of $\partial X$ is greater than some value $\rho$. For $h < \rho/\sqrt{d}$, the volume estimator $\widehat{\mathrm{Vol}}^d$ is multigrid convergent toward the volume $\mathrm{Vol}^d$ with speed $2^{d+1}\sqrt{d}\mathrm{Area}(\partial X)h$.*

*Volume and moment estimation in a local neighborhood.* For the digital integral invariant method we will require convergence results on sets that are the intersection of two sets with positive reach, more precisely on sets of the form $X \cap B_R(\mathbf{x})$, where $B_R(\mathbf{x})$ denotes the ball of center $\mathbf{x} \in \mathbb{R}^d$ and radius $R$. The previous theorem cannot be applied as is. We must first bound the volume of offsets of the boundary of $X \cap B_R(\mathbf{x})$:

**Theorem 4 ([LCLng], Theorem 3)** *Let $X$ be a compact domain of $\mathbb{R}^d$ such that the reach of $\partial X$ is greater than $\rho$. Let $\mathbf{x} \in \mathbb{R}^d$. Let $R$ (the radius of the ball) and $h$ (the gridstep) be some positive numbers such that $h \leq \frac{R}{\sqrt{2d}}$ and $2R \leq \rho$.*

$$\mathrm{Vol}^d\left((\partial(X \cap B_R(\mathbf{x})))^{\sqrt{d}h}\right) \leq K_1(d)R^{d-1}h\,, \tag{7}$$

*where $K_1(d)$ is a constant that depends only on the dimension d.*

The proof first decomposes the set with the equality $(\partial(A \cap B))^\epsilon = ((\partial A) \cap B)^\epsilon \cup (A \cap (\partial B))^\epsilon$. Then it uses differential geometry and the fact that curvatures are bounded by the reach. The good point is that the geometry of $X$ does not intervene in the constants.

We have now all the keys to upperbound the error in volume estimation, and more generally moments estimation, within a ball around the boundary of a compact domain $X$.

**Theorem 5 ([LCLng], Theorem 4)** *We take the same hypotheses as the ones of Theorem 4 and the digitization process $\mathtt{D}$ defined in Theorem 2. Then digital moments within a ball $B_R(\mathbf{x})$ are multigrid convergent toward continuous moments as follows*

$$\left|m^{p_1 \cdots p_d}(X \cap B_R(\mathbf{x})) - \hat{m}_h^{p_1 \cdots p_d}\left(\mathtt{D}_h(X \cap B_R(\mathbf{x}))\right)\right|$$
$$\leq K_1(d)R^{d-1}(\|\mathbf{x}\|_\infty + 2R)^\sigma h + \frac{h^4}{12}V_d R^d, \tag{8}$$

*where $V_d$ is the volume of the unit d-dimensional ball. Furthermore, the term in $h^4$ is only present when one $p_i$ is equal to 2.*

The proof decomposes the errors in each $h$-cube induced by digitization. Errors in the interior of the set $X \cap B_R(\mathbf{x})$ are easily solved with Lemma 4. Errors close to the boundary of $X \cap B_R(\mathbf{x})$ are bounded with Theorem 4 and some care on moments with negative value.

## 3   Curvatures with digital integral invariants

*Curvature and mean curvature estimation.* If $\mathbf{x} \in \partial X$ and $\partial X$ smooth enough, one easily notices that the volume of $X \cap B_R(\mathbf{x})$ is related to the local differential geometry of $\partial X$ around $\mathbf{x}$ for infinitesimal values of $R$. Several authors [BGCF95,PWY$^+$07,PWHY09] have made explicit this relation:

**Lemma 5 ([PWHY09])** *For a sufficiently smooth shape $X$ in $\mathbb{R}^2$, $\mathbf{x} \in \partial X$ (resp. smooth shape $X'$ in $\mathbb{R}^3$, $\mathbf{x}' \in \partial X'$), we have*

$$\text{Vol}^2(X \cap B_R(\mathbf{x})) = \frac{\pi}{2}R^2 - \frac{\kappa(X, \mathbf{x})}{3}R^3 + O(R^4), \tag{9}$$

$$\text{Vol}^3(X' \cap B_R(\mathbf{x}')) = \frac{2\pi}{3}R^3 - \frac{\pi H(X', \mathbf{x}')}{4}R^4 + O(R^5), \tag{10}$$

*where $\kappa(X, \mathbf{x})$ is the curvature of $\partial X$ at $\mathbf{x}$ and $H(X', \mathbf{x}')$ is the mean curvature of $\partial X'$ at $\mathbf{x}'$.*

Since we have seen that we can approach volumes within a ball (see Theorem 5), it is very natural to define digital curvature estimators from the volume relations Eq.(9) and Eq.(10).

