



# Single pass eikonal solver in tilted transversally anisotropic media

Ludovic Métivier, Jean-Marie Mirebeau, François Desquilbet

## ► To cite this version:

Ludovic Métivier, Jean-Marie Mirebeau, François Desquilbet. Single pass eikonal solver in tilted transversally anisotropic media. 2022. hal-03655290v1

**HAL Id: hal-03655290**

**<https://hal.science/hal-03655290v1>**

Preprint submitted on 29 Apr 2022 (v1), last revised 3 May 2022 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Single pass eikonal solver in tilted transversally anisotropic media

Francois Desquilbet\*, Ludovic Metivier†, Jean-Marie Mirebeau‡

April 29, 2022

## Abstract

We present a numerical scheme to solve the eikonal equation in a Tilted Transversely Isotropic (TTI) medium. The solution to this equation corresponds to the first arrival time of seismic pressure waves in the high frequency asymptotic regime, whose propagation speed is neither isotropic nor elliptic. Instead, the speed profile is characterized by a fourth degree polynomial equation in a rotated frame, defined in terms of the Thomsen or Hooke elasticity coefficients of the geophysical medium.

We show that TTI eikonal equations can be expressed as the maximum or minimum of a family of Riemannian eikonal equations, for which efficient discretizations are known. Based on this observation, we propose an original scheme that is causal, thus solvable in a single pass over the domain, and Eulerian, hence also mapping well to massively parallel architectures. Numerical experiments illustrate the method's accuracy, speed and robustness, on both a problem with analytical solution and a realistic synthetic instance, and compare a CPU with a GPU implementation, with the GPU being fifty times faster than the CPU implementation.

**Keywords :** Anisotropic eikonal equation, tilted transversely isotropic medium, fast marching.  
**AMS subject classification:** 65N06, 49L25, 35F30

## 1 Introduction

The eikonal Partial Differential Equation (PDE) characterizes the first arrival time of a front, whose propagation speed is locally dictated by a metric. Classical examples include isotropic metrics, which define a propagation speed depending only on the position of the front, as well as Riemannian metrics, whose propagation speed also depends on the normal to the front according to an ellipsoidal profile. In this paper we focus on the more complex *tilted transversely isotropic* (TTI) model, which commonly accounts for the velocity profiles of seismic pressure waves in complex media [LBBMV17]. Such anisotropy may originate from a variety of causes, at various physical scales: from the atomic layout in crystals, through the small scale layered structure of rocks produced by sedimentation, to homogenisation effects along geophysical fault lines [BC91, CMA<sup>+</sup>20]. We introduce a numerical scheme to solve the eikonal equation in TTI media, which is both very general - able to handle anisotropy of arbitrary strength, and to include the effects of topography, see Remark A.2 - and highly efficient - solvable in a single pass over the domain, and efficiently portable to massively parallel accelerators. The scheme

---

\*Univ. Grenoble-Alpes, LJK, F-38000, Grenoble, France

†Univ. Grenoble-Alpes, LJK, CNRS, ISTerre, F-38000, Grenoble, France

‡Univ. Paris-Saclay, CNRS, Centre Borelli, 91190, Gif-sur-Yvette, France

requires a Cartesian discretization grid, involves adaptive discretization stencils designed using algorithmic geometry for greater efficiency, and relies on a characterization of the TTI speed profile as a union or an intersection of ellipsoids, depending on the PDE coefficients, see Fig. 1. We establish the wellposedness of the method, including the scheme causality, monotony, the quasi-convexity or quasi-concavity of the involved optimization problems, and the convergence analysis. Numerical experiments illustrate our results, and include a comparison of a CPU and a GPU implementation, the validation of second order accuracy in synthetic test cases achieved using source factorization and multi-scale computations, and the fast resolution of a large and realistic three dimensional instance.

For concreteness, let us readily state the PDE that is addressed in this paper. Denote by  $\Omega \subset \mathbb{R}^3$  an open connected and bounded domain, by  $\sigma = (a, b, c, d, e) \in C^0(\overline{\Omega}, \mathbb{R}^5)$  some coefficients which are subject to the admissibility condition described in Theorem 1.2, and by  $R \in C^0(\overline{\Omega}, \text{GL}_3(\mathbb{R}))$  a continuous field of invertible matrices. Our objective is to numerically compute the viscosity [BCD08] solution  $u : \Omega \rightarrow \mathbb{R}$  of the following static first order Hamilton-Jacobi-Bellman PDE :

$$ap_r^4 + bp_z^4 + cp_r^2 p_z^2 + dp_r^2 + ep_z^2 = 1, \quad \text{where } (p_x, p_x, p_z) = R\nabla u \text{ and } p_r^2 := p_x^2 + p_y^2, \quad (1)$$

on  $\Omega \setminus \{q_0\}$ , subject to the additional condition  $u(q_0) = 0$  at a point source  $q_0 \in \Omega$ , and to outflow boundary conditions on  $\partial\Omega$ . The coefficients  $\sigma = (a, b, c, d, e)$  are derived from the local geophysical properties of the medium, and define an anelliptic (non-Riemannian) speed propagation profile, see Appendix A. The model is said *tilted* in view of the coordinate transformation  $R$ , which is usually a rotation, and *transversely isotropic* in view of symmetry in  $p_x$  and  $p_y$ . A Riemannian, or elliptic, geometry is recovered in the special case where the equation is quadratic, i.e.  $a = b = c = 0$ . Following the geophysical terminology, we refer to (1, left) as the P-SV equation and note that it defines two *slowness* surfaces, see Fig. 1, corresponding to the pressure and vertical shear wave propagation. Only the inner surface, associated with pressure waves which are the fastest, is considered in this paper.

Our numerical approach involves rephrasing the highly non-linear eikonal PDE (1), defined by a fourth degree polynomial, as a maximum or a minimum of a varying family of Riemannian eikonal PDEs, defined by quadratic polynomials and for which efficient numerical schemes have been developed [Mir19], see Fig. 1 and Section 1.1. For this reason, this work is related to multi-stencil fast marching methods [HF07], and more generally to discretizations of Hamilton-Jacobi-Bellman equations written in the Bellman extremal form [BBM20, BM21], see Section 4. For concreteness, we readily state our numerical scheme, which is presented in more detail in Sections 1.1 to 1.3 : find the solution to a finite differences equation denoted  $\mathfrak{F}u = 1$ , whose unknown  $u : \Omega_h \rightarrow \mathbb{R}$  is discretized on the Cartesian grid  $\Omega_h := \Omega \cap h\mathbb{Z}^d$  of scale  $h > 0$  with the appropriate boundary conditions, and where

$$\mathfrak{F}u(q) := \underset{\alpha \in [\alpha_*, \alpha^*]}{\text{mix}} \frac{1}{\mu(\alpha)} \sum_{1 \leq i \leq I} \rho_i(\alpha) \max \left\{ 0, \frac{u(q) - u(q + he_i)}{h}, \frac{u(q) - u(q - he_i)}{h} \right\}^2. \quad (2)$$

The notation “mix” stands for the (max) or (min) operator, if the P slowness surface is obtained as an *intersection* or as a *union* of ellipses respectively, see Fig. 1. The optimization interval  $[\alpha_*, \alpha^*]$  and multiplier  $\mu(\alpha) > 0$ , which are related to the anisotropy bounds and dilation coefficients of the ellipses respectively, have explicit algebraic expressions presented in Section 1.1. The weights  $\rho_i(\alpha) \geq 0$  and offsets  $e_i \in \mathbb{Z}^d$  are obtained using Selling’s matrix decomposition [Sel74], similarly to the Riemannian scheme [Mir19], see Section 1.2. The scheme properties, and two

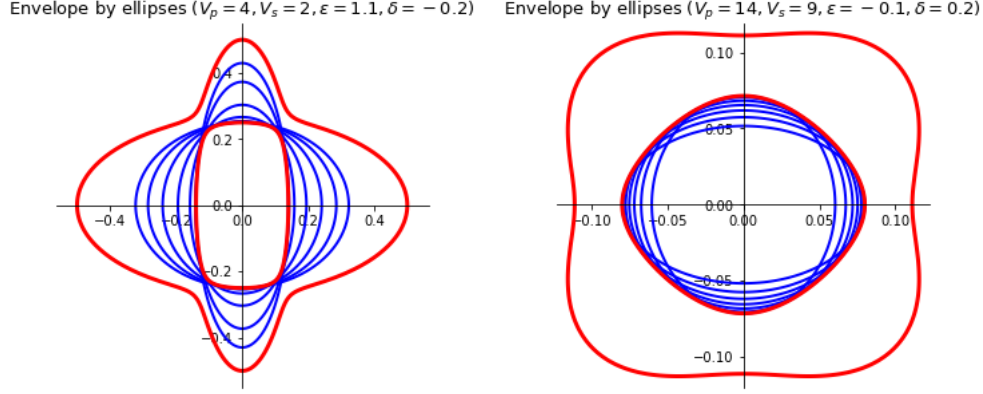


Figure 1: Slowness surfaces (red) defined by equation (1, left), in the  $(p_r, p_z)$  plane. The coefficients  $(a, b, c, d, e)$  are derived from the supplied Hooke parameters  $(V_p, V_s, \varepsilon, \delta)$ . Only the inner slowness surface is considered, and our numerical method involves its approximation by an intersection of ellipses (left) or a union of ellipses (right), shown blue. Subfigures (left) and (right) correspond respectively to the (max) and (min) alternative cases in (2) and Theorem 1.3.

strategies for the numerical optimization over  $\alpha \in [\alpha_*, \alpha^*]$ , are investigated in Section 1.3. The parameters  $\alpha_*, \alpha^*, \mu(\alpha), \rho_i(\alpha)$  are derived from the coefficients  $a, b, c, d, e, R$  of the TTI eikonal PDE (1), and similarly they implicitly depend on the position  $q \in \Omega$  according to the variations of the medium in the domain, see Remark 1.1.

The general framework of this study is seismic imaging: inferring the geophysical properties of the subsurface and its structure from the physical recordings of the seismic waves. In this context, accessing to travel-times is crucial in many steps of the workflow: macro-velocity model building through tomography, high resolution reflectivity estimation through migration techniques, and quality control along different stages of waveform based inversion techniques to estimate the time-shifts between recorded and simulated data. For this reason, efficient numerical methods to solve the high frequency asymptotics of the elastic wave equations in 3D general media are of particular interest for the seismic imaging community.

