Minimal Stencils for Discretizations of Anisotropic PDEs Preserving Causality or the Maximum Principle

Jean-Marie Mirebeau[*]

April 1, 2016

Abstract

We consider discretizations of anisotropic diffusion and of the anisotropic eikonal equation, on two dimensional cartesian grids, which preserve their structural properties: the maximum principle for diffusion, and causality for the eikonal equation. These two PDEs embed geometric information, in the form of a field of diffusion tensors and of a Riemannian metric respectively. Common knowledge is that, when these tensors are strongly anisotropic, monotonous or causal discretizations of these PDEs cannot be strictly local: numerical schemes need to involve interactions between each point and the elements of a *stencil*, which is not limited to its immediate neighbors on the discretization grid. Using tools from discrete geometry we identify the smallest valid stencils, in the sense of convex hull inclusion. We also estimate, for a fixed condition number but a random tensor orientation, the worst case and average case radius of these minimal stencils, which is relevant for numerical error analysis.

1 Introduction

The Partial Differential Equation (PDE) of diffusion obeys an important structural property, named the maximum principle. Numerical schemes which preserve this structure benefit from the discrete maximum principle [Cia70], a strong stability guarantee. The eikonal equation is the PDE formulation of an optimal control problem: to find the shortest exit path from a given domain. Discretization schemes which preserve its causal structure can be solved in a single pass using the Fast Marching algorithm [Tsi95, SV01], which has a quasi-linear complexity. Motivation for structure preservation in the discretization of PDEs is therefore plentiful, and stems from theoretical as much as practical considerations. In this intention, a variety of numerical schemes have been developed; for instance and without exhaustivity [MW53, BOZ04, Obe06, Wei98, FM13] for anisotropic diffusion, and [SV01, AM11, Mir14] for anisotropic eikonal equations.

Non-isotropic PDEs have numerous applications, of which we can only give a glimpse. Anisotropic diffusion is required in porous media simulation [Dro14], or stochastic control [BOZ04], but is also fundamental in image processing [Wei98]. The Anisotropic eikonal equation is naturally required for trajectory planning, but is also relevant for seismic imaging [SV03], and extensively used in medical image segmentation [BC10]. When discretizing PDEs, anisotropy usually comes with technical difficulties. Indeed, monotone or causal numerical schemes cannot be strictly local, but need to introduce interactions between each point and a *stencil* of potentially distant neighbors [SV03, Koc95]. The objective of this paper is to provide a complete

---

[*]Laboratoire de Mathématiques d'Orsay, University Paris-Sud, CNRS, University Paris-Saclay, 91405, Orsay

qualitative and quantitative understanding of this non-locality, in the limited setting of two dimensional discretizations on cartesian grids: we identify the smallest stencils, in the sense of convex hull inclusion, and we sharply estimate their worst case and average radius. Note that minimal stencils were similarly identified in [BCM15] for the monotone discretization of the two dimensional Monge-Ampere equation.

We introduce in §1.1 the concept of D-diffusion stencil, where D is a positive definite diffusion tensor, which allows to discretize anisotropic diffusion PDEs while obeying the maximum principle. Likewise, the concept of M-eikonal stencil, where M is a positive definite matrix encoding the metric structure, allows to discretize anisotropic eikonal PDEs while preserving their causal structure.

We identify in §1.2 the smallest D-diffusion stencil, and the smallest M-eikonal stencil, in the sense of convex hull inclusion, see Theorems 1.12 and 1.13. These discretizations turn out to coincide with numerical schemes previously introduced by the author in [FM13, Mir14], see Proposition 1.17. We refer to these earlier works for implementation details, extensive numerical studies, and comparisons with several competing methods.

We present in §1.3 quantitative estimates of the size of the optimal D-diffusion and M-eikonal stencils, which are identified in the previous subsection §1.2. For a given tensor condition number κ, the stencil radius strongly depends on the angle θ between the grid axes and the tensor eigenvectors. Our estimates provide, for a given κ, both the worst and the average stencil radius w.r.t. θ, see Theorem 1.19.

We briefly discuss in §1.4 the case of PDEs with non-constant coefficients. Indeed the results of this paper, being strictly local, are for simplicity stated for PDEs with constant coefficients. We describe problem instances to which our results apply particularly well, and show how Theorem 1.19 could be the starting point of a numerical error analysis.

Notation: Let $\langle f, g \rangle$ and $\det(f, g)$ denote respectively the scalar product and the determinant of two vectors $f, g \in \mathbb{R}^2$. The perpendicular vector $f^\perp$ to f obeys $\langle f^\perp, g \rangle = \det(f, g)$. Let S_2 denote the collection of symmetric 2×2 matrices, and let S_2^+ denote the subset of those which are positive definite. To each $D \in S_2^+$ we associate the norm defined for all $e \in \mathbb{R}^2$ by

$$\|e\|_D := \sqrt{\langle e, De \rangle}. \tag{1}$$

Let e^{T} (resp. A^{T}) denote the transpose of a vector e (resp. matrix A), and let $e \otimes e := ee^{\mathrm{T}}$ denote the tensor product of the vector $e \in \mathbb{R}^2$ with itself, which is a positive semi-definite matrix. We denote by $\mathrm{Hull}(V)$ the convex hull of a set $V \subseteq \mathbb{R}^2$, and use the shorthand

$$[x_1, \cdots, x_r] := \mathrm{Hull}(\{x_1, \cdots, x_r\}),$$

for any $x_1, \cdots, x_r \in \mathbb{R}^2$. We say that some vectors $e_1, \cdots, e_r \in \mathbb{R}^2$ are ordered trigonometrically if they are non-zero, pairwise non-positively collinear, and if the normalized vectors $e_1/\|e_1\|, \cdots, e_r/\|e_r\|$ are indexed according to the cyclic order on the unit circle. In that case we say that e_i and e_{i+1} are trigonometrically consecutive within this family of vectors, for any $0 \le i < r$ and with the convention $e_0 := e_r$. Finally, we denote by $\mathbf{Z}$ the collection of elements of $\mathbb{Z}^2$ which have co-prime coordinates

$$\mathbf{Z} := \{e \in \mathbb{Z}^2; \gcd(e) = 1\},$$

where $\gcd(e) := \gcd(a, b)$ if $e = (a, b) \in \mathbb{Z}^2$. Note that $(0, 0) \notin \mathbf{Z}$.

1.1 Structure preserving schemes for the Diffusion and Eikonal equations

We introduce in this section discretization schemes for anisotropic diffusion and eikonal PDEs, based on adequate notions of stencils. We focus on constant PDE coefficients, see §1.4 for a discussion on non-constant coefficients. This restriction is not limiting since the definitions and results presented in §1.1, §1.2 and §1.3 are purely local. As mentioned in the introduction, our main results are limited to discretizations on two dimensional cartesian grids, which take the form

$$x_0 + hR\mathbb{Z}^2,$$

where x_0 is an offset, $h > 0$ a scale, and R a rotation. Without loss of generality, we assume that these parameters take their canonical values, so that the discretization grid is simply $\mathbb{Z}^2$.

The results presented in this subsection are fairly classical, but are nevertheless required to justify the axioms chosen for the definition of diffusion and eikonal stencils. Classical references on these topics include [Cia70] on the discrete maximum principle, [Wei98] on anisotropic diffusion in image processing, and [SV03] on anisotropic eikonal equations. The results of this subsection are established in Appendix A.

Discretization of anisotropic Diffusion. We introduce the concept of D-Diffusion stencil in Proposition 1.1, show how it leads to a natural discretization of diffusion PDEs in Proposition 1.2, and justify the so-called Non-Negativity axiom by a stability property in Proposition 1.4. Throughout this section, the diffusion tensor D is given and assumed to be positive definite, but it could be severely ill conditioned.

Definition 1.1. *A D-diffusion stencil, $D \in S_2^+$, is a finitely supported map $\gamma : \mathbf{Z} \to \mathbb{R}$ obeying:*

- *Consistency: $\frac{1}{2} \sum_{e \in \mathbf{Z}} \gamma(e)\, e \otimes e = D$.*

- *Symmetry: $\gamma(-e) = \gamma(e)$ for all $e \in \mathbf{Z}$.*

- *Non-negativity: $\gamma(e) \geq 0$ for all $e \in \mathbf{Z}$.*

A D-diffusion stencil γ may also be regarded as the finite collection of points $\mathrm{supp}(\gamma)$, together with the weights $\gamma(e)$, $e \in \mathrm{supp}(\gamma)$. Our next result justifies the Consistency axiom.

Proposition 1.2. *Consider a finitely supported $\gamma : \mathbf{Z} \to \mathbb{R}$ obeying property Symmetry of Definition 1.1. Then property Consistency is equivalent to each of the following properties:*

- *Consistency': $\sum_{e \in \mathbf{Z}} \gamma(e)(u(x + e) - u(x)) = \mathrm{Tr}(D\nabla^2 u(x))$ for any quadratic $u : \mathbb{R}^2 \to \mathbb{R}$.*

- *Consistency'': $\frac{1}{2} \sum_{e \in \mathbf{Z}} \gamma(e)(u(x + e) - u(x))^2 = \|\nabla u(x)\|_D^2$ for any linear $u : \mathbb{R}^2 \to \mathbb{R}$.*

To each finitely supported $\gamma : \mathbf{Z} \to \mathbb{R}$ we attach the linear operator L_γ defined for all $u : \mathbb{Z}^2 \to \mathbb{R}$ and $x \in \mathbb{Z}^2$ by

$$L_\gamma u(x) := \sum_{e \in \mathbf{Z}} \gamma(e)(u(x + e) - u(x)). \tag{2}$$

If γ obeys properties Symmetry and Consistency of Definition 1.1, then L_γ is by Proposition 1.2 a natural discretization of $\mathrm{Tr}(D\nabla^2 u)$ and of $\mathrm{div}(D\nabla u)$. Note that these two differential operators are identical, as well as their discretizations, since the diffusion tensor D is constant over the domain. Divergence form and non-divergence form diffusion PDEs differ however, as well as their discretizations, in the case of a space varying tensor field, see §1.4.

3

Observing that $\text{Id} = e_1 \otimes e_1 + e_2 \otimes e_2$, where (e_1, e_2) denotes the canonical basis of $\mathbb{R}^2$, we obtain the Id-diffusion stencil defined by $\gamma(\pm e_1) = \gamma(\pm e_2) = 1$ and $\gamma = 0$ elsewhere. The operator (2) associated to this specific stencil is the standard five points discretization of the Laplace operator Δ.

The maximum principle is a fundamental property of the diffusion PDE: if $u \in C^2(\mathbb{R}_+ \times \mathbb{R}^2)$ is bounded and obeys $\partial_t u = \text{div}(D\nabla u)$, then for any $t \geq 0$ one has

$$\inf_{\mathbb{R}^2} u_0 \leq u_t \leq \sup_{\mathbb{R}^2} u_0,$$

where $u_t : x \mapsto u(t, x)$. Here and after, if $u : X \to \mathbb{R}$ is a map and $c \in \mathbb{R}$ is a constant, then "$u \geq c$" stands for "$u(x) \geq c$ for all $x \in X$". The next proposition shows that the Non-Negativity axiom of D-diffusion stencils is equivalent (under a CFL condition) to the discrete maximum principle for the operator $\text{Id} + \delta L_\gamma$ which is associated to an explicit time step of discretized diffusion.

Definition 1.3. *In the following, an operator is a continuous map from $L^\infty(\mathbb{Z}^2)$ to itself. An operator A obeys the discrete maximum principle iff for all $u \in L^\infty(\mathbb{Z}^2)$ one has on $\mathbb{Z}^2$*

$$\inf_{\mathbb{Z}^2} u \leq Au \leq \sup_{\mathbb{Z}^2} u.$$

Proposition 1.4. *Let $\gamma : \mathbf{Z} \to \mathbb{R}$ be finitely supported. Then the following are equivalent:*

- *The weights γ are non-negative.*

- *There exists $\delta > 0$ such that $\text{Id} + \delta L_\gamma$ obeys the maximum principle.*

Assume now that the weights γ are non-negative. Then maximum principle for $\text{Id} + \delta L_\gamma$ is equivalent to the CFL condition $0 \leq \delta \sum_{e \in \mathbf{Z}} \gamma(e) \leq 1$. In addition the operator $(\text{Id} - \delta L_\gamma)^{-1}$, describing an implicit time step of discretized diffusion, obeys the maximum principle for any $\delta > 0$.

The discrete maximum principle also makes sense for static boundary value problems, without time evolution. See the discussion in [Cia70] which is based on the concept of M-matrix.

Remark 1.5 (Stencils with support outside of $\mathbf{Z}$). *It may seem relevant to consider D-diffusion stencils $\gamma : \mathbb{Z}^2 \setminus \{0\} \to \mathbb{R}$, obeying the axioms of Definition 1.1 up to the replacement of $\mathbf{Z}$ with the larger set $\mathbb{Z}^2 \setminus \{0\}$. Note however that a standard D-diffusion stencil $\tilde{\gamma} : \mathbf{Z} \to \mathbb{R}$ may then be defined by $\tilde{\gamma}(e) := \sum_{\lambda > 0} \lambda^2 \gamma(\lambda e)$. In addition $\tilde{\gamma}$ is preferable to the original γ, since it is supported on a smaller set in the sense of convex hull inclusion. Indeed $e/\gcd(e) \in [-e, e] \subseteq \text{Hull}(\text{supp}(\gamma))$ for any $e \in \text{supp}(\gamma)$.*

Discretization of Anisotropic Eikonal equations. Given a bounded domain $\Omega \subseteq \mathbb{R}^2$, and a positive definite tensor $D \in S_2^+$, one may consider the eikonal equation

$$\forall x \in \Omega, \|\nabla u(x)\|_D = 1, \qquad\qquad \forall x \in \mathbb{R}^2 \setminus \Omega, u(x) = 0. \qquad (3)$$

This PDE is known to admit a unique viscosity solution [CL83], which is the escape time $u : \mathbb{R}^2 \to \mathbb{R}_+$ from the domain Ω, where path length is measured using the metric $\|\cdot\|_M$, with $M := D^{-1}$. The reason for the matrix inversion $D = M^{-1}$ is that the norm $\|\cdot\|_M$ is intended for measuring vectors, whereas the dual norm $\|\cdot\|_D$ is intended for measuring gradients (co-vectors).

Bellman's optimality principle, illustrated in Figure 1, expresses that in order to escape the domain Ω one must first escape any sub-domain V. Hence if $x \in V \subseteq \Omega$, then

$$u(x) = \min_{y \in \partial V} \|x - y\|_M + u(y). \tag{4}$$

The next definition introduces the notion of eikonal stencil, used in Definition 1.7 to mimic Bellman's optimality principle at the discrete level.

Definition 1.6. *An M-eikonal stencil is a finite collection of points $e_1, \cdots, e_r \in \mathbf{Z}$ obeying:*

- *Ordering: the points are indexed according to the cyclic trigonometric order. Let $e_0 := e_r$.*

- *Orientation: $\det(e_i, e_{i+1}) > 0$ for all $0 \le i < r$.*

- *Acuteness: $\langle e_i, M e_{i+1} \rangle \ge 0$ for all $0 \le i < r$.*

The properties Ordering and Orientation of this definition guarantee that the union of segments $[e_0, e_1] \cup \cdots \cup [e_{r-1}, e_r]$ delimits a neighborhood V_0 of the origin. The Hopf-Lax operator, introduced below, mimics the r.h.s. of (4) at the discrete level, on the neighborhood V_0 translated around a point $x \in \mathbb{Z}^2$ of interest, and denoted by $V(x)$ in Figure 1. Hopf formulas were introduced by Evans [?] as explicit representations of some Hamilton-Jacobi equations with constant coefficients.

Definition 1.7 (Hopf-Lax operator). *Let $e_1, \cdots, e_r \in \mathbf{Z}$ obey the properties Ordering and Orientation of Definition 1.6. Then for any $u : \mathbb{Z}^2 \to \mathbb{R}$ and any $x \in \mathbb{Z}^2$ we define*

$$\Lambda u(x) := \min_{0 \le i < r} \min_{t \in [0,1]} \|x - y_{i,t}\|_M + (1-t)u(y_i) + tu(y_{i+1}), \tag{5}$$

where $y_i := x + e_i$ and $y_{i,t} := (1-t)y_i + ty_{i+1}$, for all $0 \le i < r$ and all $t \in [0,1]$.

Semi-Lagrangian schemes [FF02, BR06] for the eikonal equation take the form of a fixed point problem: find $u : \mathbb{Z}^2 \to \mathbb{R}$ such that

$$\forall x \in X, \ u(x) = \Lambda u(x), \qquad \forall x \in \mathbb{Z}^2 \setminus X, \ u(x) = 0, \tag{6}$$

where the discrete domain $X \subseteq \mathbb{Z}^2$ approximates the PDE domain Ω in (3). The eikonal PDE has a causal structure: denoting by y_* a point at which the minimum in Bellman's optimality principle (4) is attained, the PDE solution satisfies $u(x) > u(y_*)$. The discrete counterpart of this property, which may or may not hold, is introduced in the next definition.