**Definition 2 ([CLL13])** *For any positive radius $R$, we define the 2D integral digital curvature estimator $\hat{\kappa}^R$ (resp. 3D integral mean digital curvature estimator $\hat{H}^R$) of a digital shape $Z \subset \mathbb{Z}^2$ at any point $\mathbf{x} \in \mathbb{R}^2$ (resp. of $Z' \subset \mathbb{Z}^3$ at any point $\mathbf{x}' \in \mathbb{R}^3$) and for a grid step $h > 0$ as:*

$$\forall 0 < h < R, \quad \hat{\kappa}^R(Z, \mathbf{x}, h) := \frac{3\pi}{2R} - \frac{3\widehat{\text{Vol}}^2(B_{R/h}(\mathbf{x}/h) \cap Z, h)}{R^3}, \tag{11}$$

$$\hat{H}^R(Z', \mathbf{x}', h) := \frac{8}{3R} - \frac{4\widehat{\text{Vol}}^3(B_{R/h}(\mathbf{x}'/h) \cap Z', h)}{\pi R^4}. \tag{12}$$

We can bound the error between these estimators and their associated geometric quantities as ([CLL13], but precise constants in [LCLng]):

**Theorem 6** *Let $X$ be a compact domain of $\mathbb{R}^2$ such that its boundary $\partial X$ is $C^3$-smooth and has reach greater than $\rho$. (Respectively let $X'$ be a compact domain of $\mathbb{R}^3$ such that its boundary $\partial X'$ is $C^3$-smooth and has reach greater than $\rho$).*

*The following bounds hold for $R < \rho/2$:*

$\forall 0 < h \leq R/2, \ \forall \mathbf{x} \in \partial X, \ \forall \hat{\mathbf{x}} \in \partial[\mathsf{G}_h(X)]_h \ \text{with} \ \|\hat{\mathbf{x}} - \mathbf{x}\|_\infty \leq h,$

$$\left| \hat{\kappa}^R(\mathsf{G}_h(X), \hat{\mathbf{x}}, h) - \kappa(X, \mathbf{x}) \right| \leq \left( 27\pi\sqrt{2}/4 + 3K_1(2) \right) R^{-2}h + O(R). \qquad (13)$$

$\forall 0 < h \leq R/\sqrt{6}, \ \forall \mathbf{x}' \in \partial X', \ \forall \hat{\mathbf{x}}' \in \partial[\mathsf{G}_h(X')]_h \ \text{with} \ \|\hat{\mathbf{x}}' - \mathbf{x}'\|_\infty \leq h,$

$$\left| \hat{H}^R(\mathsf{G}_h(X'), \hat{\mathbf{x}}', h) - H(X', \mathbf{x}') \right| \leq \left( 18\sqrt{3} + 4K_1(3)/\pi \right) R^{-2}h + O(R). \quad (14)$$

*Chosing $R = \Theta(h^{\frac{1}{3}})$ implies the multigrid convergence of estimator $\hat{\kappa}^R$ (resp. estimator $\hat{H}^R$) toward curvature $\kappa$ (resp. mean curvature $H$) with speed $O(h^{\frac{1}{3}})$.*

The first term in each error bound comes from the error done in volume estimation of $X \cap B_R(\mathbf{x})$ because point $\hat{\mathbf{x}}$ is not exactly on $\mathbf{x}$ but at distance $O(h)$ (Theorem 1). The second term comes from the digitization in the volume estimation (Theorem 5). The third term in the error comes from the Taylor expansion of Eq.(9) and Eq.(10). Since error terms are either decreasing with $R$ or increasing with $R$, we balance error terms to minimize the sum of errors and we get convergence of curvature estimators as immediate corollaries.

*Normal, principal curvatures and principal directions.* The same methodology can lead to estimators of principal curvatures or normal and principal directions. The idea is to use moments of zeroth, first, and second order instead of volumes. More precisely, the eigenvalues and eigenvectors of the covariance matrix of $X \cap B_R(\mathbf{x})$ were shown to hold curvature information ([PWY$^+$07], Theorem 2). The *covariance matrix* of a set $A \subset \mathbb{R}^3$ is easily defined from moments as:

$$\mathcal{V}(A) := \begin{bmatrix} m^{200}(A) & m^{110}(A) & m^{101}(A) \\ m^{110}(A) & m^{020}(A) & m^{011}(A) \\ m^{101}(A) & m^{011}(A) & m^{002}(A) \end{bmatrix}$$

$$- \frac{1}{m^{000}(A)} \begin{bmatrix} m^{100}(A) \\ m^{010}(A) \\ m^{001}(A) \end{bmatrix} \otimes \begin{bmatrix} m^{100}(A) \\ m^{010}(A) \\ m^{001}(A) \end{bmatrix}^T. \qquad (15)$$

The definition of *digital covariance matrix* $\mathcal{V}_h(Z)$ at step $h$ of a set $Z \subset \mathbb{Z}^3$ is similar, just replacing $m^{p_1 \cdots p_d}(A)$ by $\hat{m}_h^{p_1 \cdots p_d}(Z)$ in Eq.(15).