Distance maps are ubiquitous in mathematics and their applications, hence a sustained research effort and a wide variety of methods have been developed for their computation. The solution to the eikonal PDE (1) falls in this framework, since it admits an interpretation as the geodesic distance map from the source point  $q_0$  and with respect to a metric defined in terms of the parameters  $(a, b, c, d, e)$  and  $R$ , see Remark 1.5. Graph based methods can compute distances while avoiding the PDE formalism [CHK13], but they often lack the stability and high order accuracy required for seismic imaging applications. Some other approaches exploit indirect connections with different PDEs, such as the heat method which is based on the small time asymptotics of the heat or Poisson kernels [CWW13]; however, it is limited to metrics featuring a quadratic structure, and does not appear to scale well to three dimensional problems. Yet another approach to the computation of geodesics and geodesic distances is ray tracing [Sla03].

In the rest of this discussion we limit our attention to numerical methods which solve directly the eikonal PDE, either the one (1) associated to the TTI geometry, or a variant defined by another metric structure. We categorize these methods based on two criteria: *causal (single-pass) vs multi-pass solvers*, and *Eulerian vs semi-Lagrangian* approximation schemes - the method here proposed being causal and Eulerian. The first distinction is tied to a property, referred to as

causality, of the coupled system of non-linear equations arising from the PDE discretization.

- Causal, single pass methods, are often referred to as fast marching methods (FMM), see Algorithm 1. The causality property is the translation at the discrete level of a principle underlying the front propagation: the front arrival time at a given point only depends on *earlier* arrival times, see Definition 1.6. Common advantages of FMMs include faster computation times (on sequential machines), easier back propagation (thanks to the triangular structure of the Jacobian of the scheme), and opportunities for modification (adaptive stopping criteria, high order schemes, etc), see the discussion in [DCC<sup>+</sup>21]. Originally limited to isotropic eikonal equations [Set96], FMMs have been generalized to a variety of metrics [SV03, Mir14b, Mir18, Mir19, DCC<sup>+</sup>21].
- Multi-pass methods rely on fast sweeping [Zha05], the fast iterative method [FKW13], or adaptive Gauss-Siedel iterations [BR06], to solve the discretized PDE. These iterative methods miss some of the advantages of FMMs, but also avoid the severe constraints associated with the design of a causal scheme. This shift in compromises enables a wider variety of numerical approaches, and thus possibly (if properly exploited) methods with narrower stencils, or addressing more complex geometries. They are also easier to parallelize.

The second distinction is between Eulerian and semi-Lagrangian PDE discretization schemes.

- Eulerian discretizations use finite differences (or finite elements, possibly discontinuous) to approximate the derivatives of the unknown arrival time function, and to produce a consistent approximation of the eikonal PDE operator. A variety of Eulerian schemes have been developed, for isotropic [Set96, HF07], Riemannian [Mir19], TTI [LBBMV17], and curvature penalized [Mir18] geometry models.
- Semi-Lagrangian schemes mimic Bellman’s optimality principle at the discrete level, which is derived from the shortest path interpretation of the solution to the eikonal equation, see e.g. [BR06, SV03, Mir14b, Mir14a]. These methods require maintaining a complex neighborhood structure around each point, and for this reason implementing them on GPUs, while feasible [FKW13], is more cumbersome and usually less efficient than for Eulerian schemes. In addition, implementing semi-Lagrangian schemes for TTI and related models in seismology requires solving complicated algebraic equations, due to the high degree of the PDEs (1) or (72), such as an optimization problem subject to a polynomial constraint of degree six in three variables in [DCC<sup>+</sup>21]. This has a significant computational cost and typically requires double precision floating point arithmetic for stability.

Note that in the context of Eulerian methods, causality can be rephrased as a structural constraint on the numerical scheme, see Definition 1.6. Only few Eulerian schemes obey this condition beyond the standard isotropic one [Set96, Mir19, Mir18] and in particular the discretization of the TTI eikonal PDE proposed in [LBBMV17] is not causal, in contrast to the one presented in this paper. For comparison, causality in the context of semi-Lagrangian schemes is equivalent to a geometric acuteness property of the stencils [SV03], which can be ensured by refinement in two dimensions [KS98, Mir14b], but in three dimensions either requires a Riemannian structure [Mir14a], or poses a limit on the strength of the anisotropy [DCC<sup>+</sup>21], or requires impractically large stencils [SV03].

**Summary of contributions.** We present a *causal* and *Eulerian* discretization of the TTI eikonal PDE. Our approach is based on a new methodology, expressing the TTI speed profile as

a union or an intersection of ellipsoids, established in Theorem 1.3. The scheme update operator is defined by a one-dimensional optimization problem, which can be efficiently solved numerically thanks to a quasi-convexity or quasi-concavity property established in Theorem 1.11. A proof of convergence of the scheme numerical solutions is presented in Theorem 4.1, together with regularity and growth estimates established using some fine properties of Selling’s decomposition, which is a tool from discrete geometry involved in the scheme construction. The resulting scheme can be solved in a single pass over the domain using a CPU solver, but a massively parallel GPU solver is also demonstrated. Numerical experiments include a smooth synthetic test case for validating the scheme accuracy, as well as a large realistic instance.

**Paper organization.** The rest of this introduction is organized as follows: we present in Section 1.1 a reformulation of the TTI metric as an extremum of a family of Riemannian metrics, we recall in Section 1.2 an efficient discretization of the Riemannian eikonal PDE, and we combine in Section 1.3 the previous elements to obtain a discretization of the TTI eikonal equation (1). Beyond the introduction, the rest of this paper is organized as follows. Section 2 establishes the results, announced in Section 1.1, relating TTI and Riemannian geometry. We prove in Section 3 a property of the update operator of our scheme, announced in Section 1.3, which makes it numerically easy to evaluate. We establish in Section 4 some regularity and growth estimates for the scheme solutions, and prove their convergence to the viscosity solution of the PDE (1). Numerical experiments are presented in Section 5. Appendix A describes Hooke tensors, Thomsen’s elastic parameters, and their relation to the coefficients of (1). Appendix B describes Selling’s decomposition, a tool from discrete geometry used in the design of our numerical scheme in Section 1.3. Appendix C discusses various heuristic enhancements designed to improve the accuracy of the scheme solutions.

**Remark 1.1** (Varying material coefficients). *For readability, we present in Sections 1.1 to 1.3 the construction of our numerical scheme in the setting where the coefficients  $\sigma = (a, b, c, d, e)$  and the linear transformation  $R$  defining the eikonal PDE (1) are fixed over the domain  $\Omega$ . It must be clear however that, in the intended applications including the numerical experiments in Section 5, the parameters  $\sigma : \Omega \rightarrow \mathbb{R}^5$  and  $R : \Omega \rightarrow \text{GL}_3(\mathbb{R})$  vary over the domain, and the definitions below are applied independently at each discretization point.*

**Notations.** We denote by  $\text{CC}_x(X)$  the connected component of a point  $x$  in the topological set  $X$ . Position variables are usually named  $q$ , impulsions named  $p$ , and velocities named  $v$ . The set of non-negative reals is denoted  $\mathbb{R}_+ := [0, \infty[$ . Let  $\langle \cdot, \cdot \rangle$  denote the Euclidean scalar product,  $|\cdot|$  the Euclidean norm.  $S_d^{++}$  stands for the set of symmetric positive definite matrices of shape  $d \times d$ , and we let  $\|v\|_D := \sqrt{\langle v, Dv \rangle}$  for any  $D \in S_d^{++}$ ,  $v \in \mathbb{R}^d$ .

## 1.1 The eikonal equation associated to a TTI model

We study in this subsection the algebraic structure of the TTI eikonal PDE (1): we characterize in Theorem 1.2 a family of coefficients for which it is physically meaningful and mathematically well posed, and we reformulate in Theorem 1.3 and Corollary 1.4 the PDE operator in a form related to the Riemannian setting that is amenable to discretization. Theorems 1.2 and 1.3 are proved in Section 2, and Corollary 1.4 is established in Section 3.2.

Our first step is to disambiguate the PDE (1), by distinguishing the role of the inner slowness surface. For that purpose we introduce given coefficients  $\sigma = (a, b, c, d, e) \in \mathbb{R}^5$  the quadratic

function  $\mathcal{Q}_\sigma$  and the set  $\mathcal{B}_\sigma$  defined as follows

$$\mathcal{Q}_\sigma(r, z) := ar^2 + bz^2 + crz + dx + ez, \quad (3)$$

$$\mathcal{B}_\sigma := \text{CC}_0\{(p_x, p_y, p_z) \in \mathbb{R}^3; \mathcal{Q}_\sigma(p_x^2 + p_y^2, p_z^2) \leq 1\}. \quad (4)$$

By considering only the connected component of the origin, denoted  $\text{CC}_0$  in (4), we obtain that  $\partial\mathcal{B}_\sigma$  is the inner slowness surface defined by  $\mathcal{Q}_\sigma$ , as illustrated on Fig. 1. We assume that the coefficients  $\sigma$  are admissible in the sense of Theorem 1.2 below, in such way that the set  $\mathcal{B}_\sigma$  is compact and convex, and therefore this construction is well posed and physically meaningful. As a result, the eikonal PDE (1) can be reformulated in an unambiguous way

$$\mathcal{F}_\sigma^*(R\nabla u) = 1, \quad \text{where } \mathcal{F}_\sigma^*(p) := \min\{\nu > 0; p/\nu \in \mathcal{B}_\sigma\}, \quad (5)$$

with  $\mathcal{F}_\sigma^*(0) := 0$  by convention. In the geophysical context, the reformulation (5) characterizes the travel-time of the pressure wave (the fastest wave), and disregards the shear wave. Note that the eikonal PDE (5) is imposed on  $\Omega \setminus \{q_0\}$ , similarly to (1), and is combined with the point source constraint  $u(q_0) = 0$  and outflow boundary conditions on  $\partial\Omega$ . The PDE (5) admits a unique viscosity solution [BCD08], provided<sup>1</sup> the parameters  $\sigma$  and  $R$  vary continuously over the domain  $\bar{\Omega}$  and obey the admissibility condition described below in equations (7) and (8). The solution  $u(q)$  at  $q \in \Omega$  can be characterized as the minimal path length from the source point  $q_0$ , as measured by the Finsler metric  $\mathcal{F}_\sigma$  dual to  $\mathcal{F}_\sigma^*$ , see Remark 1.5.