Definition 1.8. *The Hopf-Lax operator (5) has the causality property iff, for any $u : \mathbb{Z}^2 \to \mathbb{R}$ and any $x \in \mathbb{Z}^2$, one has denoting by $0 \le i < r$ and $t \in [0,1]$ the minimizers in (5)*

$$\Lambda u(x) > u(y_i) \text{ if } t < 1, \quad \text{and} \quad \Lambda u(x) > u(y_{i+1}) \text{ if } t > 0.$$

Numerical solvers of (6) involve a mutable map $u : \mathbb{Z}^2 \to \mathbb{R} \cup \{+\infty\}$, initialized by $u = 0$ on $\mathbb{Z}^2 \setminus X$ and $u = +\infty$ on X. They iteratively substitute $u(x_i)$ with $\Lambda u(x_i)$, an operation referred to as an Hopf-Lax update [BR06], for some sequence of points $(x_i)_{i \ge 0}$ in the discrete domain X, until a convergence criterion is met. The point sequence is fixed a-priori or determined at run time.

If the causality property holds, then (6) can be efficiently solved using the Fast Marching algorithm [Tsi95], which is a variant of Dijkstra's algorithm on graphs, see Figure 1. This method
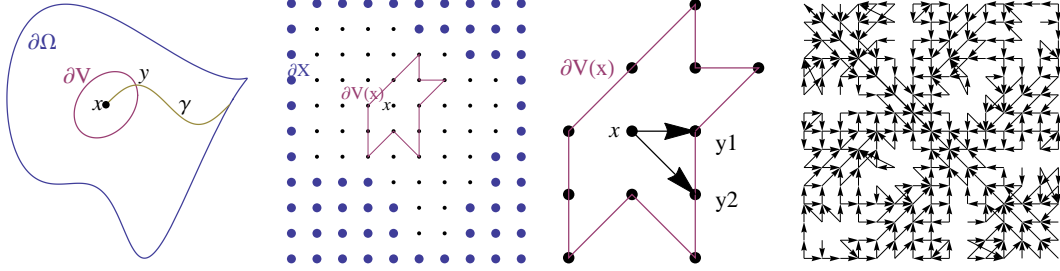
Figure 1: From left to right. I: Bellman's optimality principle (4) expresses that any path starting from x, and escaping Ω, must intersect ∂V at some point y. II: The discretization of the eikonal PDE mimics this principle, see (5) and (6). III: The minimum defining $\Lambda u(x)$ is attained on a facet of $\partial V(x)$, here a segment of endpoints y_1, y_2. The value of the solution u to (6) at x is thus sensitive to the values at y_1 and y_2. IV: The discrete *Causality property* guarantees that the solution to (6) obeys $u(x) > u(y)$, whenever the dependency graph features an arrow from x to y. This is an essential requirement for the fast marching algorithm.

determines the sequence of points $(x_i)_{i \geq 0}$ at run-time using a priority queue, and guarantees that the number of Hopf-Lax updates required at any point $x \in X$ to reach convergence is bounded a-priori by the number of elements of its stencil. Without the causality property, alternative methods must be used [BR06, **?**], which are inspired by the Bellman-Ford algorithm on graphs. They are typically slower and offer no a-priori bound on the number of Hopf-Lax updates required for each point.

The causality property, which is thus highly desirable from the algorithmic point of view, turns out to be equivalent to a geometric property of the stencils, as shown in the next proposition. This proposition can be extended to arbitrary dimension, and to arbitrary norms, which need not be of the form (1) but may be non-smooth or even asymmetric [SV03, Vla08, Mir13].

Proposition 1.9 (Causality is equivalent to Acuteness [SV03, Vla08])**.** *Under the assumptions of Definition 1.7, the causality of the Hopf-Lax operator is equivalent to the Acuteness property of the stencil in Definition 1.6.*

Remark 1.10 (Stencils with points outside **Z**)**.** *It may seem relevant to consider an M-eikonal stencil $e_1, \cdots, e_r \in \mathbb{Z}^2 \setminus \{0\}$ obeying the axioms of Definition 1.6 up to the replacement of **Z** with the larger set $\mathbb{Z}^2 \setminus \{0\}$. Note however that a standard M-eikonal stencil may then be defined by $e_1 / \gcd(e_1), \cdots, e_r / \gcd(e_r)$. In addition, the new stencil is preferable since it is supported on a smaller set in the sense of convex hull inclusion: indeed $e_i / \gcd(e_i) \in [0, e_i] \subseteq [e_1, \cdots, e_r]$.*

1.2 Minimal stencils in the sense of convex hull inclusion

We identify in this section the smallest D-diffusion and M-eikonal stencils in the sense of convex hull inclusion, see Theorems 1.12 and 1.13, for any $D, M \in S_2^+$. Empirical experience tells that the most robust and accurate PDE discretizations are typically achieved with the smallest stencils. Small stencils also limit discretization issues close to the domain boundary, and ease parallel implementations. In this light, our results characterize the best possible stencils for the discretization of anisotropic diffusion and anisotropic eikonal equations on cartesian grids - unless one is looking for discretizations with higher order consistency. The statement of our results requires the introduction of some elementary notions of lattice geometry, illustrated in Figures 2 and 3. The results presented in this section are established in §2.
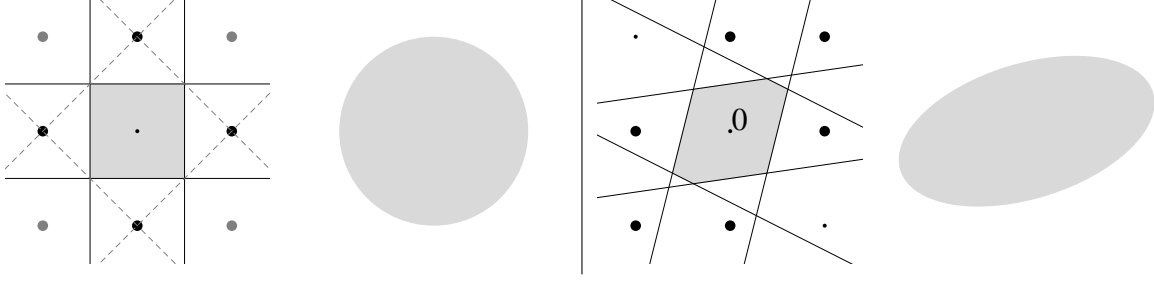
Figure 2: Voronoi cell $\mathrm{Vor}(M)$, and unit ball $\{x \in \mathbb{R}^2;\ \|x\|_M \le 1\}$, for to two matrices M. Left: $M = \mathrm{Id}$ (non generic case). Right: matrix M with condition number $\kappa = 1.4$ and orientation $\theta = 0.3$ (generic case), i.e. $M := M_\kappa(\theta)$ from Definition 1.18. Voronoi region $\mathrm{Vor}(M)$ shown in gray. Voronoi vectors shown as thick points, strict ones in black, others in gray. Lines (dashed for non-strict Voronoi vectors) mark the equality $2\langle g, Me\rangle = \|e\|_M^2$; Voronoi facets $\mathrm{Vor}(M; e)$ are their intersections with the grayed Voronoi region $\mathrm{Vor}(M)$.

Definition 1.11. *For each matrix $M \in S_2^+$, the Voronoi cell $\mathrm{Vor}(M)$, and facets $\mathrm{Vor}(M; e)$, $e \in \mathbb{Z}^2 \setminus \{0\}$, are defined by*

$$\mathrm{Vor}(M) := \{p \in \mathbb{R}^2;\ \forall e \in \mathbb{Z}^2,\ \|p\|_M \le \|p - e\|_M\},$$
$$\mathrm{Vor}(M; e) := \{p \in \mathrm{Vor}(M);\ \|p\|_M = \|p - e\|_M\}.$$

An M-Voronoi vector is an element $e \in \mathbb{Z}^2 \setminus \{0\}$ such that $\mathrm{Vor}(M; e) \ne \emptyset$. It is said strict iff $\mathrm{Vor}(M; e)$ is not reduced to a point, in which case it is a 1-dimensional segment.

Note that $\|p\|_M \le \|p - e\|_M$ if and only if $2\langle p, Me\rangle \le \|e\|_M^2$, for any $p, e \in \mathbb{R}^2$. Hence $\mathrm{Vor}(M)$ is an intersection of half-spaces, thus a convex polytope, which has one facet $\mathrm{Vor}(M; e)$ with exterior normal is Me for each strict M-Voronoi vector e. The next two theorems show that the smallest D-diffusion and M-eikonal stencils are supported precisely on the set of strict M-Voronoi vectors, with $D := M^{-1}$.

Theorem 1.12. *For any $D \in S_2^+$, there exists a unique D-diffusion stencil supported on the set V of strict M-Voronoi vectors, where $M := D^{-1}$. Furthermore, this stencil is minimal in the following sense: any other D-diffusion stencil γ satisfies $\mathrm{Hull}(V) \subseteq \mathrm{Hull}(\mathrm{supp}(\gamma))$.*

Theorem 1.13. *For any $M \in S_2^+$, the set V of strict M-Voronoi vectors, ordered trigonometrically, is an M-eikonal stencil. Furthermore this stencil is minimal in the following sense: any other M-eikonal stencil $e_1, \cdots, e_r$ satisfies $\mathrm{Hull}(V) \subseteq \mathrm{Hull}(\{e_1, \cdots, e_r\})$.*

The next two definitions distinguish, within the set S_2^+ of symmetric positive definite matrices, an open dense subset of *generic* matrices, and a complementary subset of exceptional matrices obeying a geometric property. This allows in Proposition 1.16 to characterize strict M-Voronoi vectors as the smallest elements of $\mathbf{Z}$ w.r.t. the metric $\|\cdot\|_M$, and in Proposition 1.17 to relate our stencil constructions with earlier works.

Definition 1.14. *A basis of $\mathbb{Z}^2$ is a pair $(f, g) \in (\mathbb{Z}^2)^2$ such that $|\det(f, g)| = 1$. The basis is said M-orthogonal, where $M \in S_2^+$, iff $\langle f, Mg\rangle = 0$.*

Definition 1.15. *A matrix $M \in S_2^+$ is said generic iff there exists no M-orthogonal basis of $\mathbb{Z}^2$.*

Proposition 1.16. *The set of strict M-Voronoi vectors, where $M \in S_2^+$, consists of*
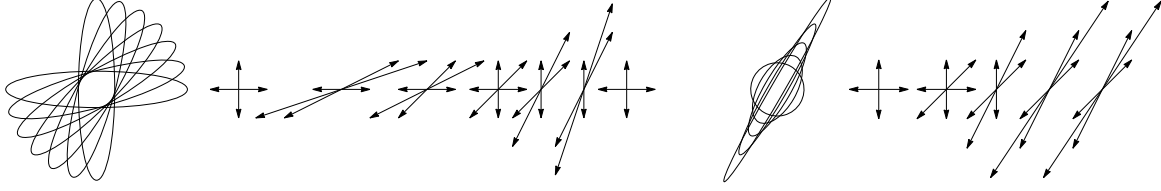
7

Figure 3: Unit circle $\{x \in \mathbb{R}^2; \|x\|_M = 1\}$, and collection of strict M-Voronoi vectors, for matrices $M = M_\kappa(\theta)$ of varying condition number κ and orientation θ, see Definition 1.18. Left: $\kappa = 5$, $\theta \in [0, \pi/2]$. Right: $\kappa \in [1, 16]$, $\theta = \pi/3$).

- *if M is non-generic: the four smallest elements of $\mathbf{Z}$ w.r.t. $\| \cdot \|_M$.*

- *if M is generic: the six smallest elements of $\mathbf{Z}$ w.r.t. $\| \cdot \|_M$.*

As far as generic matrices are concerned, the stencil constructions of Theorems 1.12 and 1.13 are not new, but coincide with those of [FM13] for diffusion and [Mir14] for eikonal PDEs. We refer to these works for convergence results, numerical experiments and comparisons with other approaches. The stencil constructions of [FM13, Mir14] are based on another tool of lattice geometry, called obtuse superbases see §2.2, and also make sense in three dimensions. However no optimality result is known in this latter case.

Proposition 1.17. *If $M \in S_2^+$ is generic, then the D-diffusion stencil of Theorem 1.12 is the one presented in [FM13], and the M-eikonal stencil of Theorem 1.13 is the one presented in [Mir14].*

Our D-diffusion stencils are also equivalent to the (two dimensional only) construction of [BOZ04]. Non-equivalent constructions can be found in [AM11] for M-eikonal stencils, and [Wei98] for D-diffusion stencils. They are, inevitably, less local and accurate, as exposed by the comparisons in [FM13, Mir14].

1.3 Average size of minimal stencils

We introduced in §1.1 the concepts of D-diffusion stencil and M-eikonal stencil, and identified in §1.2 the smallest such stencils in the sense of convex hull inclusion. They turn out to be both supported on the set of strict M-Voronoi vectors, with $M := D^{-1}$. Our next result, Theorem 1.19 is a quantitative estimate of the size these stencils, which supplements their qualitative optimality property. The result presented in this section is established in §3.

Definition 1.18. *For each $\kappa \in [1, \infty[$, $\theta \in \mathbb{R}$, we introduce the positive definite 2×2 matrix*

$$M_\kappa(\theta) := \kappa^{-1} e_\theta \otimes e_\theta + \kappa \, e_\theta^\perp \otimes e_\theta^\perp, \qquad with \; e_\theta := (\cos\theta, \sin\theta). \tag{7}$$

We denote by $V_\kappa(\theta)$ the collection of all strict $M_\kappa(\theta)$-Voronoi vectors, and define

$$R_\kappa(\theta) := \max_{e \in V_\kappa(\theta)} \|e\|, \qquad\qquad S_\kappa(\theta) := \max_{e \in V_\kappa(\theta)} \|e\|_{M_\kappa(\theta)}. \tag{8}$$

We allow ourselves a slight abuse of notation: "e_θ" always refers to the unit vector (7) of direction θ, although at times we introduce families $e_1, \cdots, e_r$ of vectors in $\mathbb{Z}^2$.

The radius of M-diffusion stencils is measured in (8) using two different norms: the "extrinsic" euclidean norm (8, left), and the "intrinsic" $\| \cdot \|_{M_\kappa(\theta)}$-norm (8, right) which is tied to the
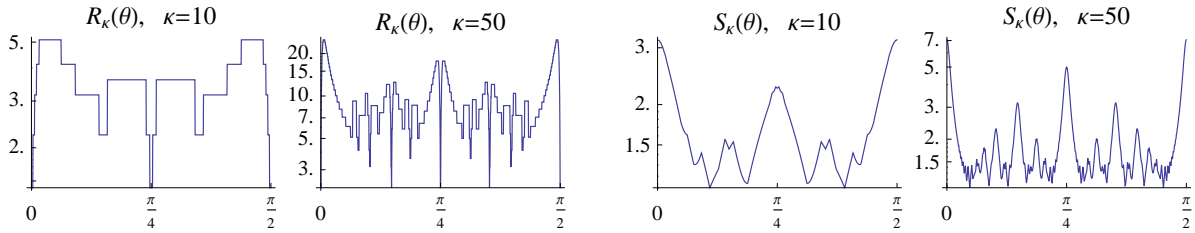
8

Figure 4: Plots of $R_\kappa(\theta)$ (left) and $S_\kappa(\theta)$ (right), for $\kappa \in \{10, 50\}$, as a function of $\theta \in [0, \pi/2]$, see Definition 1.18. Logarithmic scale.

anisotropic geometry embedded in the PDE. Several plots of $R_\kappa(\theta)$ and $S_\kappa(\theta)$, as a function of $\theta \in [0, \pi/2[$ and for a fixed κ, are shown in Figure 4. The dependence of $R_\kappa(\theta)$ and $S_\kappa(\theta)$ on the variable θ is highly irregular, and reflects how well the slope $\tan\theta$ is approximated by rationals with small denominator. Both radii measures R_κ and S_κ have their relevance and merits, as discussed in Appendix B of [Mir14] with an heuristic analysis of the accuracy of several discretizations of the eikonal PDE.

The maximum condition number κ, of the tensor field associated to an anisotropic PDE, is generally well known: it is problem data, reflecting a continuous model. In contrast we regard the angle θ, between the cartesian grid axes and the tensor eigenvectors, as a random quantity uniformly distributed in $[0, \pi[$. In other words the preferred directions of the grid and of the PDE are viewed as independent of each other. Our main result, Theorem 1.19 illustrated in Figure 5, is thus devoted to the estimation of averaged quantities w.r.t. the angular variable θ, namely the $L^p([0, \pi[)$-norms of R_κ and S_κ. The uniform equivalence of two expressions, w.r.t. the parameter κ, is denoted as follows:

$$A(\kappa) \approx B(\kappa) \quad \Leftrightarrow \quad \exists C, c > 0, \ \forall \kappa \geq 1, \ cB(\kappa) \leq A(\kappa) \leq CB(\kappa). \tag{9}$$

Theorem 1.19. *For any $p \in [1, \infty]$ one has uniformly in κ (L^p norms are on the interval $[0, \pi[$)*

$$\|R_\kappa\|_{L^p} \approx \kappa^{\frac{1}{2}} \|S_\kappa\|_{L^p}, \qquad \|S_\kappa\|_{L^p} \approx \begin{cases} \kappa^{\frac{1}{2} - \frac{1}{p}} & \text{if } p > 2, \\ (\ln\kappa)^{\frac{1}{2}} & \text{if } p = 2, \\ 1 & \text{if } p < 2. \end{cases}$$

In applications, $\kappa = 10$ is already a pronounced anisotropy, and $\kappa = 100$ is presumably the most degenerate anisotropy that can conceivably be handled with the proposed discretizations of the diffusion and eikonal PDEs. More specialized approaches are recommended for stronger anisotropies, such as asymptotic preserving formulations [**?**]. Nevertheless Theorem 1.19 is a good indicator of the effective spread of our diffusion stencils, since the asymptotic behavior is quickly attained as shown in Figure 5.