**Definition 3** *Let $Z \subset \mathbb{Z}^3$ be a digital shape and $h > 0$ be the gridstep. For $h < R$, we define the* integral principal curvature estimators $\hat{\kappa}_1^R$ and $\hat{\kappa}_2^R$ *of $Z$ at point $\mathbf{y} \in \mathbb{R}^3$, their respective* integral principal direction estimators $\hat{\mathbf{w}}_1^R$ and $\hat{\mathbf{w}}_2^R$, *and the* integral normal estimator $\hat{\mathbf{n}}^R$ *as*

$$\hat{\kappa}_1^R(Z, \mathbf{y}, h) := \frac{6}{\pi R^6}(\hat{\lambda}_2 - 3\hat{\lambda}_1) + \frac{8}{5R}, \quad \hat{\mathbf{w}}_1^R(Z, \mathbf{y}, h) := \hat{\nu}_1 \quad \hat{\mathbf{n}}^R(Z, \mathbf{y}, h) := \hat{\nu}_3,$$

$$\hat{\kappa}_2^R(Z, \mathbf{y}, h) := \frac{6}{\pi R^6}(\hat{\lambda}_1 - 3\hat{\lambda}_2) + \frac{8}{5R}, \quad \hat{\mathbf{w}}_2^R(Z, \mathbf{y}, h) := \hat{\nu}_2$$

*where $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \hat{\lambda}_3$ are the eigenvalues of $\mathcal{V}_h(B_{R/h}(\mathbf{y}/h) \cap Z)$, and $\hat{\nu}_1, \hat{\nu}_2, \hat{\nu}_3$ are their corresponding eigenvectors.*

Unfortunately, there is no hope in turning Theorem 5 into a multigrid convergence theorem for arbitrary moments, because a very small perturbation in the position of the ball center can lead to an arbitrary error on polynomial $x^\sigma$. However, due to their formulations, continuous and digital covariance matrices are invariant by translation, and error terms can thus be confined in a neighborhood around zero. Using this fact and error bounds of moments within the symmetric difference of two balls, we get:

**Theorem 7 ([CLL14])** *Let $X$ be a compact domain of $\mathbb{R}^3$ such that its boundary $\partial X$ has reach greater than $\rho$. Then the digital covariance matrix is multigrid convergent toward the covariance matrix for Gauss digitization for any radius $R < \frac{\rho}{2}$ and gridstep $h < \frac{R}{\sqrt{6}}$, with speed $O(R^4 h)$.*

The constant in $O$ is independent from the shape size or geometry. According to Definition 3, it remains to show that eigenvalues and eigenvectors of the digital covariance matrix are convergent toward the eigenvalues and eigenvectors of the continuous covariance matrix, when error on matrices tends to zero. Classical results of matrix perturbation theory (especially Lidskii-Weyl inequality and Davis-Kahan $\sin \theta$ Theorem) allow to conclude on the following relations:

**Theorem 8 ([CLL14,LCLng])** *Let $X$ be a compact domain of $\mathbb{R}^3$ such that its boundary $\partial X$ has reach greater than $\rho$ and has $C^3$-continuity. Then, for these shapes and for the Gauss digitization process, integral principal curvature estimators $\hat{\kappa}_1^R$ and $\hat{\kappa}_2^R$ are multigrid convergent toward $\kappa_1$ and $\kappa_2$ for small enough gridsteps $h$, choosing $R = kh^{\frac{1}{3}}$ and $k$ an arbitrary positive constant. Convergence speed is in $O(h^{\frac{1}{3}})$. Furthermore, integral principal direction estimators $\hat{\mathbf{w}}_1^R$ and $\hat{\mathbf{w}}_2^R$ are also convergent toward $\mathbf{w}_1$ and $\mathbf{w}_2$ with speed $O(h^{\frac{1}{3}})$ provided principal curvatures are distinct. Last the integral normal estimator $\hat{\mathbf{n}}^R$ is convergent toward the normal $\mathbf{n}$ with speed $O(h^{\frac{2}{3}})$.*

To our knowledge, these were the first estimators of principal curvatures shown to be multigrid convergent.