A fundamental special case of TTI geometry is when the equation is derived from the coefficients  $(c_{11}, c_{13}, c_{33}, c_{44})$  of a Hooke tensor with the appropriate hexagonal symmetry, see Appendix A. In this case the eikonal equation (1) reads:

$$-c_{11}c_{44}p_r^4 - c_{33}c_{44}p_z^4 + (2c_{13}c_{44} + c_{13}^2 - c_{11}c_{33})p_r^2p_z^2 + (c_{11} + c_{44})p_r^2 + (c_{33} + c_{44})p_z^2 = 1, \quad (6)$$

with  $p_r^2 := p_x^2 + p_y^2$ . In other words,  $\mathcal{Q}_\sigma(p_r^2, p_z^2) = 1$  where the parameters  $\sigma = (a, b, c, d, e)$  can be recovered by identification with (1), namely

$$\sigma = (-c_{11}c_{44}, -c_{33}c_{44}, 2c_{13}c_{44} + c_{13}^2 - c_{11}c_{33}, c_{11} + c_{44}, c_{33} + c_{44}). \quad (7)$$

Our first result uses the algebraic properties of Hooke tensors and a criterion on the coefficients to ensure that the ball  $\mathcal{B}_\sigma$  is well shaped. Some limit cases are illustrated on Fig. 5.

**Theorem 1.2.** *Define  $C_{\text{adm}} \subset \mathbb{R}^4$  as the set of Hooke tensor coefficients  $c_{11}, c_{13}, c_{33}, c_{44}$  obeying*

$$c_{11} > c_{44}, \quad c_{33} > c_{44}, \quad c_{44} > 0, \quad c_{13} + c_{44} > 0, \quad c_{11}c_{33} > c_{13}^2, \quad (8)$$

*which is open and convex. The PDE coefficients  $\sigma \in \mathbb{R}^5$  are said admissible if they take the form (7) for some  $(c_{11}, c_{13}, c_{33}, c_{44}) \in C_{\text{adm}}$ , and in that case the ball  $\mathcal{B}_\sigma$  is compact and convex.*

Geophysical elasticity properties are also often described through the Thomsen parameters  $(V_P, V_S, \varepsilon, \delta)$ , but this turns out to be equivalent to specifying  $(c_{11}, c_{13}, c_{33}, c_{44})$ , with explicit conversion formulas between the parameters, see Appendix A. We checked that all the material elasticity parameters listed in [Tho86] obey the admissibility condition. In addition, the convexity of the set  $C_{\text{adm}}$  means that admissibility is preserved when one ‘interpolates’ between those materials, hence the criterion of Theorem 1.2 does not appear to be excessively restrictive for applications in geophysics. As discussed in Remark 2.13, the eikonal PDE (1) could be studied under weaker assumptions, but in that case the slowness surfaces may not be separated see Fig. 5,

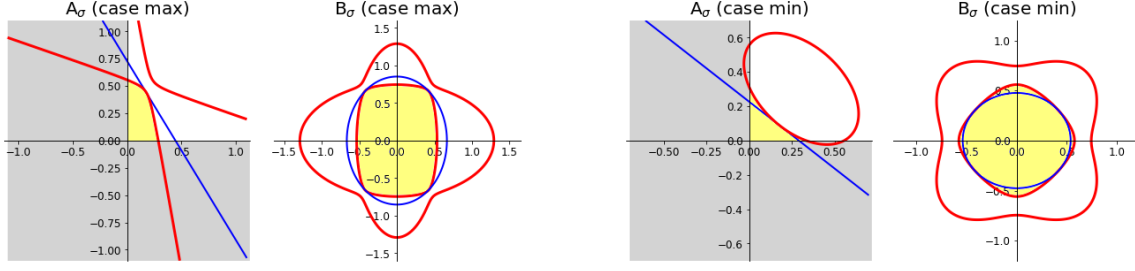


Figure 2: The set  $\mathcal{A}_\sigma \subset \mathbb{R}_+^2$  (resp.  $\mathcal{B}_\sigma \subset \mathbb{R}^2$  as defined from  $(p_r, p_z)$ ), in yellow, is bounded by a conic curve (resp. a quartic curve), in red. Tangent lines to  $\partial\mathcal{A}_\sigma$  correspond to tangent ellipses to  $\partial\mathcal{B}_\sigma$ , in blue. If the conic curve defines a convex (resp. concave) boundary, in case (max) see left (resp. case (min) see right) then the ellipses are exterior (resp. interior) tangent.

the scheme would need to be adapted and the solution  $u$  may have lower regularity. See also Remark 2.9 on the condition  $c_{13} + c_{44} > 0$ .

In order to design our numerical scheme, we need a description of  $\mathcal{F}_\sigma^*$  more tractable than (5, right). For that purpose, we consider the following set, illustrated on Fig. 2

$$\mathcal{A}_\sigma := \text{CC}_0\{(h_r, h_z) \in \mathbb{R}_+^2; \mathcal{Q}_\sigma(h_r, h_z) \leq 1\}, \quad (9)$$

which corresponds to the change of variables  $h_r = p_r^2 = p_x^2 + p_y^2$  and  $h_z = p_z^2$  in (6). The following result shows that  $\mathcal{A}_\sigma$  is either a union or an intersection of triangular regions.

**Theorem 1.3.** *Let  $\sigma \in \mathbb{R}^5$  be admissible. Then there exists  $0 < \alpha_* \leq \alpha^* < 1$  and  $\mu \in C^\infty([\alpha_*, \alpha^*], ]0, \infty[)$  such that one of the following “max” and “min” cases holds:*

(max)  $\mu$  is convex, and  $\mathcal{A}_\sigma = \{(h_r, h_z) \in \mathbb{R}_+^2; \forall \alpha \in [\alpha_*, \alpha^*], (1 - \alpha)h_r + \alpha h_z \leq \mu(\alpha)\}$ .

(min)  $\mu$  is concave, and  $\mathcal{A}_\sigma = \{(h_r, h_z) \in \mathbb{R}_+^2; \exists \alpha \in [\alpha_*, \alpha^*], (1 - \alpha)h_r + \alpha h_z \leq \mu(\alpha)\}$ .

We deduce that the TTI unit ball (4) can be obtained as a union or an intersection of ellipsoids, depending on the alternative of Theorem 1.3 and as illustrated on Figs. 1 and 2 :

$$(max) : \mathcal{B}_\sigma = \bigcap_{\alpha \in [\alpha_*, \alpha^*]} E(\alpha), \quad (min) : \mathcal{B}_\sigma = \bigcup_{\alpha \in [\alpha_*, \alpha^*]} E(\alpha), \quad E(\alpha) := \{(1 - \alpha)(p_x^2 + p_y^2) + \alpha p_z^2 \leq 1\}.$$

The denomination (max) and (min) reflects the expression of the numerical scheme (2) that we eventually obtain. The treatment of the (max) and (min) cases is remarkably symmetric in the results presented below, despite some technical differences in the proof of Corollary 1.4 and more notably in the convergence analysis of Theorem 4.1. In order to favor a unified treatment, we thus introduce a notation “mix” which stands for either “max” or “min”, according to the alternative in Theorem 1.3. However, because a number of mathematical operations (such as negation or Legendre-Fenchel duality) turn a maximum into a minimum, and conversely, we also need to introduce a complementary  $\overline{\text{mix}}$  notation. Summarizing, we denote

$$\text{Case (max): } \text{mix} := \max \text{ and } \overline{\text{mix}} := \min. \quad \text{Case (min): } \text{mix} := \min \text{ and } \overline{\text{mix}} := \max. \quad (10)$$

The proof of Theorem 1.3, presented in Section 2, relies on the observation that the boundary of  $\mathcal{A}_\sigma$  is defined by a portion of the conic curve  $\mathcal{C}_\sigma := \{(h_z, h_r) \in \mathbb{R}^2; \mathcal{Q}_\sigma(h_z, h_r) = 1\}$ . The

<sup>1</sup>Indeed, these conditions imply that  $R(z)^{-1}\mathcal{B}_{\sigma(z)}$  is a convex and compact neighborhood of the origin, depending continuously on  $z \in \overline{\Omega}$ , so that the theory of viscosity solutions for optimal control problems is applicable.



function  $\mu$  and the bounds  $\alpha_* \leq \alpha^*$  admit a simple closed form expression, established in (43), which is welcome for implementation purposes, but is not particularly enlightening for the mathematical analysis. Denoting  $\alpha := (1 - \alpha, \alpha)$  and rewriting the quadratic function (3) as  $\mathcal{Q}_\sigma(p) = \langle l, p \rangle + \frac{1}{2} \langle p, Qp \rangle$  one has

$$\mu(\alpha) = \varepsilon \sqrt{\langle \alpha, Q^{-1} \alpha \rangle (2 + \langle l, Q^{-1} l \rangle)} - \langle \alpha, Q^{-1} l \rangle, \quad \text{where } \varepsilon := \text{sign}(2 + \langle l, Q^{-1} l \rangle),$$

$$\{\alpha_*, \alpha^*\} = \left\{ \frac{l_2 + Q_{12}x}{l_1 + l_2 + Q_{11}x + Q_{12}x}, \frac{l_2 + Q_{22}y}{l_1 + l_2 + Q_{12}y + Q_{22}y} \right\}, \quad \text{where } x \text{ solves } l_1x + \frac{1}{2}Q_{11}x^2 = 1,$$

and  $l_2y + \frac{1}{2}Q_{22}y^2 = 1$ ; both  $x$  and  $y$  are the smallest root of their defining quadratic equation. An alternative expression of  $\mu$  applies when  $Q$  is degenerate (44), see Section 2.3.