1.4 PDEs with non-constant coefficients

We shortly discuss some PDE models interest, which are intrinsically anisotropic, often discretized on cartesian grids, and thus directly benefit from the results obtained in this paper.

Divergence form anisotropic diffusion. Let $\Omega \subseteq \mathbb{R}^2$ be a bounded domain, and let $\mathbf{D} : \overline{\Omega} \to S_2^+$ be a continuous field of diffusion tensors. Let $h > 0$ be a discretization scale, and let $\Omega_h := \Omega \cap h\mathbb{Z}^2$ be the discretization grid. The elliptic energy $\mathcal{E}_\mathbf{D}$ associated to $\mathbf{D}$ has a natural
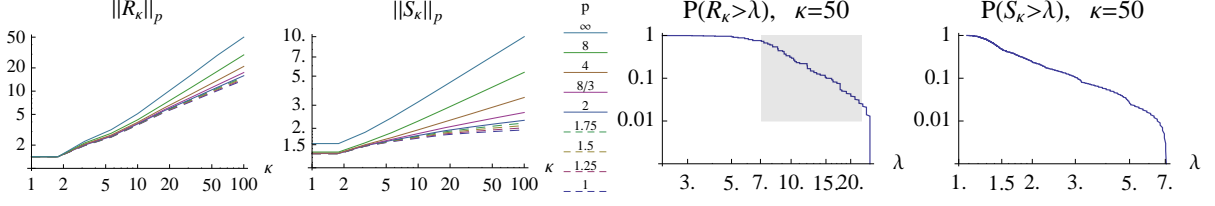
9

Figure 5: Left: $L^p([0, \pi[)$ norms of R_κ and S_κ, as a function of $\kappa \in [1, 100]$ for different exponents $p \in [1, \infty]$. The behavior for $p < 2$ (dashed) and $p > 2$ is significantly different, see Theorem 1.19. Right: Tail distribution of R_κ and S_κ. Log-log scale.

discretization $\mathcal{E}_h$. For $u : \overline{\Omega} \to \mathbb{R}$

$$\mathcal{E}_{\mathbf{D}}(u) := \int_\Omega \|\nabla u(x)\|^2_{\mathbf{D}(x)} dx, \qquad \mathcal{E}_h(u) := \frac{1}{2} \sum_{x \in \Omega_h} \sum_{e \in \mathbf{Z}} \gamma_{\mathbf{D}(x)}(u(x + he) - u(x))^2.$$

We denoted by $\gamma_D : \mathbf{Z} \to \mathbb{R}$, where $D \in S_2^+$, the minimal D-diffusion stencil of Theorem 1.12. The contributions to $\mathcal{E}_h$ involving the evaluation of u outside $\overline{\Omega}$ should be omitted for Neumann boundary conditions, or included by extending u as 0 on $\mathbb{R}^2 \setminus \Omega$ for Dirichlet boundary conditions. Taking the gradient flow of $\mathcal{E}_{\mathbf{D}}$ (resp. $\mathcal{E}_h$) w.r.t. the L^2 metric, one obtains the divergence form diffusion $\partial_t u(t, x) = \mathrm{div}(\mathbf{D}(x)\nabla u(t, x))$ (resp. a natural discretization of this PDE obeying the maximum principle). Before discussing specific applications, we illustrate the relevance of Theorem 1.19 through an heuristic accuracy analysis. By a fourth order Taylor expansion of u close to x, we obtain

$$\frac{1}{2h^2} \sum_{e \in V} \gamma_D(u(x + he) - u(x))^2 = \|\nabla u(x)\|^2_D + h^2 \sum_{e \in V} \gamma_D \Lambda(e, e, e, e) + \mathcal{O}(h^4), \qquad (10)$$

with the four-linear tensor $\Lambda = \frac{1}{2}\nabla^2 u(x) \otimes \nabla^2 u(x) + \nabla u(x) \otimes \nabla^3 u(x)$. The second order $\mathcal{O}(h^2)$ contribution is thus bounded by the three term product,

$$\sum_{e \in \mathbf{Z}} \gamma_D(e)\|\Lambda\|\|e\|^4 \le \|\Lambda\| \left(\sum_{e \in \mathbf{Z}} \gamma_D(e)\|e\|^2 \right) \left(\max_{e \in \mathrm{supp}(\gamma_D)} \|e\|^2 \right) = \|\Lambda\| \, \mathrm{Tr}(D) r(D)^2.$$

involving a contribution $\|\Lambda\|$ of the PDE solution regularity, a contribution $\mathrm{Tr}(D)$ of the PDE data, and a contribution $r(D)^2$ from the scheme, where $r(D) := \max\{\|e\|; e \in \mathrm{supp}(\gamma_D)\}$ is the radius of the stencil support. Therefore, the difference between the elliptic energy $\mathcal{E}_{\mathbf{D}}$ and its discretization $\mathcal{E}_h$ is bounded by

$$|\mathcal{E}_{\mathbf{D}}(u) - \mathcal{E}_h(u)| \le \left| \int_\Omega \|\nabla u(x)\|^2_{\mathbf{D}(x)} dx - h^2 \sum_{x \in \Omega_h} \|\nabla u(x)\|^2_{\mathbf{D}(x)} \right|$$

$$+ h^2 \left(\max_{x \in \Omega} \|\Lambda(x)\|\|\mathbf{D}(x)\| \right) \left(h^2 \sum_{x \in \Omega_h} r(\mathbf{D}(x))^2 \right) + o(h^2).$$

The first contribution, a *quadrature* error, is scheme independent and expected to be $\mathcal{O}(h^2)$ (at least on square domains, for which the boundary representation is exact). The second contribution, an $\mathcal{O}(h^2)$ *discretization* error, is proportional to the average squared stencil radius. This

10

shows the relevance of the estimate $\|R_\kappa\|_{L^2} \approx \sqrt{\kappa \ln \kappa}$ obtained in Theorem 1.19. Note that the worst case stencil radius is significantly larger: $\|R_\kappa\|_{L^\infty} \approx \kappa$.

Non-linear anisotropic diffusion PDEs take the form $\partial_t u(t,x) = \text{div}(\mathbf{D}_u(t,x)\nabla u(t,x))$, with a solution dependent tensor field $\mathbf{D}_u$. They play an important role in image processing, being used for image deblurring, sharpening, compression, ... see [Wei96] for an overview. From a practical point of view, the PDE non-linearity merely requires to update the diffusion tensor field every few time steps. Our D-diffusion stencils are used to solve numerically non-linear anisotropic diffusion PDEs in [FM13]. Note that our results are irrelevant for *non-linear isotropic* diffusion, such as the Perona-Malik model[1] with diffusion tensors $\mathbf{D}_u(t,x) := \text{Id}/(1 + \|\nabla u(t,x)\|)$.

We end this paragraph by mentioning a celebrated model of the first layer of the visual cortex [Pet03] as the three dimensional manifold $V_1 := \mathbb{R}^2 \times \mathbb{P}^1$, where $\mathbb{P}^1$ denotes $[0,\pi[$ equipped with periodic boundary conditions. This manifold is equipped with a degenerate hypo-elliptic diffusion operator, which can be approximated using a strongly anisotropic diffusion tensor field: for all $(x_0, x_1, \theta) \in V_1$, denoting by $\kappa \gg 1$ a large parameter

$$\mathbf{D}(x_0, x_1, \theta) = \left(\begin{array}{c|c} \begin{matrix} e_\theta \otimes e_\theta + \\ \kappa^{-2} e_\theta^\perp \otimes e_\theta^\perp \end{matrix} & \\ \hline & 1 \end{array} \right) = \left(\begin{array}{c|c} (\kappa M_\kappa(\theta))^{-1} & \\ \hline & 1 \end{array} \right).$$

The two-dimensional discretization scheme proposed in this paper for anisotropic diffusion easily extend to this case, thanks to the block diagonal structure. This model fits Theorem 1.19 particularly well, since the condition number κ is fixed, and all directions θ are equally represented.

Non-Divergence form diffusion. Monotone second-order differential operators can, under mild assumptions, be expressed in Hamilton-Jacobi-Bellman (HJB) form as infs and sups of linear non-divergence form diffusion operators. Denoting by A and B two parameter spaces, and by $\mathbf{D} : A \times B \to S_2^+$ a family of diffusion tensors, the continuous operator and its discretization read:

$$\inf_{a \in A} \sup_{b \in B} \text{Tr}(\mathbf{D}(a,b)\nabla^2 u(x)), \qquad \inf_{a \in A} \sup_{b \in B} \sum_{e \in \mathbf{Z}} \gamma_{\mathbf{D}(a,b)}(e) \frac{u(x+he) - u(x)}{h^2}.$$

This strategy is applied in [BOZ04] to the HJB equation of stochastic control. See also [BCM15] for a monotone and consistent discretization of the Monge-Ampere operator, which fits in this setting thanks to the identity $\det(\nabla^2 u) = \inf\{\text{Tr}(D\nabla^2 u); D \in S_2^+, \det D = 1\}$. Note that the constraint $\det(D) = 1$ includes diffusion tensors $D \in S_2^+$ of arbitrary condition number, in all orientations.

Anisotropic Eikonal equations. Anisotropic eikonal equations allow to estimate riemannian distances, and compute the associated shortest paths. Anisotropy is particularly useful in applications to image segmentation, where it is used to guide the shortest paths along one dimensional structures of interest, such as blood vessels or organ boundaries [CCM14].

Strongly anisotropic riemannian metrics can also be used to approximate degenerate, subriemannian geometries. For instance, the Reed-Shepp car model is posed on the manifold $\mathbb{R}^2 \times \mathbb{P}^1$, where again $\mathbb{P}^1$ denotes $[0,\pi[$ equipped with periodic boundary conditions. This model involves a degenerate metric which can be approximated as follows: for all $(x_0, x_1, \theta) \in \mathbb{R}^2 \times \mathbb{P}^1$, denoting by $\kappa \gg 1$ a large parameter

$$M(x_0, x_1, \theta) = \left(\begin{array}{c|c} \begin{matrix} e_\theta \otimes e_\theta + \\ \kappa^2 e_\theta^\perp \otimes e_\theta^\perp \end{matrix} & \\ \hline & 1 \end{array} \right) = \left(\begin{array}{c|c} \kappa M_\kappa(\theta) & \\ \hline & 1 \end{array} \right).$$

---

[1]Note that the Perona-Malik model is often incorrectly referred to as "anisotropic diffusion", despite the fact that its diffusion tensors are proportional to the identity matrix. See [Wei96] for a discussion on terminology.
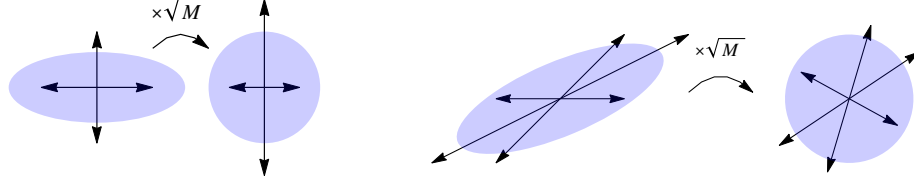
Figure 6: Unit ball $\{e \in \mathbb{R}^2; \|e\|_M \leq 1\}$ and collection of strict M-Voronoi vectors; same graphic after a linear change of coordinates by $M^{\frac{1}{2}}$. Left: diagonal matrix, Right: generic matrix. The angles of trigonometrically consecutive strict M-Voronoi vectors f, g are acute after the change of coordinates, as expected since $\langle f, Mg \rangle \geq 0$ by Corollary 2.3.

The two-dimensional numerical scheme proposed in this paper for anisotropic eikonal equations easily extends to this model, thanks to the block diagonal structure of the metric. See [SBD$^+$15] for an application to blood vessel segmentation. This model fits Theorem 1.19 particularly well, since the condition number is fixed, and all anisotropy orientations are equally represented.

2   Correctness and minimality of Voronoi based stencils

We establish in this section the results announced in §1.2. Specifically, Theorem 1.13 is proved in §2.1, Proposition 1.16 in §2.2, and Theorem 1.12 in §2.3. The results of each subsection are used in the following subsection, and in particular Theorem 1.13 is used to prove Theorem 1.12. The statement of Proposition 1.17 on eikonal stencils (resp. diffusion stencils) is proved in §2.2 (resp. §2.3). We denote by $\mathrm{Cone}(f, g)$ the convex cone spanned by two vectors $f, g \in \mathbb{R}^2$.

2.1   Correctness and minimality of eikonal stencils

We establish in Corollary 2.3 the correctness part of Theorem 1.13: strict M-Voronoi vectors, when ordered trigonometrically, do define an M-eikonal stencil. The rest of Theorem 1.13 is established in Proposition 2.6. Two preliminary lemmas are required.

Lemma 2.1. *All M-Voronoi vectors belong to the set $\mathbf{Z}$, for all $M \in S_2^+$.*

Proof. Any element of $\mathbb{Z}^2 \setminus \{0\}$ can be written under the form λe, where λ is a positive integer and $e \in \mathbf{Z}$. If $\lambda > 1$, then for any $p \in \mathbb{R}^2$ we obtain

$$\|e - p\|^2 = \frac{1}{\lambda}\|\lambda e - p\|^2 + \left(1 - \frac{1}{\lambda}\right)\|p\|^2 - (\lambda - 1)\|e\|^2 < \max\{\|\lambda e - p\|, \|p\|^2\},$$

which shows that $\mathrm{Vor}(M; \lambda e) = \emptyset$, hence that λe is not a Voronoi vector.   □

Lemma 2.2. *Let f, g be M-Voronoi vectors such that $\mathrm{Vor}(M; f) \cap \mathrm{Vor}(M; g) \neq \emptyset$, where $M \in S_2^+$. Then $\langle f, Mg \rangle \geq 0$.*

Proof. Consider $x \in \mathrm{Vor}(M; f) \cap \mathrm{Vor}(M; g)$. Then $\|x - (f + g)\|_M \geq \|x - f\|_M = \|x - g\|_M = \|x\|_M$. Thus, as announced

$$0 \leq \|x - (f + g)\|_M^2 - \|x - f\|_M^2 - \|x - g\|_M^2 + \|x\|_M^2 = 2\langle f, Mg \rangle.   □$$

Corollary 2.3. *For any $M \in S_2^+$, the collection of strict M-Voronoi vectors $(e_1, \cdots, e_r)$, ordered trigonometrically, defines an M-eikonal stencil.*

12

Proof. By Lemma 2.1, M-Voronoi vectors belongs to $\mathbf{Z}$ as required. The sets $\mathrm{Vor}(M; e_i)$, $1 \le i \le r$, are the faces of the convex polytope $\mathrm{Vor}(M)$, ordered trigonometrically. The exterior normal to $\mathrm{Vor}(M; e_i)$ is Me_i, hence $0 < \det(Me_i, Me_{i+1}) = \det(M)\det(e_i, e_{i+1})$ as required (Orientation), for all $1 \le i \le r$. Since two consecutive faces have non-empty intersection, one has $\langle e_i, Me_{i+1}\rangle$ by Lemma 2.2, as required also (Acuteness), which concludes the proof. $\qquad\square$

Our next objective is to establish the minimality property of the M-eikonal stencils of Theorem 1.13. For that purpose, achieved in Proposition 2.6, two preliminary results are required. Recall that $[p_1, \cdots, p_n] := \mathrm{Hull}(\{p_1, \cdots, p_n\})$ for any $p_1, \cdots, p_n \in \mathbb{R}^2$.