## 4  Digital voronoi covariance measure

Integral curvatures estimators are convergent. They are rather robust to the presence of Hausdorff noise in input data (see [CLL14]). However, this robustness comes with no guarantee. We present here another approach that reaches this goal, and which can even be made robust to outliers. The first idea is to use a distance function to input data, which is stable to perturbations [CCSM11]. The second idea is to notice that Voronoi cells of input data tend to align with normals of the underlying shape. To make this idea more robust, it suffices to integrate the covariance of gradient vectors of the distance function within a local neighborhood: this is called *Voronoi covariance measure* [MOG11].

**Definition 4** *A function $\delta : \mathbb{R}^d \to \mathbb{R}^+$ is called* distance-like *if (i) $\delta$ is proper, i.e. $\lim_{\|x\| \to \infty} \delta(x) = \infty$, (ii) $\delta^2$ is 1-semiconcave, that is $\delta^2(.) - \|.\|^2$ is concave.*

It is worthy to note that the standard distance $d_K$ to a compact $K$ is distance-like. Clearly this distance is robust to Hausdorff noise, i.e. $d_H(K, K') < \epsilon$ implies $\|d_K - d_{K'}\|_\infty < \epsilon$. Although we will not go into more details here, the *distance to a measure* [MOG11,CLMT15] is even resilient to outliers and the error is bounded by the Wasserstein-2 distance between measures. Let $\mathbf{N}_\delta := \frac{1}{2}\nabla\delta^2$.

**Definition 5 ($\delta$-VCM)** *The $\delta$-Voronoi Covariance Measure (VCM) is a tensor-valued measure. Given any non-negative probe function $\chi$, i.e. an integrable function on $\mathbb{R}^d$, we associate a positive semi-definite matrix defined by*

$$\mathcal{V}_\delta^R(\chi) := \int_{\delta^R} \mathbf{N}_\delta(\mathbf{x}) \otimes \mathbf{N}_\delta(\mathbf{x}) \cdot \chi(\mathbf{x} - \mathbf{N}_\delta(\mathbf{x}))\,d\mathbf{x}, \qquad (16)$$

*where $\delta^R := \delta^{-1}((-\infty, R])$.*

Note that $\mathbf{N}_\delta$ is defined almost everywhere in $\mathbb{R}^d$.

In the following, we define $\delta$ as the distance to a compact $d_K$ in all results, but one should keep in mind that these results are extensible to arbitrary distance-like functions. Then $\mathbf{N}_{d_K}$ corresponds to the vector of the projection onto $K$, except for the sign. The $d_K$-VCM corresponds then to the covariance matrix of Voronoi cells of points of $K$, restricted to a maximum distance $R$, and weighted by the probe function.

**Definition 6** *Let $Z \subset \mathbb{Z}^d$ and $h > 0$. The digital Voronoi Covariance Measure of $Z$ at step $h$ and radius $R$ associates to a probe function $\chi$ the matrix:*

$$\hat{\mathcal{V}}_{Z,h}^R(\chi) := \sum_{\mathbf{z} \in \text{vox}(Z_h, R)} h^d \mathbf{N}_{d_{Z_h}}(\mathbf{z}) \otimes \mathbf{N}_{d_{Z_h}}(\mathbf{z})\chi(\mathbf{z} - \mathbf{N}_{d_{Z_h}}(\mathbf{z})), \qquad (17)$$

*where $Z_h := h \cdot Z$ and $\text{vox}(Z_h, R)$ designates the points of $h \cdot \mathbb{Z}^d$ whose $h$-cube is completely inside the $R$-offset of $Z_h$. See Fig. 2 for an illustration.*

Since $\mathbf{N}_{d_{Z_h}}$ corresponds to the projection onto $Z_h$, the previous formulation is easily decomposed per Voronoi cells of $Z_h$ and can be computed exactly by simple summations. The digital VCM is shown to be close to the VCM for digitizations of smooth enough shapes [CLT14]. Errors are related first to the difference between $\partial X$ and its digitization $\partial[\mathtt{G}_h(X)]_h$ (essentially bounded by Theorem 5.1 of [MOG11]). Secondly they are linked to the transformation of the integral in $\mathcal{V}_\delta^R$ in a sum in $\hat{\mathcal{V}}_{Z,h}^R$ (bounded by the fact that the projection is stable and that the strip $Z_h^R \setminus \text{vox}(Z_h, R)$ is negligible). Proofs ressemble the ones of Section 2.