As a consequence of Theorem 1.3, we obtain a new expression of the TTI eikonal equation (5) operator  $\mathcal{F}_\sigma^*(R \cdot)$  as an extremum of Riemannian norms (11). We also derive its gradient and dual norm (12), which are involved in the shortest path interpretation of the PDE, see Remark 1.5.

**Corollary 1.4.** *Let  $\sigma \in \mathbb{R}^5$  be admissible, and let  $R \in \text{GL}_3(\mathbb{R})$ . Denote by  $\text{mix} \in \{\max, \min\}$  the corresponding case of Theorem 1.3. Then for any  $p \in \mathbb{R}^3$*

$$\mathcal{F}_\sigma^*(Rp) = \min_{\alpha \in [\alpha_*, \alpha^*]} \mu(\alpha)^{-\frac{1}{2}} \|p\|_{D(\alpha)}, \quad \text{where } D(\alpha) := R^\top \begin{pmatrix} 1 - \alpha & & \\ & 1 - \alpha & \\ & & \alpha \end{pmatrix} R. \quad (11)$$

Introducing the norm  $\mathcal{F}^*(p) := \mathcal{F}_\sigma^*(Rp)$ , we have the following expressions of its gradient at  $p \neq 0$ , and of the dual norm defined as  $\mathcal{F}(v) := \max\{\langle p, v \rangle; \mathcal{F}^*(p) \leq 1\}$

$$\nabla \mathcal{F}^*(p) = \mu(\alpha')^{-\frac{1}{2}} \frac{D(\alpha')p}{\|p\|_{D(\alpha')}}, \quad \mathcal{F}(v) = \overline{\min}_{\alpha \in [\alpha_*, \alpha^*]} \mu(\alpha)^{\frac{1}{2}} \|v\|_{D(\alpha)^{-1}}, \quad (12)$$

where  $\alpha'$  in (12, left) is the optimal parameter in (11, left), and where  $\{\text{mix}, \overline{\text{mix}}\} = \{\min, \max\}$ .

**Remark 1.5** (Shortest path interpretation of the TTI eikonal PDE). *We assume in this remark that the TTI parameters  $\sigma$  and  $R$  vary continuously on the PDE domain. Denote  $\mathcal{F}_q^*(p) := \mathcal{F}_{\sigma(q)}^*(R(q)p)$  for all  $q \in \overline{\Omega}$  and  $p \in \mathbb{R}^d$ , and likewise the dual norm  $\mathcal{F}$ . Then the unique viscosity solution to the eikonal equation (5) is the geodesic distance map from the source point  $q_0$  [BCD08]:*

$$u(q) = \min \{ \text{length}_{\mathcal{F}}(\gamma); \gamma(0) = q_0, \gamma(1) = q \}, \quad \text{length}_{\mathcal{F}}(\gamma) := \int_0^1 \mathcal{F}_{\gamma(t)}(\gamma'(t)) dt,$$

where the infimum is over Lipschitz paths  $\gamma : [0, 1] \rightarrow \overline{\Omega}$  with the given endpoints. Conversely, the optimal path can be obtained from the value function by solving the backtracking ODE

$$\gamma'(t) = V(\gamma(t)) \quad \text{where } V(q) := \nabla \mathcal{F}_q^*(\nabla u(q)), \quad (13)$$

backwards in time, with terminal boundary condition  $\gamma(T) = q$  where  $T = u(q)$ . Numerically the geodesic flow  $V$  is estimated in an upwind manner, using (12, left) and adapting the Riemannian case presented in [MP19, Section 3.2.1].

## 1.2 Monotony, causality, and the fast marching algorithm

We recall in this subsection the concept of finite difference scheme  $\mathfrak{F}$ , and two key structural properties known as discrete degenerate ellipticity<sup>2</sup> (DDE) and causality that enable a stable and

<sup>2</sup>This terminology is closely related to monotony, but more precise in our context.

fast numerical solution, see Definition 1.6. We also introduce Selling's matrix decomposition, see Proposition 1.7, a tool from the field of discrete geometry previously used for the discretization of the Riemannian eikonal PDE [Mir19].

**Definition 1.6.** *Let  $X, \overline{X}$  be finite sets with  $X \subset \overline{X}$ . Consider a finite difference scheme  $\mathfrak{F}$  on  $X$  taking the form*

$$\mathfrak{F}u(q) := \hat{\mathfrak{F}}(q, [u(q) - u(r)]_{r \in \overline{X}}), \quad (14)$$

where  $q \in X$ ,  $u \in \mathbb{R}^{\overline{X}}$ , and  $\hat{\mathfrak{F}} : X \times \mathbb{R}^{\overline{X}} \rightarrow \mathbb{R}$  is continuous. The scheme  $\mathfrak{F}$  is said:

- Discrete Degenerate Elliptic (DDE), if  $\hat{\mathfrak{F}}$  is non-decreasing w.r.t. its second argument.
- Causal, if  $\hat{\mathfrak{F}}$  only depends on the positive part of its second argument.

The DDE property implies a comparison principle for the equation  $\mathfrak{F}u = 1$ , and thus plays a key role in ensuring the stability and convergence of the scheme solutions, following techniques introduced in [Obe06]. Causality, on the other hand, enables the Fast Marching Method (FMM) to solve the system  $\mathfrak{F}u = 1$ ; this property was initially used in [Set96], and formalized as above in [Mir19]. These consequences are summarized in [Mir19, Theorem 2.3], and also discussed here in Section 4.3. Our GPU massively parallel solver, on the other hand, is based on a variant of the fast iterative method [JW08], which requires the DDE property but not causality, although it benefits from it.

In order to fully determine the solution, the scheme  $\mathfrak{F}$  is complemented with Dirichlet boundary conditions, of the form  $u = \psi$  on the discrete boundary  $\partial X := \overline{X} \setminus X$ , where  $\psi : \partial X \rightarrow \mathbb{R}$  is given data. The description of the FMM solver Algorithm 1 involves two additional objects derived from the scheme  $\mathfrak{F}$ : an update operator  $\Lambda$  defined implicitly, used to reformulate the system of equations  $\mathfrak{F}u = 1$  into the fixed point problem  $\Lambda u = u$ , and a stencil  $\mathcal{V}$  which describes the scheme local dependency structure, used to guide the update order. For each  $q \in X$  and  $u \in \mathbb{R}^{\overline{X}}$ , the update operator value  $\Lambda u(x) = \lambda \in \mathbb{R}$  is defined by the equation

$$\hat{\mathfrak{F}}(q, [\lambda - u(r)]_{r \in \overline{X}}) = 1. \quad (15)$$

Note that (15, l.h.s.) is a non-decreasing function of  $\lambda$  under the DDE property. In the applications of interest, one easily checks that (15) admits a unique solution, as discussed below in our case, so that the update operator  $\Lambda$  is well defined. On the other hand, for each  $q \in X$ , the stencil  $\mathcal{V}(q) \subset \overline{X}$  is defined as the collection of neighbors  $r$  such that the expression of  $\mathfrak{F}u(q)$  depends on  $u(r)$ .

In the following, we fix a grid scale  $h > 0$  for the discretization of the PDE domain  $\Omega$ , and we assume w.l.o.g. that the source point is  $q_0 = 0$ . Consistently with the addressed problem (5), we can assume that

$$\Omega_h := \Omega \cap h\mathbb{Z}^d, \quad X := \Omega_h \setminus \{q_0\}, \quad \partial X := \{q_0\},$$

with boundary data  $u(q_0) = \psi(q_0) = 0$ . Note that alternative boundary conditions may be considered, such as the null Dirichlet boundary conditions on  $\partial\Omega$  used in the convergence analysis Section 4 for simplicity. We reproduce in the rest of this section a DDE and causal scheme for Riemannian eikonal equations originally presented in [Mir19], which acts as a building block for the construction of our scheme in Section 1.3 in combination with the PDE operator description (11). For that purpose, we introduce a tool from lattice geometry, known as Selling's decomposition of positive quadratic forms [Sel74, CS92], which is particularly convenient for the

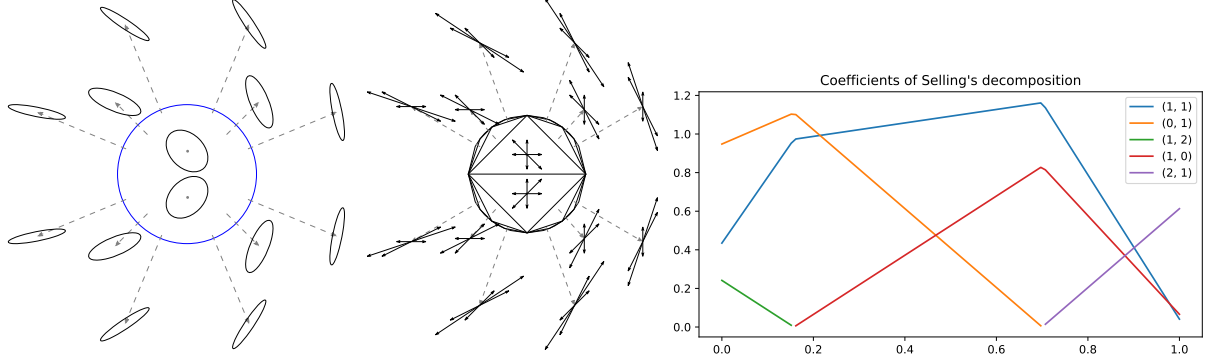


Figure 3: (Left) To each point  $(x, y)$  of the open unit disk (blue boundary), we attach the Pauli matrix  $D = \begin{pmatrix} 1+x & y \\ y & 1-x \end{pmatrix} \in S_2^{++}$ , and show the ellipse  $\{p \in \mathbb{R}^2; \langle p, D(x, y)p \rangle = 1\}$ . (Center) Offsets  $\pm e_i$ ,  $1 \leq i \leq 3$ , of Selling's decomposition (16) of the matrix  $D(x, y)$ . (Right) Coefficients and offsets of Selling's decomposition of  $(1 - \alpha)D_0 + \alpha D_1$ , as  $\alpha \in [0, 1]$ , where  $D_0, D_1 \in S_2^{++}$  are randomly chosen. The coefficient associated to a given offset is piecewise affine (17).

design of DDE discretizations of non-linear and anisotropic PDEs on Cartesian grids, both of first [Mir18, Mir19] and second order [FM14, BBM20]. We gather in the next result two properties of Selling's decomposition that are useful for our scheme, consistency (16) and piecewise linearity (17), and we refer to Appendix B for the proof and for a more constructive and detailed presentation. A function is said affine if it is the sum of a constant and of a linear map.