Lemma 2.4. *Let $f, g \in \mathbb{Z}^2$ be such that $|\det(f, g)| > 1$. Then the triangle $[0, f, g]$ contains a point $e \in \mathbf{Z}$ distinct from its vertices.*

Proof. Since $|\det(f, g)| > 1$, the map $(\alpha, \beta) \in \mathbb{Z}^2 \mapsto \alpha f + \beta g \in \mathbb{Z}^2$ is not surjective. Hence there exists $(\alpha, \beta) \in \mathbb{Q}^2$, at least one of them non-integer, such that $\alpha f + \beta g \in \mathbb{Z}^2$. Up to replacing (α, β) with $(\alpha - m, \beta - n)$, for some $(m, n) \in \mathbb{Z}^2$, we may assume that $\alpha, \beta \in [0, 1]$. Up to replacing (α, β) with $(1 - \alpha, 1 - \beta)$, we may assume that $\alpha + \beta \le 1$. The point $e := \alpha f + \beta g \in \mathbb{Z}^2$ thus belongs to T and is distinct from its vertices. This establishes the announced result, unless $\gcd(e) > 1$, in which case we can consider $e/\gcd(e)$. $\qquad\square$

Corollary 2.5. *Let $(e_1, \cdots, e_r)$ be an M-eikonal stencil. Then there exists another M-eikonal stencil $(e_1', \cdots, e_{r'}')$ such that $[e_1', \cdots, e_{r'}'] \subseteq [e_1, \cdots, e_r]$ and $\det(e_i', e_{i+1}') = 1$ for all $0 \le i < r'$.*

Proof. Define $(e_1', \cdots, e_{r'}')$ by ordering trigonometrically the elements of $[e_1, \cdots, e_r] \cap \mathbf{Z}$. By construction, the points $e_1, \cdots, e_r$ are among $e_1', \cdots, e_{r'}'$, hence for any $0 \le i < r'$ there exists $0 \le j < r$ such that the vectors $e_j, e_i', e_{i+1}', e_{j+1}$ are in trigonometric order. This implies $\langle e_i', Me_{i+1}'\rangle > 0$ (Acuteness) and $\det(e_i', e_{i+1}') > 0$ (Orientation) as required. By construction, the triangle $[0, e_i', e_{i+1}']$ contains no point of $\mathbf{Z}$ aside from its vertices, hence $|\det(e_i', e_{i+1}')| \le 1$ by Lemma 2.4, thus $\det(e_i', e_{i+1}') = 1$ as announced since this determinant is a positive integer. $\qquad\square$

Proposition 2.6. *Let $M \in S_2^+$, and let $(e_1, \cdots, e_r)$ be an M-eikonal stencil. Then all strict M-Voronoi vectors belong to $[e_1, \cdots, e_r]$.*

Proof. Let e be a strict M-Voronoi vector, and let $f = e_i$ and $g = e_{i+1}$ be trigonometrically consecutive elements of the M-eikonal stencil $(e_1, \cdots, e_r)$ such that $e \in \mathrm{Cone}(f, g)$. Then $e = \alpha f + \beta g$ for some $\alpha, \beta \in \mathbb{R}_+$. By Corollary 2.5 we can assume without loss of generality that $\det(f, g) = 1$, thus the vectors (f, g) form a basis of $\mathbb{Z}^2$, hence $\alpha, \beta \in \mathbb{Z}$. Since $e \in \mathbf{Z}$, one has $\gcd(\alpha, \beta) = 1$. Assuming for contradiction that $\alpha \ge 1$ and $\beta \ge 1$ we obtain for any $p \in \mathbb{R}^2$

$$\alpha\|p - f\|^2 + \beta\|p - g\|^2 = (\alpha + \beta)\|p\|^2 - 2\langle \alpha f + \beta g, p\rangle + \alpha\|f\|^2 + \beta\|g\|^2$$
$$= (\alpha + \beta - 1)\|p\|^2 + \|p - e\|^2 - 2\alpha\beta\langle f, g\rangle.$$

Since $\alpha\beta\langle f, g\rangle \ge 0$, and since $\alpha + \beta = (\alpha + \beta - 1) + 1$, it follows that $\min\{\|p - f\|, \|p - g\|\} \le \max\{\|p\|, \|p - e\|\}$. Choosing p in the relative interior of $\mathrm{Vor}(M; e)$, so that the elements of $\mathbb{Z}^2$ closest to p are precisely 0 and e, we obtain a contradiction. Thus (α, β) equals $(1, 0)$ or $(0, 1)$, therefore e equals f or g, which shows as announced that the Voronoi vector e belongs to the stencil convex hull $[e_1, \cdots, e_r]$. $\qquad\square$

2.2 Description of strict M-Voronoi vectors

In this section, we introduce the concept of *superbase* of the lattice $\mathbb{Z}^2$. This allows us to establish the correspondence between the M-eikonal stencil of Theorem 1.13 and the one constructed in [Mir14], as announced in Proposition 1.17. We also characterize strict M-Voronoi vectors as the smallest elements of $\mathbf{Z}$ w.r.t. $\|\cdot\|_M$, as announced in Proposition 1.16.

Definition 2.7. *A superbase of $\mathbb{Z}^2$ is a triplet (e_0, e_1, e_2) such that (e_1, e_2) is a basis of $\mathbb{Z}^2$ and $e_0 + e_1 + e_2 = 0$. It is said M-obtuse, where $M \in S_2^+$, iff $\langle e_i, M e_j \rangle \leq 0$ for all $0 \leq i < j \leq 2$.*

It is known that for each $M \in S_2^+$ there exists at least one M-obtuse superbase. Interestingly, a similar property holds in dimension 3, but not in dimension 4 or higher. The proof is constructive, and follows from the analysis of an algorithm by Selling [Sel74]. Obtuse superbases are an important tool in lattice classification [CS92]. They are also at the foundation of the M-eikonal stencil of [Mir14], described in Proposition 2.8 (ii), and of the D-diffusion stencil of [FM13], defined in Corollary 2.11. The proof of the following proposition is left as an easy exercise.

Proposition 2.8. *(i) Let $M \in S_2^+$ be non-generic, and let (e_1, e_2) be an M-orthogonal basis, sorted so that $\det(e_1, e_2) = 1$. Then $(e_1, e_2, -e_1, -e_2)$ is an M-eikonal stencil.*

(ii) Let $M \in S_2^+$, and let (e_0, e_1, e_2) be an M-obtuse superbase, sorted so that $\det(e_1, e_2) = 1$. Then $(e_0, -e_2, e_1, -e_0, e_2, -e_1)$ is an M-eikonal stencil.

Proposition 2.9. *Let $M \in S_2^+$ be non-generic and let $n = 4$ (resp. generic and let $n = 6$). Then the set of strict M-Voronoi vectors precisely consists of the n members $(f_1, \cdots, f_n)$ of the M-eikonal stencil constructed in Proposition 2.8.*

Proof. We first observe that, if (e_1, e_2) is a basis of $\mathbb{Z}^2$ (resp. (e_0, e_1, e_2) is a superbase), then $[e_1, e_2, -e_1, -e_2] \cap \mathbf{Z} = \{e_1, e_2, -e_1, -e_2\}$ (resp. $[e_0, -e_2, e_1, -e_0, e_2, -e_1] \cap \mathbf{Z} = \{e_0, -e_2, e_1, -e_0, e_2, -e_1\}$. Indeed, by linear invariance, it suffices to check the case where $e_1 = (1, 0)$ and $e_2 = (0, 1)$ (resp. and $e_0 = (-1, -1)$), which is obvious.

By Theorem 1.13, strict M-Voronoi vectors form an eikonal stencil, which is minimal for convex hull inclusion hence must be a subset of $[f_1, \cdots, f_n]$. Thus the set V of M-Voronoi vectors is included in $[f_1, \cdots, f_n] \cap \mathbf{Z} = \{f_1, \cdots, f_n\}$. Since $\langle f_{i-1}, M f_{i+1} \rangle < 0$ for all $0 \leq i < n$, and in view of the Acuteness property of the M-eikonal stencil built of the strict M-Voronoi vectors, this inclusion is an equality, which concludes the proof. $\square$

We finally characterize strict M-Voronoi vectors as announced in Proposition 1.16.

Proof of Proposition 1.16. Let $f_1, \cdots, f_n$ be the M-eikonal stencil of Proposition 2.8, which by Proposition 2.9 is also the set of strict M-Voronoi vectors. Let $e \in \mathbf{Z} \setminus \{f_1, \cdots, f_n\}$, and let $0 \leq i < n$ be such that $e \in \mathrm{Cone}(f, g)$ with $f := f_i$ and $g := f_{i+1}$. Then $e = \alpha f + \beta f$ for some coefficients α, β which must be integer (since (f, g) is a basis), non-negative (since $e \in \mathrm{Cone}(f, g)$), co-prime (since $e \in \mathbf{Z}$), and obey $|\alpha| + |\beta| > 1$ (since $e \notin \{f, g\}$). Thus $\alpha \geq 1$ and $\beta \geq 1$, and therefore $\|e\|_M \geq \|\alpha f + \beta g\|_M \geq \|f + g\|_M$ since $\langle f, M g \rangle \geq 0$.

In the non-generic case, this implies $\|e\|_M \geq \|e_1 + e_2\|_M = \sqrt{\|e_1\|_M^2 + \|e_2\|_M^2} > \max\{\|e_1\|_M, \|e_2\|_M\}$, where (e_1, e_2) is an M-orthogonal basis. In the generic case this implies $\|e\|_M \geq \|e_i - e_j\|_M$ for some $0 \leq i < j \leq 2$, where (e_0, e_1, e_2) is an M-obtuse superbase. Since $\langle e_i, M e_j \rangle < 0$, we obtain $\|e\|_M \geq \|e_i - e_j\|_M > \max\{\|e_i\|_M, \|e_j\|_M, \|e_i + e_j\|_M\} = \max\{\|e_0\|_M, \|e_1\|_M, \|e_2\|_M\}$. We have proved in both cases that $\|e\|_M > \max\{\|f_1\|_M, \cdots, \|f_n\|_M\}$ for all $e \in \mathbf{Z} \setminus \{f_1, \cdots, f_n\}$, which concludes the proof. $\square$

2.3 Correctness and minimality of diffusion stencils

We establish in Corollary 2.11 that there exists a D-diffusion stencil supported on the set of strict M-Voronoi vectors, as announced in Theorem 1.12. Its minimality, in the sense of convex hull inclusion, is established in Corollary 2.13. Our first step is an elementary algebraic lemma.

Proposition 2.10. *Let $(e_i)_{i=0}^2$ be a superbase. Then $(e_i^\perp \otimes e_i^\perp)_{i=0}^2$ is a basis of S_2, and for any $D \in S_2$ one has*

$$D = -\sum_{0 \le i \le 2} \langle e_{i+1}, D e_{i+2} \rangle \, e_i^\perp \otimes e_i^\perp. \tag{11}$$

Proof. For any $i, j \in \{0, 1, 2\}$, one has denoting by δ_{ij} the Kronecker symbol

$$\langle e_j, e_i^\perp \otimes e_i^\perp e_j \rangle = \langle e_i^\perp, e_j \rangle^2 = \det(e_i, e_j)^2 = 1 - \delta_{ij}.$$

Therefore denoting by $\tilde{D}$ the r.h.s. of (11) we obtain for any $0 \le j \le 2$

$$\langle e_j, \tilde{D} e_j \rangle = -\langle e_j, D e_{j+1} \rangle - \langle e_{j-1}, D e_j \rangle = -\langle e_{j+1} + e_{j-1}, D e_j \rangle = \langle e_j, D e_j \rangle.$$

In particular $0 = \langle e_1, (D - \tilde{D})e_1 \rangle = \langle e_2, (D - \tilde{D})e_2 \rangle = \langle (e_1 + e_2), (D - \tilde{D})(e_1 + e_2) \rangle$. Since (e_1, e_2) is a basis, this implies $D = \tilde{D}$ as announced. As a result the three element family $(e_i^\perp \otimes e_i^\perp)_{i=0}^2$ generates the three dimensional space S_2, hence it is a basis. $\square$

The next corollary describes the D-diffusion stencil used in [FM13], and shows that it is supported on strict M-Voronoi vectors, with $M := D^{-1}$, as announced in Theorem 1.12.

Corollary 2.11. *Let (e_0, e_1, e_2) be a D-obtuse superbase. Then defining $\gamma : \mathbf{Z} \to \mathbb{R}$ by $\gamma(\pm e_i^\perp) := -\langle e_{i+1}, D e_{i+2} \rangle$, $0 \le i \le 2$, and $\gamma = 0$ on $\mathbf{Z} \setminus \{\pm e_0^\perp, \pm e_1^\perp, \pm e_2^\perp\}$, we obtain the unique D-diffusion stencil which is supported on the set of strict M-Voronoi vectors, with $M := D^{-1}$.*

Proof that γ defines a D-diffusion stencil. Since (e_0, e_1, e_2) is a superbase one has $\{\pm e_1^\perp, \pm e_2^\perp, \pm e_3^\perp\} \subseteq \mathbf{Z}$, hence that $\gamma : \mathbf{Z} \to \mathbb{R}$ is well defined. Since it is D-obtuse one has $\langle e_{i+1}, D e_{i+2} \rangle \le 0$, hence γ is non-negative. By Proposition 2.10, this function γ defines a D-diffusion stencil.

Proof that γ is supported on the set of strict M-Voronoi vectors. For any $f, g \in \mathbb{Z}^2$ one has

$$\langle f, D g \rangle = \langle f^\perp, M g^\perp \rangle / \det(M).$$

Hence $\{e_0^\perp, e_1^\perp, e_2^\perp\}$ is an M-obtuse superbase, and therefore $(e_0^\perp, -e_2^\perp, e_1^\perp, -e_0^\perp, e_2^\perp, -e_1^\perp)$ is an M-eikonal stencil by Proposition 2.8. By Proposition 2.6 all strict M-Voronoi vectors belong to $[e_0^\perp, -e_2^\perp, e_1^\perp, -e_0^\perp, e_2^\perp, -e_1^\perp] \cap \mathbf{Z} = \{e_0^\perp, -e_2^\perp, e_1^\perp, -e_0^\perp, e_2^\perp, -e_1^\perp\} \supseteq \mathrm{supp}(\gamma)$.

If M is generic, then by Proposition 2.9 there are six strict M-Voronoi vectors, and the result is proved. If M is non-generic, then there are only four strict M-Voronoi vectors, which are the members of an M-orthogonal basis - w.l.o.g. $(e_1^\perp, e_2^\perp)$ - and their opposites. Observe now that $\gamma(e_0^\perp) = -\langle e_1^\perp, D e_2^\perp \rangle = -\langle e_1^\perp, M e_2^\perp \rangle / \det(M) = 0$. Thus $\gamma(\pm e_0^\perp) = 0$ and therefore γ is supported on the set $\{\pm e_1^\perp, \pm e_2^\perp\}$ of strict M-Voronoi vectors. Finally, uniqueness follows from the Symmetry axiom of D-Diffusion stencils, and from Proposition 2.10 which states that $(e_i^\perp \otimes e_i^\perp)_{i=0}^2$ is a basis of S_2. $\square$

We show in Corollary 2.13 that the support of a D-diffusion stencil, ordered trigonometrically, defines an M-diffusion stencil, with $M = D^{-1}$. The minimality of the collection of strict M-Voronoi vectors as a D-diffusion stencil, announced in Theorem 1.12, thus follows from its minimality as an M-eikonal stencil, established in §2.1. Note that, in contrast with the two-dimensional setting studied in this paper, the three-dimensional D-diffusion and M-eikonal stencils of [FM13, Mir14] are distinct, and in particular they do not have the same number of vertices. We denote $\mathbb{S}^1 := \{z \in \mathbb{R}^2; \|z\| = 1\}$.

Lemma 2.12. *Let $\gamma : \mathbb{S}^1 \to \mathbb{R}$ be symmetric and non-negative, i.e. $\gamma(-e) = \gamma(e) \geq 0$ for all $e \in \mathbb{S}^1$. Assume that γ is supported on a finite set $V \subseteq \mathbb{S}^1$, and that $\sum_{e \in V} \gamma(e) e \otimes e = \mathrm{Id}$. Then any trigonometrically consecutive $f, g \in V$ satisfy $\langle f, g \rangle \geq 0$.*

Proof. We claim that for any $\xi \in S^1$ there exists $e \in V$ such that $\langle \xi, e \rangle \geq 1/\sqrt{2}$. Indeed, since V is symmetric, we obtain otherwise $\langle \xi, e \rangle^2 < 1/2$ for all $e \in V$, but this implies the contradiction

$$1 = \|\xi\|^2 = \sum_{e \in V} \gamma(e) \langle e, \xi \rangle^2 < \sum_{e \in V} \frac{\gamma(e)}{2} = \frac{1}{2} \mathrm{Tr} \left(\sum_{e \in V} \gamma(e) \, e \otimes e \right) = \frac{1}{2} \mathrm{Tr}(\mathrm{Id}) = 1.$$

Let $f, g \in V$ be trigonometrically consecutive, and let $\xi := (f + g)/\|f + g\|$. By construction $\langle \xi, f \rangle = \langle \xi, g \rangle \geq \langle \xi, e \rangle$ for all $e \in V$. Thus $1/\sqrt{2} \leq \langle \xi, f \rangle = \langle \xi, g \rangle$ by the above. Therefore the angle between f and g is at most $2 \arccos(1/\sqrt{2}) = \pi/2$, which concludes the proof. $\square$

Corollary 2.13. *Let $M \in S_2^+$ and let $D := M^{-1}$. Then the support of any D-diffusion stencil, ordered trigonometrically, defines an M-eikonal stencil. Therefore its convex hull contains all strict M-Voronoi vectors.*

Proof. Let $\gamma : \mathbf{Z} \to \mathbb{R}$ be the D-diffusion stencil, and let A be an invertible matrix of positive determinant such that $M = A^{\mathrm{T}} A$. Apply Lemma 2.12 to $\tilde{\gamma} : S^1 \to \mathbb{R}$ defined by $\tilde{\gamma}(Ae/\|Ae\|) = \frac{1}{2} \gamma(e) \|Ae\|^2$ for all $e \in \mathbf{Z}$. Then $0 \leq \langle Af, Ag \rangle = \langle f, Mg \rangle / (\|f\|_M^2 \|g\|_M^2)$ for any trigonometrically consecutive $f, g \in \mathrm{supp}(\gamma)$, thus $\langle f, Mg \rangle \geq 0$. One has $\det(f, g) \geq 0$ since the set $\mathrm{supp}(\gamma)$ is symmetric, hence $\det(f, g) > 0$ since otherwise $f, g \in \mathbf{Z}$ would be opposite, in contradiction with $\langle f, Mg \rangle \geq 0$. As announced, we recognize an M-eikonal stencil. Finally, the convex hull of an M-eikonal stencil contains all strict M-Voronoi vectors by Proposition 2.6. $\square$

3 Average norm of Voronoi vectors

We estimate the norm of $M_\kappa(\theta)$-Voronoi vectors, as announced in Theorem 1.19, for a fixed condition number κ^2, but in the $L^p([0, \pi[)$ sense w.r.t. the anisotropy orientation θ. This type of result belongs to the well established field of lattice geometry, dating back to Lagrange [?]. Lattice geometry is currently the object of an important research activity, motivated by the numerous applications of the LLL algorithm [?] in particular in the field of cryptography. Nevertheless, the rather elementary Theorem 1.19 is original (to the author's knowledge), for the following reasons:

(a) Research on lattice geometry is mainly focused on high dimensions.