**Theorem 9 ([CLT14])** *For $X$ a compact domain of $\mathbb{Z}^3$ whose boundary is $C^2$-smooth and has reach greater than $\rho > 0$. Then, for any $R < \rho/2$ and probe function $\chi$ with finite support diameter $r$ and for small enough $h > 0$, letting $Z = \partial[\mathtt{G}_h(X)]_1 \cap (\mathbb{Z} + \frac{1}{2})^3$, we have:*

$$\|\mathcal{V}_{\partial X}^R(\chi) - \hat{\mathcal{V}}_{Z,h}^R(\chi)\|_{\text{op}} \le O\big(\text{Lip}(\chi)(r^3 R^{\frac{5}{2}} + r^2 R^3 + rR^{\frac{9}{2}})h^{\frac{1}{2}}$$
$$+ \|\chi\|_\infty[(r^3 R^{\frac{3}{2}} + r^2 R^2 + rR^{\frac{7}{2}})h^{\frac{1}{2}} + r^2 Rh]\big).$$
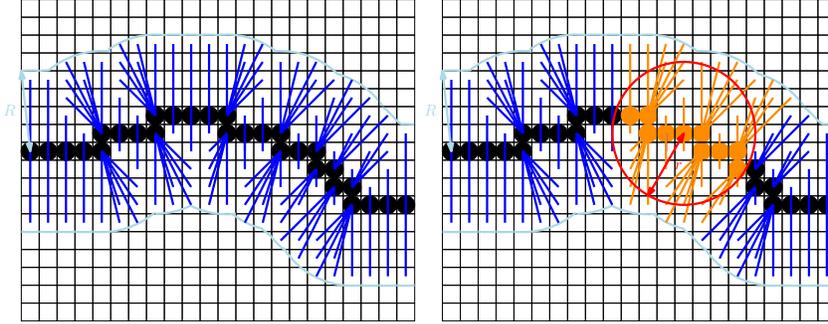
**Fig. 2.** Left: the limits of the $R$-offset of a set of digital points are drawn as a cyan contour, while vectors connecting points within the R-offset (i.e. $Z_h^R$) to their projection are drawn in deep blue. Right: Voronoi cells defining the VCM. The domain of integration for a kernel $\chi$ of radius $r$ is drawn in dark orange (both germs and projection vectors). The kernel itself is drawn in red.

As one can see, the digital VCM is an approximation of the VCM, but the quality of the approximation is related not only to the gridstep $h$ but also to two parameters, the size $r$ of the support of $\chi$ and the distance $R$ to input data, which defines the computation window.

An interesting fact about the VCM is that it is in some sense complementary to integral invariants. It does not aim at fitting a tangent plane to $\partial X$, it aims at fitting a line to the normal vector $\mathbf{n}$ to $\partial X$. It carries thus information about the normal to $\partial X$. This is shown by the relation below ([CLT14], Lemma 2, adaptation of a relation in [MOG11]):

$$\left\| \frac{3}{2\pi R^3 r^2} \mathcal{V}_{d_{\partial X}}^R (\mathbf{1}_{B_r(\mathbf{x})}) - [\mathbf{n}(\mathbf{x}) \otimes \mathbf{n}(\mathbf{x})] \right\|_{\mathrm{op}} = O(r + R^2). \tag{18}$$

Putting together Theorem 9 and Eq.(18), as well as an error term coming from the positioning error of the kernel $\chi$, provides a multigrid convergence theorem, that has the remarkable property to be stable to Hausdorff perturbation of digital data.

**Theorem 10 ([CLT14])** *With the same hypothesis as in Theorem 9, $\chi$ bounded Lipschitz, and denoting $\hat{\mathbf{n}}_\chi^R$ the eigenvector associated to the highest eigenvalue of $\hat{\mathcal{V}}_{Z,h}^R(\chi)$, then $\hat{\mathbf{n}}_\chi^R$ is multigrid convergent toward the normal vector to $\partial X$, with speed $O(h^{\frac{1}{8}})$ when both $R$ and the support $r$ of $\chi$ are chosen in $\Theta(h^{\frac{1}{4}})$.*

Experiments indicate a much faster convergence speed (close to $O(h)$) even in presence of noise. The discrepancy comes mainly from the fact that $\chi$ is any bounded Lipschitz function while Eq.(18) is valid for the characteristic function of $B_r(\mathbf{x})$.

# 5 Digital surface integration

Until now, convergence results where achieved by approximating well-chosen volume integrals around input data. What can we say if we wish to approach integrals defined over the shape boundary, given only the shape digitization. We focus here on Gauss digitization and we write $\partial_h X$ for $\partial[\mathsf{G}_h(X)]_h$. A natural answer is to define a mapping between the digitized boundary $\partial_h X$ and the continuous boundary $\partial X$. Using standard geometric integration results, a surface integral defined over $\partial X$ can be transformed into a surface integral over $\partial_h X$ by introducing the Jacobian of this mapping. However, we have to face several difficulties in our case. The first one is that, starting from 3D, $\partial_h X$ may not even be a manifold, whatever the smoothness of $\partial X$ and the gridstep $h$ [SLS07]. Hence, it is not possible to define an injective mapping between the two sets. The second difficulty is that the underlying continuous surface is unknown, so we have to define a natural mapping without further hypothesis on the shape. The best candidate is the projection $\pi^{\partial X}$ onto $\partial X$, which is defined everywhere in the $R$-offset of $\partial X$, for $R$ smaller than the reach of $\partial X$. It is nevertheless easily seen that $\pi^{\partial X}$, although surjective, is generally not injective between the digitized boundary and the continuous boundary. In the following, we bound these problematic zones in order to define convergent digital surface integrals. For simplicity, we write $\pi$ for $\pi^{\partial X}$ and $\pi'$ for its restriction to $\partial_h X$.