**Proposition 1.7** (Selling's decomposition). *Let  $D \in S_d^{++}$ , where  $d \in \{2, 3\}$ . Then Selling's decomposition defines weights  $\rho_i \geq 0$ , and offsets  $e_i \in \mathbb{Z}^d \setminus \{0\}$ , where  $1 \leq i \leq I = d(d+1)/2$ , such that*

$$D = \sum_{1 \leq i \leq I} \rho_i e_i e_i^\top. \quad (16)$$

*In addition, assume that  $D(\alpha) \in S_d^{++}$  depends in an affine manner of a parameter  $\alpha \in [\alpha_*, \alpha^*]$ , and let  $(\rho_i(\alpha), e_i)_{i=1}^I$  be the weights and offsets of Selling's decomposition<sup>3</sup>. Then there exists  $\alpha_* = \alpha_0 < \dots < \alpha_K = \alpha^*$ , such that for all  $0 \leq k < K$  and all  $1 \leq i \leq I$ ,*

$$\rho_i(\alpha) \text{ is affine as } \alpha \in [\alpha_k, \alpha_{k+1}]. \quad (17)$$

If  $D$  is a diagonal matrix, then Selling's decomposition (16) is particularly simple:  $\rho_1, \dots, \rho_d$  are the diagonal coefficients,  $\rho_{d+1} = \dots = \rho_I = 0$ , and  $e_1, \dots, e_d$  is the canonical basis of  $\mathbb{R}^d$ . In contrast, if  $D$  is not diagonal, then Selling's decomposition differs for the eigenvalue-eigenvector decomposition, and crucially it only involves offsets  $(e_i)_{i=1}^I$  with integer coordinates. The piecewise linearity of Selling's decomposition is used in Theorem 1.11 below to establish that the numerical scheme proposed in this paper benefits from a property, known as quasi-convexity or quasi-concavity, which allows to evaluate it efficiently numerically.

We devote the rest of this subsection to the description and analysis of a discretization scheme denoted  $\mathfrak{F}^D$  for the Riemannian eikonal equation [Mir19], which generalizes the classical isotropic fast marching scheme [Set96] using Selling's decomposition. Given  $D \in S_d^{++}$ , and  $u : \Omega_h \rightarrow \mathbb{R}$ ,

<sup>3</sup>Here  $I$  is arbitrary. Yet for each  $\alpha \in [\alpha_*, \alpha^*]$ , at most  $d(d+1)/2$  of the weights  $\rho_i(\alpha)$ ,  $1 \leq i \leq I$ , are positive.

$q \in \Omega_h \setminus \{q_0\}$ , we define

$$\mathfrak{F}^D u(q) := \sum_{1 \leq i \leq I} \rho_i \max \left\{ 0, \frac{u(q) - u(q + h e_i)}{h}, \frac{u(q) - u(q - h e_i)}{h} \right\}^2 \quad (18)$$

where  $D = \sum_{i=1}^I \rho_i e_i e_i^\top$  is Selling's decomposition. Since the offsets  $e_i$  have integer coordinates, the scheme  $\mathfrak{F}^D$  only involves values of the unknown  $u$  on the Cartesian discretization grid:  $q + \varepsilon h e_i \in h\mathbb{Z}^d$  for all  $1 \leq i \leq I$  and  $\varepsilon \in \{-1, 1\}$ . By convention, the terms associated to points  $q + \varepsilon h e_i \notin \Omega_h$  are discarded<sup>4</sup>, which implements outflow boundary conditions. The scheme stencil

$$\mathcal{V}^D(q) := \{q + \varepsilon h e_i; 1 \leq i \leq I, \rho_i > 0, \varepsilon = \pm 1\} \cap \Omega_h,$$

is reasonably small since  $\|e_i\| \leq C \sqrt{\|D\| \|D^{-1}\|}$ , by Proposition B.4 (Offset boundedness).

The DDE property of the scheme  $\mathfrak{F}^D$  follows from the non-negativity of the weights  $(\rho_i)_{i=1}^I$ , and the observation that  $s \in \mathbb{R} \mapsto \max\{0, s\}^2$  is non-decreasing. Causality holds as well, since by construction  $\mathfrak{F}^D u(q)$  only depends on the positive part of the finite differences  $u(q) - u(r)$ , where  $r \in \mathcal{V}^D(q)$ . Observing that (18) defines a strictly increasing function of  $u(q)$  over the interval  $[u_{\min}, +\infty[$ , where  $u_{\min} := \min\{u(r); r \in \mathcal{V}(q)\}$ , we obtain that the corresponding update operator  $\Lambda^D$  is uniquely defined by (15). In practice, solving (15) amounts to computing the roots of  $I$  univariate quadratic equations, see [MGB<sup>+</sup>21].

Finally, recalling that  $D := \sum_{i=1}^I \rho_i e_i e_i^\top$ , we obtain for smooth  $u$  the consistency relation

$$\|\nabla u(q)\|_D^2 = \sum_{1 \leq i \leq I} \rho_i \langle \nabla u, e_i \rangle^2 = \sum_{1 \leq i \leq I} \rho_i \max\{0, \langle \nabla u, e_i \rangle, -\langle \nabla u, e_i \rangle\}^2 = \mathfrak{F}^D u(q) + \mathcal{O}(h^r), \quad (19)$$

with  $r = 1$  for the straightforward implementation (18). Some scheme modifications improve its accuracy, such as source factorization [LQ12], multiscale computation [WFNBZ20], and second order finite differences [Set99] which yield  $r = 2$ . However they break the DDE and causality properties, hence must be used carefully, see Appendix C.

---

**Algorithm 1** The Fast Marching algorithm, solving  $\mathfrak{F}u = 1$  for a DDE and causal scheme  $\mathfrak{F}$ .

---

**Input:** The update operator  $\Lambda$  and stencils  $\mathcal{V}$  associated to  $\mathfrak{F}$ . Boundary conditions  $\psi$ .

**Initialize:**  $u = +\infty$  on the domain  $X$ , and  $u = \psi$  on  $\partial X$ . Tag all points as non-accepted.

**While** a non-accepted point remains:

Denote by  $q \in \overline{X}$  the non-accepted point minimizing  $u(q)$ . 1.

Tag  $q$  as accepted. (And optionally, for e.g. higher order methods: PostProcess( $q$ ) ). 2.

**For each** non-accepted point  $r \in X$  such that  $q \in \mathcal{V}(r)$ : 3.

$u(r) \leftarrow \tilde{\Lambda}u(r)$  (modified operator using only the values from accepted points). 4.

$u(r) \leftarrow \tilde{\Lambda}u(r)$  (modified operator using only the values from accepted points). 5.

---

### 1.3 Discretization scheme for the TTI model

We combine the description of the TTI norm as an extremum of Riemannian norms (11), with the Riemannian scheme (18), to obtain a DDE and causal discretization of the TTI eikonal PDE. We then discuss its efficient numerical implementation, and the convergence of its solutions as

---

<sup>4</sup>Formally, we use the boundary condition  $u = +\infty$  on  $h\mathbb{Z}^d \setminus \Omega_h$ .

the grid scale  $h > 0$  is refined. Specifically, and consistently with (2), define for any  $u : \Omega_h \rightarrow \mathbb{R}$  and any  $q \in \Omega_h$ , where  $\Omega_h := \Omega \cap h\mathbb{Z}^d$  is a Cartesian discretization grid

$$\mathfrak{F}u(q) := \underset{\alpha \in [\alpha_*, \alpha^*]}{\text{mix}} \mathfrak{F}^\alpha u(q), \quad \text{where } \mathfrak{F}^\alpha := \frac{1}{\mu(\alpha)} \mathfrak{F}^{D(\alpha)}. \quad (20)$$

We used the notations  $0 < \alpha_* \leq \alpha^* < 1$  and  $\mu(\alpha) > 0$  from Theorem 1.3, the matrix  $D(\alpha)$  and extremum operator  $\text{mix} \in \{\min, \max\}$  from (11), and the Riemannian scheme  $\mathfrak{F}^D$  defined in (18). For sufficiently smooth  $u$  one has

$$\mathcal{F}_\sigma(R\nabla u(q))^2 = \underset{\alpha \in [\alpha_*, \alpha^*]}{\text{mix}} \frac{1}{\mu(\alpha)} \|\nabla u(q)\|_{D(\alpha)}^2 = \mathfrak{F}u(q) + \mathcal{O}(h^r), \quad (21)$$

using the consistency relation in the Riemannian case (19), with the same order  $r \in \{1, 2\}$ , and the expression of the TTI norm (11).

The proposed scheme  $\mathfrak{F}$  for the TTI eikonal PDE is defined as the maximum or minimum of the infinite family of schemes  $\mathfrak{F}^{D(\alpha)}$ ,  $\alpha \in [\alpha_*, \alpha^*]$ , hence inherits their DDE and causality properties as shown in Proposition 1.8 below. The existence and uniqueness of the scheme solutions, and their convergence as  $h \rightarrow 0$  to the viscosity solution of (5), can then be established following a common scheme of proof, see Theorem 4.1 below. In a similar fashion, a DDE scheme for the second order fully non-linear Monge-Ampere [BM21] and Pucci [BBM20] PDEs, in two dimensions, is obtained as the maximum of an infinite family of linear schemes. In the same spirit again, and in the context of fast marching methods, multi-stencil schemes [HF07] are defined as the maximum of a finite number of discretizations of the eikonal equation (with identical anisotropy, unlike here). Note however that considering the *minimum* of several schemes, as we do in (20) for case (min) of Theorem 1.3, is uncommon and leads to a few additional difficulties in the analysis in comparison with the (max) case, see Section 4.