(b) For many applications of current interest, such as cryptography, it is natural to assume that the lattice Gram matrix M has integer coefficients, in contrast with the present setting where $M = M_\kappa(\theta)$.

(c) The proposed random lattice model, which can be reformulated as $\Lambda = DR\mathbb{Z}^d$ with D diagonal of eigenvalues $\kappa^{\pm \frac{1}{2}}$ and R a random rotation, differs from classical [?] studies.

(d) The proposed result intertwines two mutually "singular" norms: the euclidean norm and the anisotropic $\| \cdot \|_M$ norm. Indeed we estimate the euclidean norm of M-Voronoi vectors. In contrast, research in lattice geometry most commonly focuses on a single norm.

3.1 Reduced basis of a lattice

We introduce the concept of Minkowski basis [Ngu04], and use it to describe the set of M-Voronoi vectors and to derive uniform bounds on their norms. To each $M \in S_2^+$ we associate an $\|\cdot\|_M$-shortest non-zero vector $e_1(M) \in \mathbb{Z}^2$, and an $\|\cdot\|_M$-shortest linearly independent vector $e_2(M) \in \mathbb{Z}^2$

$$e_1(M) \in \underset{e \in \mathbb{Z}^2 \setminus \{0\}}{\operatorname{argmin}} \|e\|_M, \qquad\qquad e_2(M) \in \underset{f \in \mathbb{Z}^2 \setminus e_1(M)\mathbb{Z}}{\operatorname{argmin}} \|f\|_M. \qquad (12)$$

The vectors $e_1(M)$ and $e_2(M)$ are uniquely determined, up to a change of sign, for all matrices M except for a Lebesgue negligible set. For non-generic matrices, an arbitrary minimizer is selected. In applications, the vectors (13) can be computed via Lagrange's algorithm [Ngu04], a generalization of Euclid's gcd algorithm. The next lemma shows that the two vectors (12) form a basis of $\mathbb{Z}^2$, often referred to as a Minkowski reduced basis [Ngu04].

Proposition 3.1. *For any $M \in S_2^+$, the Minkowski vectors $e_1(M), e_2(M)$ form a basis of $\mathbb{Z}^2$.*

Proof. Let $e_1 := e_1(M)$ and $e_2 := e_2(M)$. One has $\gcd(e_1) = \gcd(e_2) = 1$, since otherwise the vector $e_i/\gcd(e_i)$ would violate the minimality property of e_i, $i \in \{1,2\}$, see (12). Since $e_2 \in \mathbb{Z}^2 \setminus e_1\mathbb{Z}$ and $\gcd(e_1) = 1$, the vectors e_1, e_2 are not collinear, hence $\det(e_1, e_2) \neq 0$. Assume for contradiction that $|\det(e_1, e_2)| > 1$ and, by Lemma 2.4, consider an element $e \in \mathbb{Z}^2$ of the triangle $[0, e_1, e_2]$ distinct from its vertices: $e = \alpha e_1 + \beta e_2$, $\alpha, \beta \geq 0$, $\alpha + \beta \leq 1$. Since $\gcd(e_1) = \gcd(e_2) = 1$, the vector e does not lie on an edge $[0, e_i]$, $i \in \{1, 2\}$. Hence by the triangular inequality, which is strict since e_1, e_2 are not collinear and α, β are positive, we obtain $\|e\|_M < \alpha\|e_1\|_M + \beta\|e_2\|_M \leq \|e_2\|_M$. This contradicts the minimality of $\|e_2\|_M$, hence $|\det(e_1, e_2)| = 1$ which concludes the proof. $\qquad\square$

The following lemma upper bounds the orthogonality defect of a Minkowski reduced basis. See [Ngu04] for extensions of this result to higher dimension.

Lemma 3.2. *Let $M \in S_2^+$ and let $e_1 := e_1(M)$ and $e_2 := e_2(M)$. Then $2|\langle e_1, Me_2\rangle| \leq \|e_1\|^2$.*

Proof. By construction $\|e_2\|_M^2 \leq \|e_2 + e_1\|_M^2$, and $\|e_2\|_M^2 \leq \|e_1 - e_2\|_M^2$. The result follows. $\quad\square$

Our next lemma describes the collection of strict M-Voronoi vectors, in terms of Minkowski's shortest vectors (12).

Proposition 3.3. *Let $M \in S_2^+$ and $e_i := e_i(M)$, $i \in \{1,2\}$. The strict M-Voronoi vectors are:*

(i) if $\langle e_1, Me_2\rangle = 0$ then $\{\pm e_1, \pm e_2\}$.

(ii) if $\langle e_1, Me_2\rangle < 0$ then $\{\pm e_0, \pm e_1, \pm e_2\}$, with $e_0 := -(e_1 + e_2)$. In addition $\langle e_0, Me_1\rangle < 0$ and $\langle e_0, Me_2\rangle < 0$. (In the case $\langle e_1, Me_2\rangle > 0$, replace e_2 with its opposite.)

Proof. If M is non-generic, then strict Voronoi vectors are the four smallest elements of $\mathbf{Z}$ by Proposition 1.16, hence they are precisely $\{\pm e_1, \pm e_2\}$ by construction (12). Furthermore strict Voronoi vectors feature an M-orthogonal basis by Proposition 2.9, hence case (i) holds.

If M-is generic, then $\langle e_1, Me_2\rangle \neq 0$ hence $\langle e_1, Me_2\rangle < 0$ up to replacing e_2 with $-e_2$. By Lemma 3.2 we obtain $\langle e_0, Me_1\rangle = -\langle e_2, Me_1\rangle - \|e_1\|_M^2 < 0$ and $\langle e_0, Me_2\rangle = -\langle e_1, Me_2\rangle - \|e_2\|_M^2 < 0$, with $e_0 := -(e_1 + e_2)$. This shows that (e_0, e_1, e_2) is an M-obtuse superbase, and case (ii) holds by Proposition 2.9. $\qquad\square$
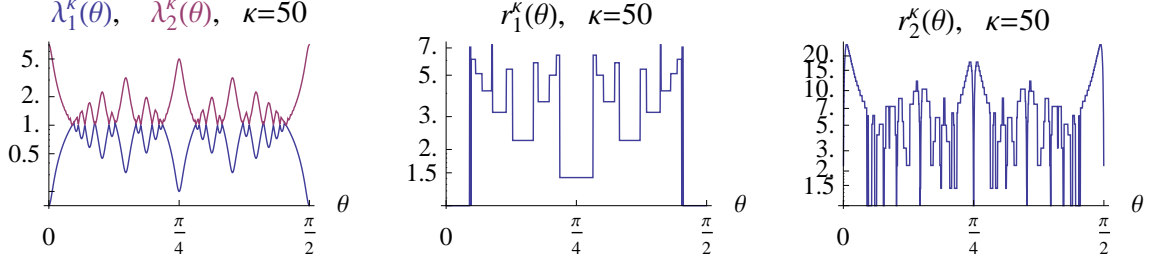
Figure 7: Minkowski minima λ_1^κ, λ_2^κ, and radii of Minkowski's vectors r_1^κ, r_2^κ. Logarithmic scale.

Our objective, in Theorem 1.19, is to estimate the radius of the set of strict M-Voronoi vectors. In light of the previous lemma, we introduce for $M \in S_2^+$, $i \in \{1,2\}$, the norms

$$\lambda_i(M) := \|e_i(M)\|_M, \qquad\qquad r_i(M) := \|e_i(M)\|. \qquad (13)$$

The scalars $\lambda_i(M)$ are called the Minkowski minima, and depend continuously on the matrix $M \in S_2^+$, in contrast with $r_i(M)$ and $e_i(M)$ which take discrete values. Minkowski minima are tied together by Minkowski's second theorem [?], a fundamental result of lattice geometry stating that:

$$2 \leq \frac{\pi}{\sqrt{\det M}} \lambda_1(M)\lambda_2(M) \leq 4. \qquad (14)$$

For each $\kappa \geq 1$, $\theta \in \mathbb{R}$, $i \in \{1,2\}$, we denote by $e_i^\kappa(\theta)$, $\lambda_i^\kappa(\theta)$, $r_i^\kappa(\theta)$, the vectors (12) and their norms (13) attached to the matrix $M_\kappa(\theta)$, defined in (7). Inequality (14) rewrites as

$$2/\pi \leq \lambda_1^\kappa(\theta)\lambda_2^\kappa(\theta) \leq 4/\pi. \qquad (15)$$

This explains the approximate vertical symmetry in the plot of λ_1^κ, λ_2^κ, Figure 7 (left, log scale). We denote by $\sphericalangle(f,g) \in [0,\pi/2]$ the unoriented angle between two vectors $f,g \in \mathbb{R}^2 \setminus \{0\}$, considered up to a change of sign:

$$\sphericalangle(f,g) := \arccos\left(\frac{|\langle f,g\rangle|}{\|f\|\|g\|}\right) \in [0,\pi/2]. \qquad (16)$$

Let $e \in \mathbb{R}^2 \setminus \{0\}$, let $\theta \in \mathbb{R}$, and let $\varphi = \sphericalangle(e,e_\theta)$ where $e_\theta := (\cos\theta, \sin\theta)$. Then

$$\|e\|_{M_\kappa(\theta)}^2 = \|e\|^2(\kappa \sin^2\varphi + \kappa^{-1}\cos^2\varphi), \qquad \text{hence } \kappa^{-\frac{1}{2}}\|e\| \leq \|e\|_{M_\kappa(\theta)} \leq \kappa^{\frac{1}{2}}\|e\|. \qquad (17)$$

Lemma 3.4 (Uniform bounds on Minkowski's vectors)**.** *For any $\kappa \in [1,\infty[$, and any $\theta \in \mathbb{R}$ (omitting the argument θ for readability)*

$$\lambda_1^\kappa \in [\kappa^{-\frac{1}{2}}, c_1] \qquad\qquad\qquad \lambda_2^\kappa \in [c_2, \kappa^{\frac{1}{2}}],$$
$$r_1^\kappa \in [1, c_1\kappa^{\frac{1}{2}}] \qquad\qquad\qquad r_2^\kappa \in [1, \kappa],$$

where $c_1 := 2/\sqrt{\pi}$, $c_2 := \sqrt{2/\pi}$.

Proof. First line. Since $\lambda_1^\kappa \leq \lambda_2^\kappa$ by construction, and $c_2^2 \leq \lambda_1^\kappa\lambda_2^\kappa \leq c_1^2$ by (15), we obtain $\lambda_1^\kappa \leq c_1$ and $\lambda_2^\kappa \geq c_2$. By minimality (13) of $\|e_2^\kappa\|_{M_\kappa}$, and since at least one of the two independent vectors $(1,0)$ or $(0,1)$ is non-collinear with e_1^κ, one has $\lambda_2^\kappa = \|e_2^\kappa\|_{M_\kappa} \leq \max\{\|(1,0)\|_{M_\kappa}, \|(0,1)\|_{M_\kappa}\} \leq \kappa^{\frac{1}{2}}\max\{\|(1,0)\|, \|(0,1)\|\} = \kappa^{\frac{1}{2}}$ as announced. The lower bound for λ_1^κ is shown below.

Second line. One has $1 \leq \|e_1^\kappa\| =: r_1^\kappa$ since $e_1^\kappa \in \mathbb{Z}^2 \setminus \{0\}$, and likewise $r_2^\kappa \geq 1$. From (17, right) we obtain $\lambda_\kappa^1 \geq \kappa^{-\frac{1}{2}}r_1^\kappa \geq \kappa^{-\frac{1}{2}}$ which completes the first line, and $r_1^\kappa \leq \kappa^{\frac{1}{2}}\lambda_1^\kappa \leq c_1\kappa^{\frac{1}{2}}$ and $r_2^\kappa \leq \kappa^{\frac{1}{2}}\lambda_2^\kappa \leq \kappa$ which concludes the proof. $\qquad\square$

18

Corollary 3.5. *With the notations of Lemma 3.4, one has for $\kappa \geq c_1^2$*

$$\frac{1}{2}R_\kappa \leq \max\{r_1^\kappa, r_2^\kappa\} \leq R_\kappa, \qquad\qquad \frac{1}{2}S_\kappa \leq \lambda_2^\kappa \leq S_\kappa, \qquad (18)$$

$$R_\kappa \in [1, 2\kappa], \qquad\qquad S_\kappa \in [c_2, 2\kappa^{\frac{1}{2}}].$$

Proof. By Proposition 3.3, one has $\max\{r_1^\kappa, r_2^\kappa\} \leq R_\kappa \leq \max\{r_1^\kappa, r_2^\kappa, r_1^\kappa + r_2^\kappa\} \leq 2\max\{r_1^\kappa, r_2^\kappa\}$, as announced in (18, left). Proceeding likewise with $S_\kappa, \lambda_1^\kappa, \lambda_2^\kappa$, and observing that $\lambda_1^\kappa \leq \lambda_2^\kappa$ by construction (12), we obtain (18, right). The announced uniform bounds on R_κ, S_k, follow from those of Lemma 3.4. $\qquad\square$

We introduce a set Z, which collects one representative vector for each rational direction in the plane

$$Z := \{e \in \mathbb{Z}^2; \gcd(e) = 1, e \succ (0,0) \text{ lexicographically}\}.$$

For any $e \in \mathbb{Z}^2 \setminus \{0\}$, the line $e\mathbb{R}$ intersects Z at a single point. Note also that $\{-1, 1\} \times Z \to \mathbf{Z}$: $(\varepsilon, e) \mapsto \varepsilon e$ is clearly a bijection, where $\mathbf{Z}$ is defined in (1). Up to changing their sign, we assume from this point that Minkowski's vectors $e_i(M)$, $i \in \{1, 2\}$, $M \in S_2^+$, belong to Z. We end this subsection by estimating a sum, over pairs of co-prime integers, which repeatedly appears in the sequel.

Proposition 3.6. *One has*

$$\lim_{r \to \infty} \frac{\Sigma(r)}{r} = \frac{6}{\pi}, \qquad\qquad \text{where } \Sigma(r) := \sum_{e \in Z, \|e\| \leq r} \frac{1}{\|e\|}.$$

In particular, there exists $0 < c \leq C < \infty$ such that $cr \leq \Sigma(r) \leq Cr$ for all $r \geq 1$.