**Definition 7 ([LT15])** *Let $Z \subset \mathbb{Z}^d$ be a digital set, with gridstep $h > 0$ between samples. Let $\mathrm{Bd}(Z) = \{(\mathbf{z}_1 + \mathbf{z}_2)/2, \mathbf{z}_1 \in Z, \mathbf{z}_2 \notin Z, \|\mathbf{z}_1 - \mathbf{z}_2\|_1 = 1\}$. It corresponds to centroid of $d - 1$-cells separating $Z$ from $\mathbb{Z}^d \setminus Z$. Let $f : \mathbb{R}^d \to \mathbb{R}$ be an integrable function and $\hat{\mathbf{n}}$ be a digital normal estimator. We define the* digital surface integral *by*

$$\mathrm{DI}_{f,\hat{\mathbf{n}}}(Z, h) := \sum_{\mathbf{y} = h\mathbf{z}, \mathbf{z} \in \mathrm{Bd}(Z)} f(\mathbf{y}) |\hat{\mathbf{n}}(\mathbf{y}) \cdot \mathbf{n}_h(\mathbf{y})|,$$

*where $\mathbf{y}$ visits faces of $\partial[Z]_h$ and $\mathbf{n}_h(\mathbf{y})$ is the trivial normal of this face.*

We show the following facts in [LT15], if $X$ is a compact domain of $\mathbb{R}^d$ such that $\mathrm{reach}(\partial X) > \rho > 0$:

- for $d = 3$ and $h < 0.198\rho$, $\partial_h X$ may not be a manifold only at places at distance lower than $h$ to parts of $\partial X$ whose normal makes an angle smaller than $1.26h/\rho$ to some axis;
- for arbitrary $d \geq 2$ and $h < \rho/\sqrt{d}$, let $\mathbf{y} \in \partial_h X$ and $\mathbf{n}_h(\mathbf{y})$ be its (trivial) normal vector; then the angle between the normal $\mathbf{n}(\mathbf{x})$ to $\partial X$ at $\mathbf{x} = \pi(\mathbf{y})$ and $\mathbf{n}_h(\mathbf{y})$ cannot be much greater than $\pi/2$, since $\mathbf{n}(\mathbf{x}) \cdot \mathbf{n}_h(\mathbf{y}) \geq -\sqrt{3d}h/\rho$;
- let $\mathrm{Mult}(\partial X)$ be the points of $\partial X$ that are images of several points of $\partial_h X$ by $\pi'$, then it holds for at least one of the point $\mathbf{y}$ in the fiber of $\mathbf{x} \in \mathrm{Mult}(\partial X)$ under $\pi'$ that $\mathbf{n}(\mathbf{x}) \cdot \mathbf{n}_h(\mathbf{y})$ is not positive;
- the Jacobian of $\pi'$ is almost everywhere $|\mathbf{n}(\mathbf{x}) \cdot \mathbf{n}_h(\mathbf{y})|(1 + O(h))$;
- areas are related with $\mathrm{Area}(\partial_h X) \leq 2^{d+2} d^{\frac{3}{2}} \mathrm{Area}(\partial X)$;