**Proposition 1.8.** *Let  $A$  be a compact set and let  $\mathfrak{F}^\alpha$ , for each  $\alpha \in A$ , be a finite difference scheme on a finite set  $X$ , depending continuously on the parameter  $\alpha$ . Define*

$$\mathfrak{F}u(q) := \max_{\alpha \in A} \mathfrak{F}^\alpha u(q) \quad \left( \text{resp. } \mathfrak{F}u(q) := \min_{\alpha \in A} \mathfrak{F}^\alpha u(q) \right)$$

*for all  $u \in \mathbb{R}^{\overline{X}}$  and all  $q \in X$ . If  $\mathfrak{F}^\alpha$  is DDE (resp. causal) for all  $\alpha \in A$ , then so is  $\mathfrak{F}$ . Furthermore, denoting by  $\Lambda^\alpha$  the update operator for  $\mathfrak{F}^\alpha$  (which is assumed to exist and to depend continuously on  $\alpha \in A$ ), one has*

$$\Lambda u(q) = \min_{\alpha \in A} \Lambda^\alpha u(q) \quad \left( \text{resp. } \Lambda u(q) = \max_{\alpha \in A} \Lambda^\alpha u(q) \right). \quad (22)$$

*Proof.* A maximum or a minimum over a compact set of a continuously depending family of functions which are continuous (resp. non-decreasing) (resp. depend only on the positive part of their arguments), clearly obeys the same property. The first claim follows.

We focus on (22, left) since the other case is proved similarly, and denote  $\lambda_* := \min_{\alpha \in A} \Lambda^\alpha u(q)$ . If  $\lambda < \lambda_*$ , then  $\hat{\mathfrak{F}}^\alpha(q, [\lambda - u(r)]_{r \in X}) < 0$  for all  $\alpha \in A$  by the DDE property, hence  $\max_{\alpha \in A} \hat{\mathfrak{F}}^\alpha(q, [\lambda - u(r)]_{r \in X}) < 0$  by compactness. If  $\lambda > \lambda_*$  on the other hand, then  $\max_{\alpha \in A} \hat{\mathfrak{F}}^\alpha(q, [\lambda - u(r)]_{r \in X}) > 0$  by definition. Thus, by continuity,  $\lambda_*$  is the unique solution to  $\hat{\mathfrak{F}}(q, [\lambda - u(r)]_{r \in X})$ , hence  $\lambda_* = \Lambda u(q)$  as announced.  $\square$

Proposition 1.8 immediately implies that the TTI scheme (20) is DDE and causal, and thus solvable using the FMM Algorithm 1. The numerical implementation however requires to efficiently evaluate the update operator  $\Lambda$  associated with the scheme, which is defined as an

extremum (22) over a continuous set of parameters  $A = [\alpha_*, \alpha^*]$ . We compare in the following two strategies for solving this optimization problem, used respectively in our GPU and CPU eikonal solver.

**Optimization by grid search.** In this approach, the maximization or minimization problem (20) over  $[\alpha_*, \alpha^*]$ , is approximated using an exhaustive search over a regular sampling of this real interval with  $K + 1$  elements, where the integer  $K \geq 1$  is fixed by the user. More explicitly, we introduce the scheme  $\mathfrak{F}_K$  and update operator  $\Lambda_K$  defined as

$$\mathfrak{F}_K u(q) := \underset{0 \leq k \leq K}{\text{mix}} \mathfrak{F}^{\alpha_k} u(q), \quad \Lambda_K u(q) := \overline{\underset{0 \leq k \leq K}{\text{mix}}} \mathfrak{F}^{\alpha_k} u(q), \quad \text{with } \alpha_k := (1 - \frac{k}{K})\alpha_* + \frac{k}{K}\alpha^*, \quad (23)$$

following the notations of (20). In particular  $\text{mix} \in \{\max, \min\}$  is the suitable extremum, and  $\overline{\text{mix}} \in \{\min, \max\}$  is the opposite extremum, following (10) and Proposition 1.8. The use of an equispaced sampling of parameters  $\alpha_0 \leq \dots \leq \alpha_K$  in the interval  $[\alpha_*, \alpha^*]$  is quasi-optimal for consistency, see Proposition 1.9, and corresponds to an envelope of the TTI slowness surface by a family of ellipses with regularly varying aspect ratios, which is visually pleasing, see Figs. 1 and 5.

The eikonal solvers [MP19, MGB<sup>+</sup>21], originally limited to  $K = 2$  and to the (max) case, are easily adapted to address the scheme  $\mathfrak{F}_K$ , using an exhaustive search over  $0 \leq k \leq K$  to evaluate the update operator  $\Lambda_K$ . This approach is well suited to massively parallel accelerators such as GPUs, since those have (i) enough horsepower to accommodate the computational cost of exhaustive search, and (ii) a SIMT<sup>5</sup> architecture that is not well suited to the multiple conditional branchings found in more sophisticated optimization procedures.

**Proposition 1.9.** *For smooth  $u$ , one has  $\mathfrak{F}_K u(q) = \mathfrak{F}u(q) + \mathcal{O}(h^r + K^{-2})$ .*

*Proof.* By the consistency relation in the Riemannian case (19),

$$\underset{0 \leq k \leq K}{\text{mix}} g(\alpha_k) = \mathfrak{F}_K u(q) + \mathcal{O}(h^r), \quad \text{where } g(\alpha) := \frac{1}{\mu(\alpha)} \|\nabla u(q)\|_{D(\alpha)}^2. \quad (24)$$

The function  $g : [\alpha_*, \alpha^*] \rightarrow \mathbb{R}$  is smooth by (11) and Theorem 1.3. Since  $\alpha_0 = \alpha_*$ ,  $\alpha_K = \alpha^*$ , and  $\alpha_{k+1} - \alpha_k = (\alpha^* - \alpha_*)/K = \mathcal{O}(1/K)$  for all  $0 \leq k < K$ , one has  $\text{mix}\{g(\alpha_k); 0 \leq k \leq K\} = \text{mix}\{g(\alpha); \alpha \in [\alpha_*, \alpha^*]\} + \mathcal{O}(1/K^2)$ . From this point, the announced result follows from (21).  $\square$

For numerical efficiency, one usually balances the errors  $\mathcal{O}(h^r + K^{-2})$  associated to the discretization scale  $h$  and to the consistency of the operator approximation with  $K$  terms. This suggests the parameter choice  $K \approx h^{-\frac{1}{2}}$  with a first order scheme ( $r = 1$ ), and  $K \approx h^{-1}$  with a second order scheme ( $r = 2$ ). For instance, in the synthetic numerical experiment presented on Fig. 10, the TTI scheme needs to be defined as the extremum of  $K + 1 = 26$  Riemannian schemes to ensure good accuracy. In the second order case, the evaluation cost of the update operator (23) thus becomes non-negligible, and for this reason an optimization procedure more efficient than exhaustive search is described in the next paragraph.

**Quasi-convex optimization.** This approach relies on a fine property of Selling's matrix decomposition, namely the piecewise linearity of its coefficients (17) established in Proposition 1.7, which is used in this paper for the first time in the discretization of a three dimensional PDE. The same property is exploited in [BM21, BBM20] to obtain a DDE, second order consistent, and

---

<sup>5</sup>Single instruction multiple threads

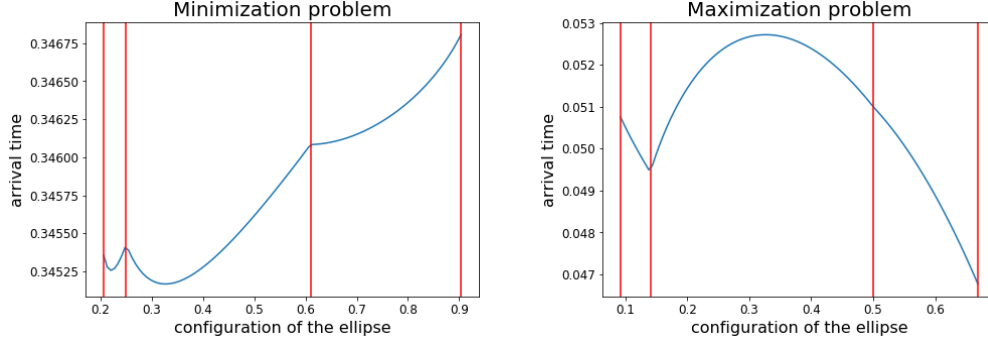


Figure 4: Mapping  $\alpha \in [\alpha_*, \alpha^*] \mapsto f(\alpha) := \Lambda^\alpha u(q)$  obtained for some TTI parameters  $\sigma$ ,  $R$ , a point  $q \in h\mathbb{Z}^d$ , and an arbitrary mapping  $u : \Omega_h \rightarrow \mathbb{R}$ . The vertical red lines correspond to the abscissas  $\alpha_0 \leq \dots \leq \alpha_K$  of Proposition 1.7, here with  $K = 3$ . Left (resp. Right) subfigure illustrates case (max) (resp. case (min)), where by Theorem 1.11 the function  $f$  is quasi-convex (resp. quasi-concave) on each sub-interval  $[\alpha_k, \alpha_{k+1}]$ ,  $0 \leq k \leq K$ , and must be minimized (resp. maximized).

numerically efficient scheme for the Pucci and Monge-Ampere PDEs in two dimensions. In those previous works, the non-linear PDE operator can be expressed as the maximum of an infinite family of linear operators, each discretized using Selling's decomposition, in a spirit similar to (20); a closed form expression is then obtained using the piecewise linear property of Selling's decomposition. We do not obtain such a closed form here, but nevertheless we derive a property known as quasi-convexity, allowing for an efficient implementation.

**Definition 1.10.** A function  $f : A \rightarrow \mathbb{R}$ , where  $A$  is a convex subset of a vector space, is said *quasi-convex* if for each  $\lambda \in \mathbb{R}$  the set  $\{x \in I; f(x) \leq \lambda\}$  is convex.

By construction, the set of minimizers of a quasi-convex function is convex, and in particular there is at most one isolated local minimum. If  $A = [a, b]$  is a segment of  $\mathbb{R}$ , as in our application, and if  $f$  is continuous and quasi-convex, then the classical golden search method [PTVF07, Section 10.2] produces an interval of length  $(b - a)\phi^N$  containing its minimizer using  $N + 1$  evaluations of  $f$ , where  $\phi^{-1} = \frac{1+\sqrt{5}}{2}$  is the golden ratio. This is considerably more efficient than optimization by grid search, considered previously, which only yields an interval of length  $2(b - a)/N$  for the same numerical cost. A function  $f$  is said *quasi-concave* if  $-f$  is convex, and in that case by the previous discussion it can be efficiently *maximized* numerically.