Proof. Denote by $p(r)$ the the relative density of points with co-prime coordinates in the ball of radius r:

$$p(r) := \frac{1}{\pi r^2}\#\{e \in \mathbb{Z}^2; 0 < \|e\| \leq r, \gcd(e) = 1\}.$$

It is well known [HW79] that $p(r) \to 1/\zeta(2) = 6/\pi^2$ as $r \to \infty$. The announced result then follows from the observations

$$\frac{1}{\|e\|} = \int_{\|e\|}^r \frac{1}{s^2}ds + \frac{1}{r}, \qquad \text{hence } \frac{2\Sigma(r)}{\pi r} = \frac{1}{r}\int_0^r p(s)ds + p(r) = \frac{2}{\zeta(2)} + o(1). \qquad \square$$

3.2 Probabilistic estimate of the Minkowksi minima

We estimate the cumulative distribution of λ_1^κ, seen as a random variable of $\theta \in [0, \pi[$, for a fixed $\kappa \geq 1$, see Figure 9. The tail distribution of λ_2^κ is then inferred using Minkowski's second theorem (14). For that purpose we introduce the following angular sectors, illustrated on Figure 8 (left): for each $\kappa \geq 1$, $e \in Z$, $\alpha > 0$

$$\Phi_e^\kappa(\alpha) := \left\{\theta \in [0, \pi[; \sin \sphericalangle(e, e_\theta) \leq \frac{\alpha}{\|e\|\sqrt{\kappa}}\right\}. \qquad (19)$$

Lemma 3.7. *Let $\kappa \geq 1$, let $e \in Z$, let $\theta \in [0, \pi[$ and let $c > 0$. We have the implications:*

$$\|e\|_{M_\kappa(\theta)} \leq c \implies \left(\theta \in \Phi_e^\kappa(c) \text{ and } \|e\| \leq c\sqrt{\kappa}\right) \implies \|e\|_{M_\kappa(\theta)} \leq c\sqrt{2}.$$
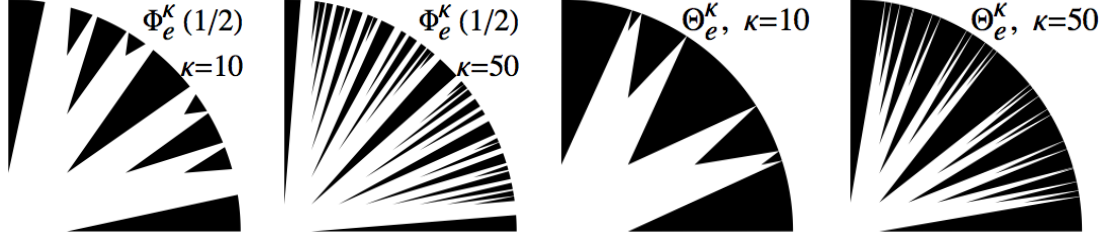
19

Figure 8: Left: representation of the sets $\Phi_e^\kappa(1/2)$, with $e \in Z$ and $\|e\| \leq \sqrt{\kappa}$, defined in (19). Right: sets Θ_e^κ defined in (23) below. In this representation, we show in black the convex hull of each vector e and of the corresponding angular sector on the circle of radius κ.

Proof. First implication. If $\|e\|_{M_\kappa(\theta)} \leq c$ then from (17, left), neglecting the contribution of $\kappa^{-1} \cos^2(\varphi)$, we obtain $\theta \in \Phi_e^\kappa(c)$. From (17, right) we get $\|e\| \leq c\sqrt{\kappa}$.

Second implication. Using again (17, left), we obtain:

$$\|e\|_{M_\kappa(\theta)}^2 \leq \|e\|^2 \left(\kappa \left(\frac{c}{\|e\|\sqrt{\kappa}} \right)^2 + \kappa^{-1} \right) = c^2 + \|e\|^2 \kappa^{-1} \leq 2c^2. \qquad \square$$

Corollary 3.8. *For any $\kappa \geq 1$, and any $\lambda \geq \kappa^{-\frac{1}{2}}$, one has with $\lambda' := \lambda/\sqrt{2}$*

$$\bigcup_{\substack{e \in Z \\ \|e\| \leq \lambda'\sqrt{\kappa}}} \Phi_e^\kappa(\lambda') \quad \subseteq \quad \{\theta \in [0,\pi[; \lambda_1^\kappa(\theta) \leq \lambda\} \quad \subseteq \quad \bigcup_{\substack{e \in Z \\ \|e\| \leq \lambda\sqrt{\kappa}}} \Phi_e^\kappa(\lambda). \tag{20}$$

Proof. The right implication in Lemma 3.7, with $c = \lambda' = \lambda/\sqrt{2}$, yields the first inclusion, since $\lambda_1^\kappa(\theta) \leq \|e\|_{M_\kappa(\theta)}$. The left implication in Lemma 3.7, with $c = \lambda$, yields the second inclusion. $\square$

We estimate in Proposition 3.10 the distribution function $P(\lambda_1^\kappa \leq \lambda)$. The main ingredient of the proof is that, under suitable assumptions, the union appearing in (20, left) is of disjoint sets. Probabilities considered here and in the following are for a uniformly distributed angle $\theta \in [0, \pi[$. We recall that for any $\varphi \in]0, \pi/2]$ one has

$$\frac{2}{\pi}\varphi \leq \sin\varphi < \varphi. \tag{21}$$

Denote by $|E|$ the Lebesgue measure of a set $E \subseteq \mathbb{R}$. Using (21) we obtain for any $e \in Z$, $\kappa \geq 1$ and $\alpha \leq 1$, denoting $\sigma := \frac{\alpha}{\|e\|\sqrt{\kappa}}$

$$\frac{2\alpha}{\|e\|\sqrt{\kappa}} \leq |\Phi_e^\kappa(\alpha)| = |[-\arcsin\sigma, \arcsin\sigma]| = 2\arcsin\sigma \leq \frac{\pi\alpha}{\|e\|\sqrt{\kappa}}. \tag{22}$$

Lemma 3.9. *Let f, g be distinct elements of Z such that such that $\max\{\|f\|, \|g\|\} \leq c\sqrt{\kappa}$, where $c = 1/\pi$ and $\kappa \geq 1$. Then $\Phi_f^\kappa(1) \cap \Phi_g^\kappa(1) = \emptyset$.*

Proof. Since f, g are distinct elements of Z, they are not collinear. Hence $\det(f, g) \neq 0$, and therefore $|\det(f, g)| \geq 1$ since this determinant is a positive integer. Thus

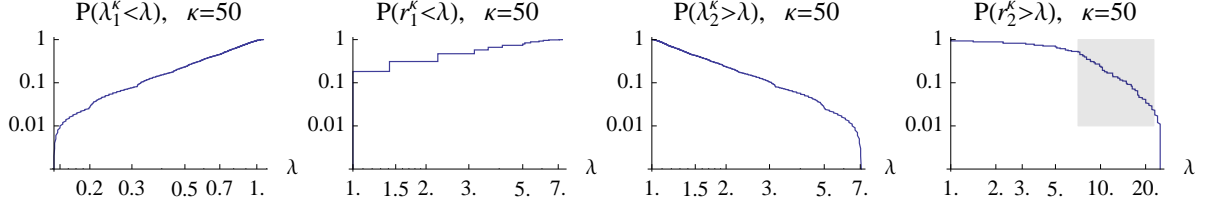$$\sphericalangle(f, g) > \sin\sphericalangle(f, g) = \frac{|\det(f, g)|}{\|f\|\|g\|} \geq \frac{1}{\|f\|\|g\|}.$$

20

Figure 9: Cumulative distributions of λ_1^κ, r_1^κ, and tail distributions of λ_2^κ, r_2^κ. Log-log scales.

Assuming that $\theta \in \Phi_f^\kappa(1) \cap \Phi_g^\kappa(1)$, we obtain using (21) the contradiction

$$\sphericalangle(f,g) \leq \sphericalangle(f,e_\theta) + \sphericalangle(e_\theta,g) \leq \frac{\pi}{2}(\sin \sphericalangle(f,e_\theta) + \sin \sphericalangle(e_\theta,g))$$

$$\leq \frac{\pi}{2}\left(\frac{1}{\|f\|\sqrt{\kappa}} + \frac{1}{\|g\|\sqrt{\kappa}}\right) \leq \frac{\pi c}{\|f\|\|g\|}. \qquad \square$$

Proposition 3.10 (Cumulative distribution of λ_1^κ)**.** *There exists constants $C, c > 0$ such that: for all $\kappa \geq 1$, and all $\lambda \in [C\kappa^{-\frac{1}{2}}, c]$, one has $c\lambda^2 \leq P(\lambda_1^\kappa \leq \lambda) \leq C\lambda^2$.*

Proof. Let c_0 be the constant of Lemma 3.9, and let $\lambda' := \lambda/\sqrt{2}$. For $\lambda' \leq c_0$, the left term of (20) is a disjoint union, by Lemma 3.9. Hence we obtain successively, using (22) in the second line, and introducing Σ of Proposition 3.6 in the third

$$\sum_{\substack{e \in Z \\ \|e\| \leq \lambda'\sqrt{\kappa}}} |\Phi_e^\kappa(\lambda')| \ \leq \ \pi P(\lambda_1^\kappa \leq \lambda) \ \leq \ \sum_{\substack{e \in Z \\ \|e\| \leq \lambda\sqrt{\kappa}}} |\Phi_e^\kappa(\lambda)|,$$

$$\sum_{\substack{e \in Z \\ \|e\| \leq \lambda'\sqrt{\kappa}}} \frac{2\lambda'}{\|e\|\sqrt{\kappa}} \ \leq \ \pi P(\lambda_1^\kappa \leq \lambda) \ \leq \ \sum_{\substack{e \in Z \\ \|e\| \leq \lambda\sqrt{\kappa}}} \frac{\pi\lambda}{\|e\|\sqrt{\kappa}},$$

$$2\frac{\lambda'}{\sqrt{\kappa}}\Sigma(\lambda'\sqrt{\kappa}) \ \leq \ \pi P(\lambda_1^\kappa \leq \lambda) \ \leq \ \pi\frac{\lambda}{\sqrt{\kappa}}\Sigma(\lambda\sqrt{\kappa}).$$

Assume in addition that $\lambda'\sqrt{\kappa} \geq 1$. Then denoting by c_1, C_1 the constants of Proposition 3.6

$$c_1\lambda^2 = 2\frac{\lambda'}{\sqrt{\kappa}}c_1\lambda'\sqrt{\kappa} \ \leq \ \pi P(\lambda_1^\kappa \leq \lambda) \ \leq \ \pi\frac{\lambda}{\sqrt{\kappa}}C_1\lambda\sqrt{\kappa} = \pi C_1\lambda^2.$$

This concludes the proof, with $c = \min\{c_0\sqrt{2}, c_1/\pi\}$ and $C = \max\{\sqrt{2}, C_1\}$. $\qquad \square$

Corollary 3.11 (Tail distribution of λ_2^κ)**.** *There exists constants $C, c > 0$ such that: for any $\kappa \geq 1$, and any $\lambda \in [C, c\kappa^{\frac{1}{2}}]$ one has $c\lambda^{-2} \leq P(\lambda_2^\kappa \geq \lambda) \leq C\lambda^{-2}$.*

Proof. The result follows from Proposition 3.10, and from (15) which implies for any $\lambda > 0$:

$$P(\lambda_1^\kappa \leq 2/(\lambda\pi)) \ \leq \ P(\lambda_2^\kappa \geq \lambda) \ \leq \ P(\lambda_1^\kappa \leq 4/(\lambda\pi)). \qquad \square$$

3.3 Euclidean norm of Minkowski's shortest vector

We estimate the distribution of the euclidean norm $r_1^\kappa(\theta)$ of Minkowski's vector $e_1^\kappa(\theta) \in Z$, which is defined in (12) as the shortest vector with respect to the anisotropic norm $\|\cdot\|_{M_\kappa(\theta)}$. For that

21

purpose, we need to precisely identify this vector, hence we introduce for each $\kappa \geq 1$, $e \in Z$, the set

$$\Theta_e^\kappa := \{\theta \in [0, \pi[; \; e = e_1^\kappa(\theta)\}, \tag{23}$$

Our first two lemmas compare Θ_e^κ with the angular sector $\Phi_e^\kappa(c)$ introduced in (19).

Lemma 3.12. *Let $\kappa \geq 1$, $e \in Z$, and $C = 2/\sqrt{\pi}$. Then $\Theta_e^\kappa \subseteq \Phi_e^\kappa(C)$, and $\left(\Theta_e^\kappa \neq \emptyset \Rightarrow \|e\| \leq C\kappa^{\frac{1}{2}}\right)$.*

Proof. Using successively (i) the definition of Θ_e^κ, (ii) the uniform bound on $\lambda_1^\kappa(\theta)$ obtained in Lemma 3.4, for which $C = c_1$, and (iii) Lemma 3.7 (left implication) we obtain the implications

$$\theta \in \Theta_e^\kappa \Rightarrow \left(\|e\|_{M_\kappa(\theta)} = \lambda_1^\kappa(\theta)\right) \Rightarrow \left(\|e\|_{M_\kappa(\theta)} \leq C\right) \Rightarrow \left(\theta \in \Phi_e^\kappa(C) \text{ and } \|e\| \leq C\sqrt{\kappa}\right).$$

This establishes the announced inclusion and bound on $\|e\|$. $\qquad\square$

Lemma 3.13. *Let $\kappa \geq 1$, $e \in Z$, and $c < 1/(\pi\sqrt{2})$. If $\|e\| \leq c\sqrt{\kappa}$ then $\Theta_e^\kappa \supseteq \Phi_e^\kappa(c)$.*

Proof. Let $\theta \in \Phi_e^\kappa(c)$, and $M := M_\kappa(\theta)$. By Lemma 3.7 (right implication) one has $\|e\|_M \leq c\sqrt{2}$.

Let $e' \in Z \setminus \{e\}$, and let $c_0 := 1/\pi$ be the constant from Lemma 3.9. If $\|e'\| \geq c_0\sqrt{\kappa}$, then $\|e'\|_M \geq \kappa^{-\frac{1}{2}}\|e'\| \geq c_0 > c\sqrt{2} \geq \|e\|_M$. On the other hand, if $\|e'\| \leq c_0\sqrt{\kappa}$, then $\Phi_e^\kappa(1) \cap \Phi_{e'}^\kappa(1) = \emptyset$ by Lemma 3.9. Thus $\theta \notin \Phi_{e'}^\kappa(1)$ and therefore $\|e'\|_M \geq 1 > c\sqrt{2} \geq \|e\|_M$ by Lemma 3.7 (left implication). We have shown that $\|e\|_M < \|e'\|_M$ for any $e' \in Z \setminus \{e\}$. Thus $e = e_1(M) = e_1^\kappa(\theta)$, in other words $\theta \in \Theta_e^\kappa$ as announced. $\qquad\square$

Our next proposition estimate the distribution of r_1^κ. Strictly speaking this is not needed for the proof of Theorem 1.19, but we provide for completeness the argument.

Proposition 3.14 (Cumulative distribution of r_1^κ)**.** *There exists $C, c > 0$ such that for any $\kappa \geq 1$ and any $\lambda \in [\kappa^{-\frac{1}{2}}, c]$, one has $c\lambda \leq P(r_1^\kappa \leq \lambda\sqrt{\kappa}) \leq C\lambda$.*

Proof. Note that for any fixed $\kappa \geq 1$, the sets Θ_e^κ, $e \in Z$ are pairwise disjoint. Denoting by C_0 and c_0 the constants of Lemmas 3.12 and 3.13 respectively, one thus has for $\lambda \leq c_0$

$$\sum_{\substack{e \in Z \\ \|e\| \leq \lambda\sqrt{\kappa}}} |\Phi_e^\kappa(c_0)| \leq \sum_{\substack{e \in Z \\ \|e\| \leq \lambda\sqrt{\kappa}}} |\Theta_e^\kappa| = \pi P(r_1^\kappa \leq \lambda\sqrt{\kappa}) \leq \sum_{\substack{e \in Z \\ \|e\| \leq \lambda\sqrt{\kappa}}} |\Phi_e^\kappa(C_0)|$$

Using (22) for the first line, and denoting by c_1, C_1 the constants of Proposition 3.6 for the second line, we conclude that when $\lambda\sqrt{\kappa} \geq 1$

$$\sum_{\substack{e \in Z \\ \|e\| \leq \lambda\sqrt{\kappa}}} \frac{2c_0}{\|e\|\sqrt{\kappa}} \leq \pi P(r_1^\kappa \leq \lambda\sqrt{\kappa}) \leq \sum_{\substack{e \in Z \\ \|e\| \leq \lambda\sqrt{\kappa}}} \frac{\pi C_0}{\|e\|\sqrt{\kappa}}.$$

$$2c_0 c_1 \lambda \leq 2c_0 \frac{\Sigma(\lambda\sqrt{\kappa})}{\sqrt{\kappa}} \leq \pi P(r_1^\kappa \leq \lambda\sqrt{\kappa}) \leq \pi C_0 \frac{\Sigma(\lambda\sqrt{\kappa})}{\sqrt{\kappa}} \leq \pi C_0 C_1 \lambda. \qquad\square$$

3.4 Euclidean norm of Minkowski's second shortest vector

We estimate the tail distribution of the euclidean norm $r_2^\kappa(\theta)$ of Minkowski's vector $e_2^\kappa(\theta)$, the second shortest vector with respect to the anisotropic norm $\|\cdot\|_{M_\kappa(\theta)}$. The upper bound for the tail distribution of r_2^κ directly follows from the one obtained in Corollary 3.11 for λ_2^κ, as shown in the next lemma.

Lemma 3.15 (Tail distribution of r_2^κ. Upper bound)**.** *There exists $C, c > 0$ such that for any $\kappa \geq 1$ and any $\lambda \in [C, c\kappa^{\frac{1}{2}}]$, one has $P(r_2^\kappa \geq \lambda\sqrt{\kappa}) \leq C\lambda^{-2}$.*

Proof. By (17, right) one has $r_2^\kappa(\theta) = \|e_2^\kappa(\theta)\| \leq \sqrt{\kappa}\|e_2^\kappa(\theta)\|_{M_\kappa(\theta)} = \lambda_2^\kappa(\theta)\sqrt{\kappa}$, for any $\kappa \geq 1$, $\theta \in \mathbb{R}$. Hence denoting by $C, c > 0$ the constants of Corollary 3.11 we obtain $P(r_2^\kappa \geq \lambda\sqrt{\kappa}) \leq P(\lambda_2^\kappa \geq \lambda) \leq C\lambda^{-2}$, for all $\lambda \in [C\kappa^{\frac{1}{2}}, c]$. $\qquad\square$

Note that the estimate of the tail distribution $P(r_2^\kappa \geq \mu)$ obtained in Lemma 3.15 is limited to $\mu \in [C\sqrt{\kappa}, c\kappa]$, whereas by Lemma 3.4 the radius r_2^κ may take values in the larger interval $[1, \kappa]$. This estimate is nevertheless sufficient for our purposes; the regime in which it applies is shown in the grayed region in Figure 9.