(a) $h = 0.1$
$A_\pi = 58.40\%, A_{nm} = 1.57\%$

(b) $h = 0.05$
$A_\pi = 30.71\%, A_{nm} = 0.38\%$

(c) $h = 0.025$
$A_\pi = 15.88\%, A_{nm} = 0.09\%$



(a) $h = 0.04$
$A_\pi = 62.22\%, A_{nm} = 0.88\%$

(b) $h = 0.02$
$A_\pi = 30.54\%, A_{nm} = 0.12\%$

(c) $h = 0.01$
$A_\pi = 16.67\%, A_{nm} = 0.03\%$
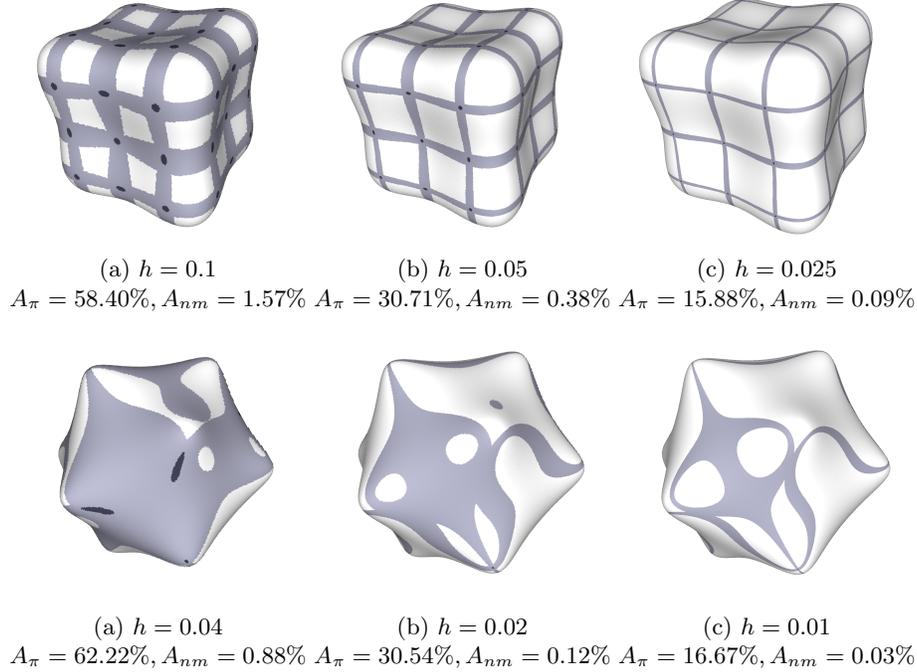
**Fig. 3.** Properties of several Gauss digitizations of two polynomial surfaces (top row displays a Goursat's smooth cube and bottom row displays Goursat's smooth icosahedron). Zones in dark grey indicates the surface parts where the Gauss digitization might be non-manifold (their relative area is denoted by $A_{nm}$). Zones in light grey (and dark grey) indicates the surface parts where projection $\pi$ might not be a homeomorphism (their relative area is denoted by $A_\pi$). Clearly, both zones tends to area zero as the gridstep gets finer and finer, while parts where digitization might not be manifold are much smaller than parts where $\pi$ might not be homeomorphic.

–  hence, when $h \leq R/\sqrt{d}$, the area of the non-injective part of $\pi'$ decreases with $h$: Area(Mult($\partial X$)) $\leq K_2$ Area($\partial X$) $h$, with $K_2 \leq 2\sqrt{3}\, d^2\, 4^d/\rho$.

The preceding properties show that problematic places on $\partial_h X$ for the integration have decreasing area. Furthermore, if $E(\hat{\mathbf{n}}, \mathbf{n}) := \sup_{\mathbf{y} \in \partial_h X} \|\mathbf{n}(\pi(\mathbf{y})) - \hat{\mathbf{n}}(\mathbf{y})\|$, errors in normal estimation induce propotional errors in surface integration. We can then prove the multigrid convergence of the digital surface integral toward the surface integral.

**Theorem 11** *Let $X$ be a compact domain whose boundary has positive reach $\rho$. For $h \leq \rho/\sqrt{d}$, the digital surface integral is multigrid convergent toward the integral over $\partial X$. More precisely, for any integrable function $f : \mathbb{R}^d \to \mathbb{R}$ with*

*bounded Lipschitz norm* $\|f\|_{\mathrm{BL}} := \mathrm{Lip}(f) + \|f\|_\infty$, *one gets*

$$\left| \int_{\partial X} f(x)dx - \mathrm{DI}_{f,\hat{\mathbf{n}}}(\mathtt{G}_h(X), h) \right| \leq \mathrm{Area}(\partial X)\, \|f\|_{\mathrm{BL}} \left( O(h) + O(E(\hat{\mathbf{n}}, \mathbf{n})) \right).$$

The constant involved in the notation $O(.)$ only depends on the dimension $d$ and the reach $\rho$. Note that Theorem 8 and Theorem 10 have shown that there exist normal estimators such that $E(\hat{\mathbf{n}}, \mathbf{n})$ tends to zero as $h$ tends to zero. Experimental evaluation of the digital surface integral shows a better convergence in practice for area analysis, with convergence speed close to $O(h^2)$ both for integral normal estimator and digital VCM normal estimator.

## References

BGCF95.  J. W. Bullard, E. J. Garboczi, W. C. Carter, and E. R. Fullet. Numerical methods for computing interfacial mean curvature. *Computational materials science*, 4:103–116, 1995.