**Theorem 1.11.** Let  $\alpha_* \leq \alpha_0 \leq \dots \leq \alpha_K = \alpha^*$  be such that Selling's decomposition of the matrix  $D(\alpha)$ , see (11, right), is piecewise linear on each interval  $[\alpha_k, \alpha_{k+1}]$ ,  $0 \leq k < K$ , in the sense of (17). Fix  $u : \Omega_h \rightarrow \mathbb{R}$ ,  $q \in \Omega_h$  and define  $f(\alpha) := \Lambda^\alpha u(q)$  for all  $\alpha \in [\alpha_*, \alpha^*]$ . Then the following alternative holds, whose cases match those of Theorem 1.3

(max)  $f$  is quasi-convex on each interval  $[\alpha_k, \alpha_{k+1}]$ ,  $0 \leq k < K$ .

(min)  $f$  is quasi-concave on each interval  $[\alpha_k, \alpha_{k+1}]$ ,  $0 \leq k < K$ .

This result allows to extremize the function  $f(\alpha) := \Lambda^\alpha u(q)$  over the interval  $[\alpha_*, \alpha^*]$  in a numerically efficient manner, and thus to evaluate the update operator (22). A possible allure of  $f$  is illustrated on Fig. 4.

## 2 Properties and guarantees of TTI models

This section is devoted to the proof of the results announced in Section 1.1. We introduce in Section 2.1 several properties of Hooke tensors, known as ellipticity, positivity and separability, and relate them with the admissibility conditions (8) of the TTI parameters  $\sigma$ . We establish in Section 2.2 that the TTI unit ball  $\mathcal{B}_\sigma$  is convex and compact, thus concluding the proof of Theorem 1.2. We prove Theorem 1.3 in Section 2.3, where we also derive the closed form expression of the weight function  $\mu : [\alpha_*, \alpha^*] \rightarrow ]0, \infty[$ .

**Notations.** The TTI parameters  $\sigma \in \mathbb{R}^5$  are regarded as fixed throughout this section, and thus for readability the sets  $\mathcal{A}_\sigma, \mathcal{B}_\sigma, \mathcal{C}_\sigma, \mathcal{Q}_\sigma$  introduced in Section 1.1 are simply denoted  $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{Q}$ . The symbol  $\propto$  denotes *positive proportionality*, i.e.  $v \propto w$  iff  $v = \lambda w$  for some  $\lambda > 0$ .

### 2.1 Admissible coefficients, and properties of Hooke tensors

We relate the TTI eikonal PDE with a two-dimensional Hooke tensor (28), and investigate its algebraic properties as a preliminary step to Theorem 1.2. See Appendix A and references therein for a more physically oriented discussion of these elasticity parameters. A Hooke tensor is a 4-th order tensor,  $\mathbf{c} = (\mathbf{c}_{ijkl})$  where  $i, j, k, l \in \{1, \dots, d\}$  in dimension  $d$ , which characterizes the anisotropy properties of a linear elastic material, hence also the propagation speed of elastic waves through it (72). Hooke tensors are subject to the major and minor symmetry relations  $\mathbf{c}_{ijkl} = \mathbf{c}_{jikl} = \mathbf{c}_{klij}$ , and for this reason a Hooke tensor  $\mathbf{c}$  can be represented compactly as a symmetric matrix  $\mathfrak{C}$  of shape  $3 \times 3$  if  $d = 2$  (resp.  $6 \times 6$  if  $d = 3$ ) using Voigt's matrix of indices denoted  $\mathbf{v}$ :

$$\mathbf{c}_{ijkl} := \mathfrak{C}_{\mathbf{v}_{ij}\mathbf{v}_{kl}} \quad \text{where } \mathbf{v} = \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix} \quad \left( \text{resp. } \mathbf{v} = \begin{pmatrix} 1 & 6 & 5 \\ 6 & 2 & 4 \\ 5 & 4 & 3 \end{pmatrix} \right). \quad (25)$$

Following [BST83] we recall the notion of a positive or elliptic Hooke tensor in Definition 2.1, and the relation between these properties in Lemma 2.2.

**Definition 2.1.** A Hooke tensor  $\mathbf{c}$  is said *strictly positive* (resp. *strictly elliptic*) if

$$\sum_{i,j,k,l} \mathbf{c}_{ijkl} m_{ij} m_{kl} > 0 \quad \left( \text{resp. } \sum_{i,j,k,l} \mathbf{c}_{ijkl} p_i q_j p_k q_l > 0 \right), \quad (26)$$

for all  $m \in S_d \setminus \{0\}$  (resp.  $p, q \in \mathbb{R}^d \setminus \{0\}$ ), where the sums implicitly range over  $i, j, k, l \in \{1, \dots, d\}$ .

In order to describe further properties of Hooke tensors, we introduce for all  $p \in \mathbb{R}^d$  a symmetric matrix  $\mathbf{c}(p) \in S_d$  defined as follows: for all  $j, l \in \{1, \dots, d\}$

$$\mathbf{c}(p)_{jl} := \sum_{i,k \in \{1, \dots, d\}} \mathbf{c}_{ijkl} p_i p_k, \quad \text{thus } \langle q, \mathbf{c}(p)q \rangle = \sum_{i,j,k,l} \mathbf{c}_{ijkl} p_i q_j p_k q_l. \quad (27)$$

The following lemma rephrases the positivity and ellipticity properties of Hooke tensors in terms of usual matrix positive definiteness.

**Lemma 2.2.** A Hooke tensor  $\mathbf{c}$  is strictly positive iff  $\mathfrak{C} \in S_D^{++}$  with  $D = d(d+1)/2$ , where  $\mathfrak{C}$  is defined by Voigt's notation (25). It is strictly elliptic iff  $\mathbf{c}(p) \in S_d^{++}$  for all  $p \neq 0$ . Strict positivity implies strict ellipticity.



*Proof.* By definition (26, left) a Hooke tensor is strictly positive iff it defines a positive definite quadratic form over the space  $S_d$  of  $d \times d$  symmetric matrices, which has dimension  $D$ . Noting that  $\mathfrak{C}$  is the matrix of this quadratic form, in the basis  $E_{11}$ ,  $E_{22}$ ,  $(E_{12} + E_{21})/2$  if  $d = 2$  where  $E_{ij}$  is the null matrix except for a single coefficient 1 at position  $(i, j)$ , and likewise in the case  $d = 3$ , we establish the first point. On the other hand, the definition (26, right) of ellipticity can be rephrased using the identity (27, right) as  $\langle q, \mathfrak{c}(p)q \rangle > 0$  for all  $p, q \neq 0$ , in other words  $\mathfrak{c}(p) \in S_d^{++}$  for all  $p \neq 0$ , as announced. Finally, given a strictly positive Hooke tensor and  $p, q \in \mathbb{R}^d \setminus \{0\}$ , define  $m \in S_d$  by  $m_{ij} = p_i q_j + q_j p_i$ , equivalently  $m = pq^\top + qp^\top$ , and note that  $\text{Tr}(m^2) = 2(\langle p, q \rangle^2 + \|p\|^2 \|q\|^2) > 0$ . Thus  $m \neq 0$  and therefore  $0 < \sum_{i,j,k,l} \mathfrak{c}_{ijkl} m_{ij} m_{kl} = 4 \sum_{i,j,k,l} \mathfrak{c}_{ijkl} p_i q_j p_k q_l$ , showing that  $\mathfrak{c}$  is strictly elliptic as announced.  $\square$

We introduce in Definition 2.3 a non-degeneracy property of Hooke tensors referred to as separability [DCC<sup>+</sup>21]. This property ensures that the slowness surfaces of the pressure and shear waves are topologically separated from each other, see Figs. 1 and 2 for examples and Fig. 5 for counter-examples.

**Definition 2.3.** *A Hooke tensor  $\mathfrak{c}$  is said separable iff the largest eigenvalue of  $\mathfrak{c}(p)$  has multiplicity one, for all  $p \in \mathbb{R}^d \setminus \{0\}$ .*

In the rest of this section, we limit our attention to the following two dimensional Hooke tensor, whose coefficients are assumed to belong to the admissible set  $C_{\text{adm}}$ , and are related to the TTI coefficients  $\sigma$  by (7):

$$\mathfrak{c} := \begin{pmatrix} c_{11} & c_{13} & 0 \\ c_{13} & c_{33} & 0 \\ 0 & 0 & c_{44} \end{pmatrix}, \quad \begin{cases} c_{11} > c_{44}, & c_{33} > c_{44}, \\ c_{44} > 0, & c_{13} + c_{44} > 0, \\ c_{11}c_{33} > c_{13}^2. \end{cases} \quad (28)$$

We establish in Proposition 2.4 that  $\mathfrak{c}$  is strictly positive, hence strictly elliptic by Lemma 2.2. We also prove that admissible coefficients form a convex set, as announced in Theorem 1.2.