In order to estimate r_2^κ from below, we tie it to r_1^κ, a strategy similar to the one used for the Minkowski minima λ_1^κ and λ_2^κ in §3.2. For any $\alpha \in \mathbb{R}$ we abusively denote $\alpha + \mathbb{Z} := \{\alpha + z; \, z \in \mathbb{Z}\}$

Lemma 3.16. *Let $M \in S_2^+$, let $e, f \in \mathbb{R}^d \setminus \{0\}$, let $\alpha \in \mathbb{R}$, and let k be a minimizer $\|f - ke\|_M$ in $\alpha + \mathbb{Z}$. Then $|k| + \frac{1}{2} \geq |\langle e, Mf \rangle| / \|e\|_M^2$.*

Proof. Up to a linear change of coordinates, we may assume that $M = \text{Id}$. On the whole real line $\mathbb{R}$, the minimizer of $\lambda(l) := \|f - le\|^2 = \|f\|^2 - 2l\langle e, f \rangle + l^2\|e\|^2$ is $l_* := \langle f, e \rangle / \|e\|^2$. Since $\lambda(l_* + \delta) = \lambda(l_*) + \delta^2\|e\|^2$ for any $\delta \in \mathbb{R}$, we find that k is the (an) element of $(\alpha + \mathbb{Z})$ closest to l_*. Thus $|k - l_*| \leq \frac{1}{2}$, which concludes the proof. $\qquad\square$

Corollary 3.17. *Let $\kappa \geq 1$ and $\theta \in \mathbb{R}$. Let us denote $e_1 := e_1^\kappa(\theta)$, $M := M_\kappa(\theta)$, and*

$$\mu_\kappa(\theta) := \frac{|\langle e_1, Me_1^\perp \rangle|}{\|e_1\|_M^2 \|e_1\|^2}.$$

Then $r_2^\kappa(\theta) \geq r_1^\kappa(\theta)(\mu_\kappa(\theta) - 1/2)$.

Proof. Let $f := e_1^\perp / \|e_1\|^2$, let $e_2 := e_2^\kappa(\theta)$, and let us note that $|\det(e_1, e_2)| = 1 = \det(e_1, f)$. Up to replacing f with $-f$ we may thus assume that $\det(e_1, e_2 - f) = 0$. Therefore there exists $\alpha \in \mathbb{R}$ such that $e_2 = f - \alpha e_1$. For any $k \in \mathbb{R}$ one has ($f - ke_1 \in \mathbb{Z}^2$ iff $k \in \alpha + \mathbb{Z}$). By minimality (12) of $\|e_2\|_M$ among elements of $\mathbb{Z}^2$ non-collinear with e_1, we thus have $\|f - ke_1\|_M \geq \|f - \alpha e_1\|_M$ for any $k \in \alpha + \mathbb{Z}$. By Lemma 3.16

$$|\alpha| + \frac{1}{2} \geq |\langle e_1, Mf \rangle| / \|e_1\|_M^2 = \mu_\kappa(\theta).$$

Finally, $r_2^\kappa(\theta) = \|e_2\| = \|f - \alpha e_1\| \geq |\alpha|\|e_1\|$, since $\langle e_1, f \rangle = 0$, which concludes the proof. $\qquad\square$

Note that $\mu_\kappa(\theta)$ vanishes precisely when $e_1^\kappa(\theta)$ is an eigenvector of $M_\kappa(\theta)$. We study in Lemma 3.18 the angular sectors where in contrast $\mu_\kappa(\theta)$ is large. This implies in Corollary 3.19 a lower bound on the tail distribution of r_2^κ, which is the counterpart of Lemma 3.15.

Lemma 3.18. *Let $\kappa \geq \sqrt{2}$, let $\lambda \in [C, \kappa^{\frac{1}{2}}]$, and let $e \in Z$ be such that $\lambda\|e\| \leq c\sqrt{\kappa}$, where we denoted $C = 2$ and $c = 1/8$. Then*

$$\left| \left\{ \theta \in \Theta_e^\kappa; \, \mu_\kappa(\theta) \geq \frac{\lambda\sqrt{\kappa}}{\|e\|} \right\} \right| \geq \frac{c}{\|e\|\lambda\sqrt{\kappa}}. \tag{24}$$

23

Proof. Let κ, λ, e be as in the lemma statement. Let $\theta \in [0, \pi[$, let $M := M_\kappa(\theta)$, and let $\varphi = \sphericalangle(e_\theta, e)$. We define, and estimate using (17), the scalar

$$\mu := \frac{|\langle e, Me^\perp\rangle|}{\|e\|_M^2 \|e\|^2} = \|e\|^{-2} \frac{(\kappa - \frac{1}{\kappa})\cos\varphi \sin\varphi}{\kappa \sin^2\varphi + \frac{1}{\kappa}\cos^2\varphi} = \|e\|^{-2} \frac{(\kappa^2 - 1)}{\kappa^2 \tan\varphi + \cotan\varphi}. \tag{25}$$

One has the equivalence

$$\mu \geq \frac{\lambda\sqrt{\kappa}}{\|e\|} \quad \Leftrightarrow \quad \frac{\kappa^2 - 1}{\|e\|^2} \times \frac{\|e\|}{\lambda\sqrt{\kappa}} \geq \kappa^2 \tan\varphi + \cotan\varphi. \tag{26}$$

Observe that $\kappa^2 - 1 \geq \kappa^2/2$, simplify powers of κ and $\|e\|$, and bound above the sum in the r.h.s. by twice its largest element. Then (26) is the consequence of the equivalent conditions.

$$\frac{\kappa^{\frac{3}{2}}}{2\lambda\|e\|} \geq 2\max\{\kappa^2\tan\varphi, \cotan\varphi\} \quad \Leftrightarrow \quad \tan\varphi \in \left[\frac{4\lambda\|e\|}{\kappa^{\frac{3}{2}}}, \frac{1}{4\lambda\|e\|\kappa^{\frac{1}{2}}}\right] = \left[\frac{1}{\rho\kappa}, \frac{\rho}{\kappa}\right], \tag{27}$$

where we introduced the ratio $\rho := \frac{\sqrt{\kappa}}{4\lambda\|e\|} = \frac{2c\sqrt{\kappa}}{\lambda\|e\|} \geq 2$. We assume that condition (27) holds and observe that since $\lambda \geq 2$

$$\sin\varphi \leq \tan\varphi \leq \frac{2c}{\lambda\|e\|\sqrt{\kappa}} \leq \frac{c}{\|e\|\sqrt{\kappa}},$$

thus $\theta \in \Phi_e^\kappa(c)$. By Lemma 3.13, and the assumption $\|e\| \leq c\sqrt{\kappa}$, we have $\Phi_e^\kappa(c) \subseteq \Theta_e^\kappa$. Thus $\theta \in \Theta_e^\kappa$, which implies $e = e_1^\kappa(\theta)$, and $\mu = \mu_\kappa(\theta)$. Denoting by I the interval (27), the measure (24) is thus at least

$$|\tan^{-1}(I)| + |\tan^{-1}(-I)| \geq 2\frac{1}{\kappa}\left(\rho - \frac{1}{\rho}\right) \times (\tan^{-1})'\left(\frac{\rho}{\kappa}\right) \geq \frac{\rho}{2\kappa}$$

We used fact that $\rho \geq 2$, so that $\rho - \frac{1}{\rho} \geq \rho/2$. In addition $\frac{\rho}{\kappa} = \frac{2c}{\lambda\|e\|\sqrt{\kappa}} \leq c$, since $\lambda \geq 2$, $\|e\| \geq 1$ and $\kappa \geq 1$, hence $(\tan^{-1})'(\rho/\kappa) \geq 1/(1 + c^2) \geq 1/2$. This concludes the proof. $\quad\square$

Corollary 3.19 (Tail distribution of r_2^κ. Lower bound)**.** *There exists $0 < c \leq C < \infty$ such that for any $\kappa \geq 1$, and any $\lambda \in [C, c\kappa^{\frac{1}{2}}]$, one has $P(r_2^\kappa \geq \lambda\sqrt{\kappa}) \geq c\lambda^{-2}$.*

Proof. Let c_0, C_0 be the constants of Lemma 3.18, let κ, λ, e, satisfy the conditions of its statement, and let θ be in the set (24). Then

$$\mu_\kappa(\theta) \geq \frac{\lambda\sqrt{\kappa}}{\|e\|} \geq \frac{\lambda^2}{c_0} \geq 1, \qquad\qquad \text{thus } \mu_\kappa(\theta) - \frac{1}{2} \geq \frac{1}{2}\mu_\kappa(\theta) \geq \frac{\lambda\sqrt{\kappa}}{2\|e\|}.$$

Therefore, using Corollary 3.17 and since $r_1^\kappa(\theta) = \|e\|$

$$r_2^\kappa(\theta) \geq \|e\|(\mu_\kappa(\theta) - 1/2) \geq \frac{1}{2}\lambda\sqrt{\kappa}. \tag{28}$$

Denoting by c_1 and Σ the constant and function from Proposition 3.6, we conclude assuming $\lambda \leq c_0\sqrt{\kappa}$ for the last inequality

$$P(r_2^\kappa(\theta) \geq \frac{1}{2}\lambda\sqrt{\kappa}) \geq \sum_{\substack{e \in Z \\ \lambda\|e\| \leq c_0\sqrt{\kappa}}} |\{\theta \in \Theta_e^\kappa; \mu_\kappa(\theta) \geq \frac{\lambda\sqrt{\kappa}}{\|e\|}\}|$$

$$\geq \sum_{\substack{e \in Z \\ \lambda\|e\| \leq c_0\sqrt{\kappa}}} \frac{c_0}{\|e\|\lambda\sqrt{\kappa}} = \frac{c_0}{\lambda\sqrt{\kappa}}\Sigma(c_0\lambda^{-1}\sqrt{\kappa}) \geq c_0^2 c_1\lambda^{-2}. \qquad\square$$

3.5 Moments of the Voronoi radii

We conclude in this section the proof of Theorem 1.19, we estimate the moments of the euclidean radius $R_\kappa(\theta)$ and intrinsic radius $S_\kappa(\theta)$ of the set of strict $M_\kappa(\theta)$-Voronoi vectors. As usual θ is regarded as a random variable uniformly distributed over $\theta \in [0, \pi[$, and the the parameter $\kappa \geq 1$ is fixed. Our first step is to estimate the tail distributions of R_κ and S_κ.

Proposition 3.20. *There exists $C, c > 0$ such that for all $\kappa \geq 1$, and all $\lambda \in [C, c\kappa^{\frac{1}{2}}]$, one has*

$$c\lambda^{-2} \leq P(S_\kappa \geq \lambda) \leq C\lambda^{-2}, \qquad\qquad c\lambda^{-2} \leq P(R_\kappa \geq \lambda\kappa^{\frac{1}{2}}) \leq C\lambda^{-2}.$$

Proof. For readability, we omit the argument $\theta \in [0, \pi[$ of the functions involved. The estimate for S_κ immediately follows from the equivalence $S_\kappa/2 \leq \lambda_2^\kappa \leq S_\kappa$, see Corollary 3.5, and from the estimate of the tail distribution of λ_2^κ obtained in Corollary 3.11. Regarding R_κ we have $R_\kappa/2 \leq \max\{r_1^\kappa, r_2^\kappa\} \leq R_\kappa$, but $r_1^\kappa \in [1, c_1\kappa^{\frac{1}{2}}]$ by Lemma 3.4. Hence, assuming $C > 2c_1$, we obtain for $\lambda \geq C$ the implications

$$R_\kappa \geq \lambda\kappa^{\frac{1}{2}} \;\Rightarrow\; \max\{r_1^\kappa, r_2^\kappa\} \geq \tfrac{1}{2}\lambda\kappa^{\frac{1}{2}} > c_1\kappa^{\frac{1}{2}} \geq r_1^\kappa \;\Rightarrow\; \max\{r_1^\kappa, r_2^\kappa\} = r_2^\kappa \;\Rightarrow\; R_\kappa/2 \leq r_2^\kappa \leq R_\kappa.$$

The announced estimate for $P(R_\kappa \geq \lambda\kappa^{\frac{1}{2}})$ then follows from the estimates of the tail distribution of r_2^κ obtained in Lemma 3.15 and Corollary 3.19. $\qquad\square$

Proposition 3.20 implies in particular that $P(S_\kappa \geq c\kappa^{\frac{1}{2}}) > 0$ for all $\kappa \geq (C/c)^2$, but we also know that $S_\kappa \leq 2\kappa^{\frac{1}{2}}$ by Corollary 3.5. Hence $\|S_\kappa\|_{L^\infty([0,\pi[)} \approx \kappa^{\frac{1}{2}}$ as announced in Theorem 1.19. For a finite exponent $p \in [1, \infty[$, one has

$$\frac{1}{\pi}\|S_\kappa\|^p_{L^p([0,\pi[)} = \frac{1}{\pi}\int_0^\pi S_\kappa^p(\theta)d\theta = p\int_0^\infty P(S_\kappa \geq \lambda)\lambda^{p-1}d\lambda \tag{29}$$

In view of the uniform bounds $S_\kappa \in [c_2, 2\kappa^{\frac{1}{2}}]$ obtained in Corollary 3.5, and of the estimates of Proposition 3.20 which hold over a range $[C, c\kappa^{\frac{1}{2}}]$, we cut the integration range $]0, \infty[$ of (29, right) into four sub-intervals. Assuming without loss of generality that $c_2 \leq C$, $c \leq 2$, and $\kappa \geq (C/c)^2$

$$I_1 :=]0, C[, \; I_2 := [C, c\kappa^{\frac{1}{2}}[, \; I_3 := [c\kappa^{\frac{1}{2}}, 2\kappa^{\frac{1}{2}}[, \; I_4 := [2\kappa^{\frac{1}{2}}, \infty[.$$

Interval I_4 does contribute to the integral (29). Since I_1 is bounded independently of κ, its contribution to (29) also is. Since the tail distribution $P(S_\kappa \geq \lambda)$ is decreasing in λ, the contribution to (29) of $I_3 = [c\kappa^{\frac{1}{2}}, 2\kappa^{\frac{1}{2}}[= [\lambda_-, \lambda_+[$ is bounded by

$$\int_{\lambda_-}^{\lambda_+} P(S_\kappa \geq \lambda)\lambda^{p-1}d\lambda \leq P(S_\kappa \geq \lambda_-)\lambda_+^{p-1}(\lambda_+ - \lambda_-) \leq C\lambda_+^{-2}\lambda_+^p = Cc^{-2}2^p\kappa^{\frac{p}{2}-1}.$$

Finally, and most importantly, the contribution of I_2 is

$$p\int_C^{c\kappa^{\frac{1}{2}}} P(S_\kappa \geq \lambda)\lambda^{p-1}d\lambda \approx \int_C^{c\kappa^{\frac{1}{2}}} \lambda^{p-3}d\lambda \approx \begin{cases} 1 & \text{if } p < 2, \\ \ln\kappa & \text{if } p = 2, \\ \kappa^{\frac{p}{2}-1} & \text{if } p > 2, \end{cases}$$

for sufficiently large κ, with equivalence constants (9) depending only on p. This implies the estimate on $\|S_\kappa\|_{L^p([0,\pi[)}$ announced in Theorem 1.19. The case of $\|R_\kappa\|_{L^p([0,\pi[)}$ is analogous.

References

[AM11]   Ken Alton and Ian M Mitchell.   An Ordered Upwind Method with Precomputed
           Stencil and Monotone Node Acceptance for Solving Static Convex Hamilton-Jacobi
           Equations. *Journal of Scientific Computing*, 51(2):313–348, July 2011.

[BC10]   Fethallah Benmansour and Laurent D. Cohen. Tubular Structure Segmentation Based
           on Minimal Path Method and Anisotropic Enhancement. *International Journal of
           Computer Vision*, 92(2):192–210, March 2010.

[BCM15]   Jean-David Benamou, Francis Collino, and Jean-Marie Mirebeau. Monotone and Con-
           sistent discretization of the Monge-Ampere operator. *Mathematics of computation*,
           September 2015.

[BOZ04]   J Frederic Bonnans, Elisabeth Ottenwaelter, and Hasnaa Zidani. A fast algorithm for
           the two dimensional HJB equation of stochastic control. Technical report, 2004.

[BR06]   Folkmar Bornemann and Christian Rasch.   Finite-element Discretization of Static
           Hamilton-Jacobi Equations based on a Local Variational Principle. *Computing and
           Visualization in Science*, 9(2):57–69, June 2006.