BLM15.  B. Buet, G. P. Leonardi, and S. Masnou. Discrete varifolds: A unified framework for discrete approximations of surfaces and mean curvature. In J. F. Aujol et al., editor, *Proc. Scale Space and Variational Methods in Computer Vision*, volume 9087 of *LNCS*, pages 513–524. Springer, 2015.

Bue14.  B. Buet. *Approximation de surfaces par des varifolds discrets: représentation, courbure, rectifiabilité.* PhD thesis, Université Claude Bernard-Lyon I, France, 2014.

CCSM11.  F. Chazal, D. Cohen-Steiner, and Q. Mérigot. Geometric inference for probability measures. *Foundations of Computational Mathematics*, 11(6):733–751, December 2011.

CK04.  D. Coeurjolly and R. Klette. A comparative evaluation of length estimators of digital curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2):252–258, 2004.

CLL13.  D. Coeurjolly, J. O. Lachaud, and J. Levallois. Integral based curvature estimators in digital geometry. In *Discrete Geometry for Computer Imagery*, number 7749 in LNCS, pages 215–227. Springer, 2013.

CLL14.  D. Coeurjolly, J.-O. Lachaud, and J. Levallois. Multigrid convergent principal curvature estimators in digital geometry. *Computer Vision and Image Understanding*, 129:27–41, 2014.

CLMT15.  L. Cuel, J.-O. Lachaud, Q. Mérigot, and B. Thibert. Robust geometry estimation using the generalized voronoi covariance measure. *SIAM Journal on Imaging Science*, 8(2):1293–1314, 2015.

CLR12.  D. Coeurjolly, J.-O. Lachaud, and T. Roussillon. *Digital Geometry Algorithms, Theoretical Foundations and Applications of Computational Imaging*, volume 2 of *LNCVB*, chapter Multigrid convergence of discrete geometric estimators, pages 395–424. Springer, 2012.

CLT14.  L. Cuel, J.-O. Lachaud, and B. Thibert. Voronoi-based geometry estimator for 3d digital surfaces. In *Discrete Geometry for Computer Imagery*, pages 134–149. Springer, 2014.

dVL09.  F. de Vieilleville and J.-O. Lachaud. Comparison and improvement of tangent estimators on digital curves. *Pattern Recognition*, 42(8):1693–1707, aug 2009.

EP16.  H. Edelsbrunner and F. Pausinger. Approximation and convergence of the intrinsic volume. *Advances in Mathematics*, 287:674–703, 2016.

Guo10.  J. Guo. On lattice points in large convex bodies. *ArXiv e-prints*, 2010.

HKS15.  D. Hug, M. Kiderlen, and A. M. Svane. Voronoi-based estimation of minkowski tensors from finite point samples, 2015.

Hux90.  M. N. Huxley. Exponential sums and lattice points. *Proc. London Math. Soc.*, 60:471–502, 1990.

KN91.  E. Krätzel and W. G. Nowak. Lattice points in large convex bodies. *Monatshefte für Mathematik*, 112:61–72, 1991.

KR04.  R. Klette and A. Rosenfeld. *Digital Geometry: Geometric Methods for Digital Picture Analysis*. Series in Computer Graphics and Geometric Modelin. Morgan Kaufmann, 2004.

KŽ00.  R. Klette and J. Žunić. Multigrid convergence of calculated features in image analysis. *Journal of Mathematical Imaging and Vision*, 13:173–191, 2000.

LCLng.  J.-O. Lachaud, D. Coeurjolly, and J. Levallois. Robust and convergent curvature and normal estimators with digital integral invariants. In *Discrete Curvature*, Lecture Notes in Mathematics. Springer International Publishing, 2016, forthcoming.

LT15.  J.-O. Lachaud and B. Thibert. Properties of gauss digitized sets and digital surface integration. *Journal of Mathematical Imaging and Vision*, 2015. Online since 21 August 2015.

M̈99.  W. Müller. Lattice points in large convex bodies. *Monatshefte für Mathematik*, 128:315–330, 1999.

MOG11.  Q. Mérigot, M. Ovsjanikov, and L. Guibas. Voronoi-based curvature and feature estimation from point clouds. *Visualization and Computer Graphics, IEEE Transactions on*, 17(6):743–756, 2011.

PWHY09.  H. Pottmann, J. Wallner, Q. Huang, and Y. Yang. Integral invariants for robust geometry processing. *Computer Aided Geometric Design*, 26(1):37–60, 2009.

PWY+07.  H. Pottmann, J. Wallner, Y. Yang, Y. Lai, and S. Hu. Principal curvatures from the integral invariant viewpoint. *Computer Aided Geometric Design*, 24(8-9):428–442, 2007.

SLS07.  P. Stelldinger, L. J. Latecki, and M. Siqueira. Topological equivalence between a 3d object and the reconstruction of its digital image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(1):126–140, 2007.