[CCM14]   Da Chen, Laurent D. Cohen, and Jean-Marie Mirebeau.   Vessel Extraction Using
           Anisotropic Minimal Paths and Path Score. In *IEEE International Conference on
           Image Processing (ICIP 2014)*, paris, France, 2014. IEEE.

[Cia70]   Philippe G Ciarlet. Discrete maximum principle for finite-difference operators. *Ae-
           quationes Mathematicae*, 4(3):338–352, 1970.

[CL83]   MG Crandall and PL Lions. Viscosity solutions of Hamilton-Jacobi equations. 1983.

[CS92]   J H Conway and N J A Sloane. Low-Dimensional Lattices. VI. Voronoi Reduction
           of Three-Dimensional Lattices. *Proceedings of the Royal Society A: Mathematical,
           Physical and Engineering Sciences*, 436(1896):55–68, January 1992.

[Dro14]   Jérôme Droniou. Finite volume schemes for diffusion equations: Introduction to and
           review of modern methods. 24(08):1575–1619, 2014.

[FF02]   M Falcone and R Ferretti. Semi-Lagrangian schemes for Hamilton-Jacobi equations,
           discrete representation formulae and Godunov methods. *Journal of Computational
           Physics*, 175(2):559–575, 2002.

[FM13]   Jérôme Fehrenbach and Jean-Marie Mirebeau.   Sparse Non-negative Stencils for
           Anisotropic Diffusion. *Journal of Mathematical Imaging and Vision*, pages 1–25,
           2013.

[HW79]   GGH Hardy and E M Wright. *An Introduction to the Theory of Numbers*. Oxford
           Science Publications, 1979.

[Koc95]   Maciej Kocan. Approximation of viscosity solutions of elliptic partial differential
           equations on minimal grids. *Numerische Mathematik*, 72(1):73–92, November 1995.

[Mir13]   Jean-Marie Mirebeau. Efficient fast marching with Finsler metrics. *Numerische Math-
           ematik*, pages 1–43, 2013.

[Mir14] Jean-Marie Mirebeau. Anisotropic Fast-Marching on cartesian grids using Lattice Basis Reduction. *SIAM Journal on Numerical Analysis*, 52(4):1573–1599, January 2014.

[MW53] T S Motzkin and W Wasow. On the approximation of linear elliptic differential equations by difference equations with positive coefficients. *J. Math. Physics*, 31:253–259, 1953.

[Ngu04] P Nguyen. Low-dimensional lattice basis reduction revisited. *Algorithmic Number Theory*, 2004.

[Obe06] A M Oberman. Convergent Difference Schemes for Degenerate Elliptic and Parabolic Equations: Hamilton–Jacobi Equations and Free Boundary Problems. *SIAM Journal on Numerical Analysis*, 44(2):879–895, January 2006.

[Pet03] Jean Petitot. The neurogeometry of pinwheels as a sub-Riemannian contact structure. *Journal of Physiology-Paris*, 97(2-3):265–309, March 2003.

[SBD$^+$15] Gonzalo Sanguinetti, Erik Bekkers, Remco Duits, Michiel Janssen, Alexey Mashtakov, and Jean-Marie Mirebeau. Sub-Riemannian Fast Marching in SE(2). *arXiv.org*, August 2015.

[Sel74] Eduard Selling. Ueber die binären und ternären quadratischen Formen. *Journal fur die Reine und Angewandte Mathematik*, 77:143–229, 1874.

[SV01] J. A. Sethian and A. Vladimirsky. Ordered upwind methods for static Hamilton-Jacobi equations. *Proceedings of the National Academy of Sciences of the United States of America*, 98(20):11069–11074 (electronic), 2001.

[SV03] James A. Sethian and Alexander Vladimirsky. Ordered upwind methods for static Hamilton-Jacobi equations: theory and algorithms. *SIAM Journal on Numerical Analysis*, 41(1):325–363, 2003.

[Tsi95] J.N. Tsitsiklis. Efficient algorithms for globally optimal trajectories. 40(9):1528–1538, September 1995.

[Vla08] Alexander Vladimirsky. Label-setting methods for Multimode Stochastic Shortest Path problems on graphs. *Mathematics of Operations Research*, 33(4):821–838, 2008.

[Wei96] J Weickert. Theoretical foundations of anisotropic diffusion in image processing. *Computing-Wien-Supplements*, 1996.

[Wei98] Joachim Weickert. *Anisotropic diffusion in image processing*, volume 1. Teubner Stuttgart, 1998.

A Discretizations of anisotropic PDEs based on stencils

We establish in this appendix the results announced in §1.1, on the consistency of numerical schemes based on D-diffusion and M-eikonal stencils, and their structure preservation properties. These results are fairly classical, but the proofs are provided for completeness. Results of similar nature can be found in [Wei96] for diffusion with a focus on image processing, [Cia70] for static diffusion problems, or [SV03, Vla08] for eikonal equations. Note also the restriction to two

dimensional discretizations on cartesian grids is not necessary, and that in both cases extensions to arbitrary dimensional discretizations on unstructured point sets can easily be formulated.

A.1 Diffusion stencils

We first establish that the Consistency axiom of D-diffusion stencils, is indeed equivalent to the consistency of two discretizations of $\|\nabla u(x)\|_D^2$ and $\mathrm{Tr}(\nabla^2 u(x))$ respectively.

Proof of proposition 1.2, on property Consistency. Consider a finitely supported $\gamma : \mathbf{Z} \to \mathbb{R}$ obeying property Symmetry of Definition 1.1, and define $D := \sum_{e \in \mathbf{Z}} \gamma(e) e \otimes e$. Consider also the linear and quadratic functions defined for all $x \in \mathbb{R}^2$ by

$$u_1(x) := C + \langle L, x \rangle, \qquad\qquad u_2(x) := C + \langle L, x \rangle + \frac{1}{2}\langle x, Qx \rangle,$$

where $C \in \mathbb{R}$, $L \in \mathbb{R}^2$ and $Q \in S_2$. Then by elementary linear algebra we obtain

$$\sum_{e \in \mathbf{Z}} \gamma(e)(u_1(x+e) - u_1(x))^2 = \sum_{e \in \mathbf{Z}} \gamma(e)\langle L, e \rangle^2 = \sum_{e \in \mathbf{Z}} \gamma(e)\langle L, e \otimes e\, L \rangle = 2\langle L, DL \rangle.$$

Similarly, using the symmetry of the weights,

$$2\sum_{e \in \mathbf{Z}} \gamma(e)(u_2(x+e) - u_2(x)) = \sum_{e \in \mathbf{Z}} \gamma(e)\left(u(x+e) - 2u(x) + u(x-e)\right)$$

$$= \sum_{e \in \mathbf{Z}} \gamma(e)\langle e, Qe \rangle = \sum_{e \in \mathbf{Z}} \gamma(e)\,\mathrm{Tr}(e \otimes e\, Q) = 2\,\mathrm{Tr}(DQ).$$

Thus the Consistency axiom implies properties Consistency' and Consistency". Conversely, observing that the quadratic form $L \in \mathbb{R}^2 \mapsto \langle L, DL \rangle$ (resp. the linear form $Q \in S_2 \mapsto \mathrm{Tr}(DQ)$) uniquely determines the matrix $D \in S_2$, we obtain that Consistency' (resp. Consistency") implies Consistency, which concludes the proof. $\qquad\square$

Our next objective is to establish Proposition 1.4 on the stability of the explicit and implicit time steps for anisotropic diffusion. A definition and two intermediate results are required. In this paper, the term *operator* always refers to a continuous map from $L^\infty(\mathbb{Z}^2)$ to itself, and the maximum principle is understood in the sense of Definition 1.3.

Definition A.1. *An operator L preserves constants iff $Lu = u$ for any identically constant $u \in L^\infty(\mathbb{Z}^2)$. An operator L is non-negative iff $u \geq 0 \Rightarrow Lu \geq 0$ for any $u \in L^\infty(\mathbb{Z}^2)$.*

Lemma A.2. *A linear operator obeys the maximum principle iff it is non-negative and preserves constants.*

Proof. Denote by B this operator, and by u an arbitrary element of $L^\infty(\mathbb{Z}^2)$. Sups and Infs are taken over the set $\mathbb{Z}^2$.

Proof of implication $\Rightarrow$. If u is constant, then from $\inf u \leq Bu \leq \sup u$ we obtain $Bu = u$. If u is non-negative, then from $Bu \geq \inf u \geq 0$ we obtain that Bu is non-negative.

Proof of implication $\Leftarrow$. For any $u \in L^\infty(\mathbb{Z}^2)$, on obtains $0 \leq B(u - \inf u) = Bu - \inf u$, and likewise $0 \leq B(\sup u - u) = \sup u - Bu$, thus $\inf u \leq Bu \leq \sup u$ as announced. $\qquad\square$

The next corollary is a states a simple necessary and sufficient condition for the maximum principle for linear and translation invariant operators.

Corollary A.3 (Stability criterion)**.** *Consider $\eta \in L^1(\mathbb{Z}^2)$, and the operator defined by $A_\eta u(x) := \sum_{e \in \mathbb{Z}^2} \eta(e)u(x + e)$. Then A_η obeys the maximum principle iff $\eta \geq 0$ and $\sum_{e \in \mathbb{Z}^2} \eta(e) = 1$.*

Proof. The non-negativity of A_η is clearly equivalent to $\eta \geq 0$, and the fact that it preserves constants to $\sum_{e \in \mathbb{Z}^2} \eta(e) = 1$. $\qquad\square$

Finally we establish Proposition 1.4. For that purpose we separate the claims on the explicit diffusion time step $\mathrm{Id} + \delta L_\gamma$ and the implicit one $(\mathrm{Id} - \delta L_\gamma)^{-1}$.

Proof of Proposition 1.4, on explicit diffusion. Let $\gamma : \mathbf{Z} \to \mathbb{R}$ be finitely supported, and let $\delta > 0$. Observing that

$$(\mathrm{Id} + \delta L_\gamma)u(x) = \left(1 - \delta \sum_{e \in \mathbf{Z}} \gamma(e)\right)u(x) + \sum_{e \in \mathbf{Z}} \gamma(e)u(x + e),$$

and applying Lemma A.3 we obtain the equivalence of the two points of Proposition 1.4, as well as the optimality of the CFL condition $0 \leq \delta \sum_{e \in \mathbf{Z}} \gamma(e) \leq 1$. $\qquad\square$

Proof of Proposition 1.4, on implicit diffusion. Let $\gamma : \mathbf{Z} \to \mathbb{R}$ be finitely supported and non-negative, and let $\delta > 0$. Consider the operator defined for $u : \mathbb{Z}^2 \to \mathbb{R}$ and $x \in \mathbb{Z}^2$ by

$$Au(x) := \lambda \sum_{e \in \mathbf{Z}} \gamma(e)u(x + e), \qquad \text{where } \lambda := \left(1 + \delta \sum_{e \in \mathbf{Z}} \gamma(e)\right)^{-1}.$$

Then A is non-negative: $Au \geq 0$ if $u \geq 0$, and A is strictly Lipschitz in the L^∞ norm: $\|Au\|_\infty \leq (1 - \lambda)\|u\|_\infty$. As a result, the operator associated to the implicit time step

$$B := (\mathrm{Id} - \delta L_\gamma)^{-1} = \lambda(\mathrm{Id} - A)^{-1} = \lambda \sum_{k \geq 0} A^k$$

is well defined in $L^\infty(\mathbb{Z}^2)$ and non-negative. In addition the operator $(\mathrm{Id} - \delta L_\gamma)^{-1}$ preserves constants, since its inverse does, hence it obeys the maximum principle by Lemma A.2. $\qquad\square$

A.2 Eikonal stencils

We provide for completeness the proof Proposition 1.9, stating that the causality property of the Hopf-Lax update operator is equivalent to a geometrical acuteness property of the stencils. We mostly follow the steps of the appendix of [SV03].

Lemma A.4. *Let $N \in S_2^+$. Then the following are equivalent:*

(i) N has non-negative entries.

(ii) $N\xi$ has positive entries whenever $\xi \in \mathbb{R}^2$ has positive entries.

Proof. Proof that (i) $\Rightarrow$ (ii). Since N has non-negative coefficients, the product $N\xi$ also has. Observing that N has positive diagonal entries, since it is positive definite, we obtain that $N\xi$ has positive entries. Proof that (ii) $\Rightarrow$ (i). Let $1 \leq i \leq 2$. Choosing $\xi = (\varepsilon, \cdots, \varepsilon, 1, \varepsilon, \cdots, \varepsilon)$, where the coefficient 1 is in the i-th position, and letting $\varepsilon \to 0$, we obtain that the i-th column of N is non-negative. The result follows. $\qquad\square$

Lemma A.5. *Let $N \in S_2^+$, and for all $u \in \mathbb{R}^2$ let*

$$\lambda_N(u) := \min_{\xi \in \Xi} \|\xi\|_N + \langle \xi, u \rangle, \tag{30}$$

where $\Xi := \{\xi \in \mathbb{R}_+^2; \langle \xi, 1 \rangle = 1\}$ and $1 := (1,1) \in \mathbb{R}^2$. Then the following are equivalent:

(i) *the matrix N has non-negative entries.*

(ii) *for any $u \in \mathbb{R}^2$ and any $i \in \{1, 2\}$, such that the minimizer in (30) obeys $\xi_i > 0$, one has $\lambda_N(u) > u_i$.*

Proof. Note that the quantity $\lambda_N(u)$ is defined as the minimum (30) of a strictly convex functional on a compact interval, hence the minimizer exists and is unique. Let $u \in \mathbb{R}^2$, and let us assume that the minimizer ξ in (30) has positive entries. By the Kuhn-Tucker optimality conditions, there exists $\lambda \in \mathbb{R}$ such that

$$\lambda 1 = N\xi / \|\xi\|_N + u. \tag{31}$$

Taking the scalar product with ξ we obtain

$$\lambda = \lambda \langle 1, \xi \rangle = \langle \xi, N\xi \rangle / \|\xi\|_N + \langle \xi, u \rangle = \|\xi\|_N + \langle \xi, u \rangle = \lambda_N(u).$$

Thus combining the last two equations

$$\lambda_N(u) 1 - u = N\xi / \|\xi\|_N. \tag{32}$$

Proof that (i) $\Rightarrow$ (ii). The announced result is clear if the minimizer ξ in (30) equals $(1, 0)$ or $(0, 1)$, hence as above we can assume that ξ has positive entries. By assumption N has non-negative entries, hence $N\xi$ has positive entries by Lemma A.4, thus also $\lambda_N(u) 1 - u$ by (32), which concludes the proof of this implication.

Proof that not (i) $\Rightarrow$ not (ii). Assume that N has a negative entry. By Lemma A.4 there exists ξ with positive entries such that $N\xi$ has at least one non-positive entry. Up to multiplication by a positive constant, we may assume that $\xi \in \Xi$. Let $u := -N\xi / \|\xi\|_N$. Since (30) is the optimization of a strictly convex functional over a convex set, its minimizer is characterized by the first order Kuhn-Tucker condition (31). Hence $\lambda = \lambda_N(u) = 0$ and as announced we obtain not (ii). $\square$

Proof of Proposition 1.9, Acuteness is equivalent to Causality. By construction, one has

$$\Lambda u(x) = \min_{0 \le i < r} \lambda_{N_i}(u(y_i), u(y_{i+1})), \qquad \text{where } N_i := \begin{pmatrix} \langle e_i, M e_i \rangle & \langle e_i, M e_{i+1} \rangle \\ \langle e_i, M e_{i+1} \rangle & \langle e_{i+1}, M e_{i+1} \rangle \end{pmatrix},$$

and $y_i := x + e_i$. The matrices N_i, $0 \le i < r$, are positive definite by construction, and the Acuteness geometric property is equivalent to the non-negativity of their coefficients. The first implication Acuteness $\Rightarrow$ Causality, is thus a direct consequence of Lemma A.5.

Proof that not Acuteness $\Rightarrow$ not Causality. Without loss of generality, we may assume that $\langle e_0, M e_1 \rangle < 0$, hence that N_0 has one negative entry. Choose $u(y_0)$ and $u(y_1)$ so as to contradict point (ii) of Lemma A.5, in other words such that defining $\lambda := \lambda_{N_0}(u(y_0), u(y_1))$ one has $\lambda < \max\{u(y_0), u(y_1)\}$ and the associated minimizer $\xi_0 = (1 - t_0, t_0)$ for (30) has positive entries. Note also that $\lambda < \min\{\|e_0\|_M + u(y_0), \|e_0\|_M + u(y_0)\}$, since these values correspond to $t = 0$ and $t = 1$.

For any $0 \le i < r$, one has $\lambda_{N_i}(u(y_i), u(y_{i+1})) \ge \min\{u(y_i), u(y_{i+1})\}$, and $\lambda_{N_i}(u(y_i), u(y_{i+1})) \to \|e_i\|_M + u(y_i)$ as $u(y_{i+1}) \to \infty$. Hence choosing $u(y_i)$ sufficiently large, for all $i \notin \{0, 1\}$, we obtain $\Lambda u(x) = \lambda$, this minimum being attained solely for $i = 0$ and $t = t_0$. This contradicts the causality property, and concludes the proof. $\square$