**Proposition 2.4.** *The following sets of coefficients  $(c_{11}, c_{13}, c_{33}, c_{44}) \in \mathbb{R}^4$  are open and convex:*

$$C_{\text{adm}}^1 := \{c_{11} > 0, c_{33} > 0, c_{11}c_{33} > c_{13}^2, c_{44} > 0\}, \quad C_{\text{adm}}^2 := \{c_{11} > c_{44}, c_{33} > c_{44}, c_{13} > -c_{44}\},$$

*thus also their intersection  $C_{\text{adm}} = C_{\text{adm}}^1 \cap C_{\text{adm}}^2$ . In addition the Hooke tensor  $\mathfrak{c}$  defined by (28, left) is strictly positive for any  $(c_{11}, c_{13}, c_{33}, c_{44}) \in C_{\text{adm}}^1 \supset C_{\text{adm}}$ .*

*Proof.* The openness properties follow from the definition of  $C_{\text{adm}}^1$  and  $C_{\text{adm}}^2$  by strict inequalities. Recall that  $[c_{11} > 0, c_{33} > 0 \text{ and } c_{11}c_{33} > c_{13}^2]$  iff  $\begin{pmatrix} c_{11} & c_{13} \\ c_{13} & c_{33} \end{pmatrix} \in S_2^{++}$ . Thus  $C_{\text{adm}}^1$  characterizes the positive definiteness of the block matrix (28, left), as announced. This also shows that  $C_{\text{adm}}^1$  is in linear bijection with  $S_2^{++} \times ]0, \infty[$ , hence is a convex set. The set  $C_{\text{adm}}^2$  is convex since it is defined by linear inequalities.  $\square$

The rest of this subsection is devoted to the proof that  $\mathfrak{c}$  is separable, which is concluded in Corollary 2.8. For that purpose, we introduce the quadratic function  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  defined as

$$F(x, z) := 1 + c_{11}c_{44}x^2 + c_{33}c_{44}z^2 - (2c_{13}c_{44} + c_{13}^2 - c_{11}c_{33})xz - (c_{11} + c_{44})x - (c_{33} + c_{44})z. \quad (29)$$

The next identity (30, left) relates the function  $F$  with the Hooke tensor  $\mathfrak{c}$ , whereas (30, right) links it with the TTI eikonal equation (5). For all  $p = (p_x, p_z) \in \mathbb{R}^2$  and all  $x, z \in \mathbb{R}$

$$\det(\mathfrak{c}(p) - \text{Id}) = F(p_x^2, p_z^2), \quad F(x, z) = 1 - \mathcal{Q}(x, z). \quad (30)$$

Note that the quadratic function  $\mathcal{Q}$  defined in (3) is based on the admissible coefficients (7).

**Remark 2.5** (Validation of the polynomial identities). *Checking polynomial identities such as (30) by hand is simultaneously trivial, tedious, and error prone. For this reason, a companion notebook is provided containing those verifications in the Wolfram Mathematica<sup>®</sup> language.*

Proposition 2.7 below, which is the most technical result of this subsection, establishes that the function  $F$  vanishes exactly twice in each direction of the positive quadrant. In the following, the letter  $r$  is used to denote a radius in the two-dimensional  $(x, z)$  plane. This usage is *distinct* from the three dimensional cylindrical coordinates considered in Section 1.

**Lemma 2.6.** *Define the polynomial  $P(r) := ar^2 + br + c$ , for all  $r \in \mathbb{R}$ . If  $a > 0$ ,  $b < 0$ ,  $c > 0$ , and the discriminant  $\Delta := b^2 - 4ac > 0$  is positive, then  $P$  has two distinct positive roots.*

*Proof.* Noting that  $0 < \Delta < b^2$  we obtain  $\sqrt{\Delta} < |b| = -b$ , and thus  $(-b \pm \sqrt{\Delta})/(2a) > 0$ .  $\square$

**Proposition 2.7.** *For any  $\theta \in [0, \pi/2]$ , the quadratic polynomial  $r \in \mathbb{R} \mapsto F(r \cos \theta, r \sin \theta)$  has two distinct positive roots. Denoting by  $R(\theta)$  the smallest root, one has  $R \in C^\infty([0, \pi/2], ]0, \infty[)$ .*

*Proof.* By symmetry, we may assume that  $0 \leq \theta < \pi/2$ . Using the change of variables  $r' = r/\cos \theta$ , and denoting  $\alpha := \tan \theta \geq 0$ , we can limit our attention to the following polynomial:

$$F_\alpha(r) := F(r, \alpha r) = 1 + r^2(c_{11}c_{44} + \alpha^2 c_{33}c_{44} - 2\alpha c_{13}c_{44} - \alpha c_{13}^2 + \alpha c_{11}c_{33}) - r(c_{11} + c_{44} + \alpha c_{33} + \alpha c_{44}).$$

One has  $F_\alpha(0) = 1 > 0$  and  $F'_\alpha(0) = -(c_{11} + c_{44} + \alpha c_{33} + \alpha c_{44}) < 0$ . The coefficient of  $r^2$  in  $F_\alpha(r)$  reads

$$c_{44}(c_{11} + \alpha^2 c_{33} - 2\alpha c_{13}) + \alpha(c_{11}c_{33} - c_{13}^2)$$

which is positive, by the admissibility conditions  $c_{44} > 0$  and  $c_{11}c_{33} > c_{13}^2$ . Indeed one has  $(c_{11} + \alpha^2 c_{33})/2 \geq \sqrt{c_{11}\alpha^2 c_{33}} \geq \alpha|c_{13}|$ , by the arithmetic geometric mean inequality. The discriminant of  $F_\alpha(r)$  reads (after suitably grouping the terms)

$$\Delta(\alpha) = \alpha^2(c_{33} - c_{44})^2 + (c_{11} - c_{44})^2 + 2\alpha c_*, \quad c_* := 2c_{13}^2 - c_{11}c_{33} + c_{44}(c_{11} + 4c_{13} + c_{33} + c_{44}).$$

Distinguishing two cases, depending on the sign of  $c_*$ , we establish below that  $\Delta(\alpha) > 0$ .

- Case  $c_* \geq 0$ . Then  $\Delta(\alpha) \geq (c_{11} - c_{44})^2 > 0$  for any  $\alpha \geq 0$ , as announced.
- Case  $c_* < 0$ . Then we consider the discriminant of the polynomial  $\alpha \mapsto \Delta(\alpha)$ , which reads

$$16 \times (c_{13} + c_{44})^2 (c_* - (c_{13} + c_{44})^2) \tag{31}$$

(after factorization), and is thus negative since  $c_{13} + c_{44} > 0$  by admissibility. Therefore  $\Delta(\alpha) \neq 0$  for all  $\alpha \in \mathbb{R}$ , and thus  $\Delta(\alpha)$  has the same sign as  $\Delta(0) = (c_{11} - c_{44})^2 > 0$ .

By Lemma 2.6, the polynomial  $F_\alpha$  admits two positive roots, as announced. Finally we note that the smallest root  $(-b - \sqrt{\Delta})/(2a)$  of a polynomial of degree two  $P(r) = a + br + cr^2$  is a smooth function of its coefficients so long as the discriminant  $\Delta = b^2 - 4ac$  and the dominant coefficient  $a$  remain positive. By composition the smallest root of  $F_\alpha$  depends smoothly on  $\alpha \in [0, \infty[$ , which concludes.  $\square$

**Corollary 2.8.** *The Hooke tensor  $\mathfrak{c}$  is separable.*

*Proof.* Fix an arbitrary  $p = (p_x, p_z) \in \mathbb{R}^2 \setminus \{0\}$ . Then for any  $r > 0$ , one has

$$F(rp_x^2, rp_z^2) = \det(\mathfrak{c}(\sqrt{r}p) - \text{Id}) = \det(r\mathfrak{c}(p) - \text{Id}) = r^2 \det(\mathfrak{c}(p) - r^{-1} \text{Id}). \tag{32}$$

By Proposition 2.7, the polynomial  $r \mapsto F(rp_x^2, rp_z^2)$  admits two positive roots  $0 < r_1 < r_2$ , and thus the matrix  $\mathfrak{c}(p)$  admits two positive eigenvalues  $0 < r_2^{-1} < r_1^{-1}$ , which concludes.  $\square$

**Remark 2.9** (The condition  $c_{13} + c_{44} > 0$ ). Consider a material obeying the admissibility conditions  $C_{\text{adm}}$ , except that  $c_{13} + c_{44} < 0$  rather than the opposite. Define  $c'_{13} := -c_{13} - 2c_{44}$ , in such way that  $c'_{13} + c_{44} = -(c_{13} + c_{44}) > 0$ .

The modified Hooke tensor coefficients  $(c_{11}, c'_{13}, c_{33}, c_{44})$  yield the same eikonal PDE as  $(c_{11}, c_{13}, c_{33}, c_{44})$ , since  $c_{13}^2 + 2c_{13}c_{44} - c_{11}c_{33} = (c_{13} + c_{44})^2 - c_{44}^2 - c_{11}c_{33}$  only depends on  $(c_{13} + c_{44})^2$  and the other terms of (6) are independent of  $c_{13}$ . The modified coefficients also meet the totality of the admissibility conditions (8), noting that  $c_{13}^{\prime 2} = c_{13}^2 + 4c_{44}(c_{44} + c_{13}) \leq c_{13}^2$ .

On the positive side, this discussion shows that our numerical method can handle (hypothetical) materials such that  $c_{13} + c_{44} < 0$ , through modified coefficients. On the negative side, this phenomenon illustrates an invariance of the TTI eikonal PDE, which therefore cannot be used to reconstruct the sign of  $c_{13} + c_{44}$  in a tomography context.

In the degenerate case where  $c_{13} + c_{44} = 0$ , the eikonal equation factors as  $F(x, z) = (1 - c_{44}x - c_{33}z)(1 - c_{11}x - c_{44}z)$ . Subject to the other admissibility conditions, the conic  $\mathcal{C}$  is then a union of two lines intersecting at the point  $(c_{33} - c_{44}, c_{11} - c_{44})/(c_{11}c_{33} - c_{13}^2)$  of the positive quadrant, as illustrated on Fig. 5 (right).

## 2.2 Region delimited by a conic

In this section we conclude the proof of Theorem 1.2, which describes the shape of slowness profile  $\mathcal{B}$  of the pressure waves, see Corollary 2.12. The ellipticity and separability of Hooke tensors defining TTI models, established in Section 2.1, are the key ingredient of the first result Proposition 2.10. In this section, we assume that  $(c_{11}, c_{13}, c_{33}, c_{44}) \in C_{\text{adm}}$  are admissible TTI parameters, see Theorem 1.2, and that  $\mathbf{c}$  and  $\sigma$  are the corresponding Hooke tensor (28) and coefficients (7) of the eikonal equation. The quadratic form  $\mathcal{Q} = \mathcal{Q}_\sigma$  and set  $\mathcal{B} = \mathcal{B}_\sigma$  are defined in (4). The regions  $\mathcal{A}$ ,  $\mathcal{B}$ , and conic  $\mathcal{C}$  are illustrated on Fig. 6.

**Proposition 2.10.** *The set  $\mathcal{B}' := \text{CC}_0\{(p_x, p_z) \in \mathbb{R}^2; \mathcal{Q}(p_x^2, p_z^2) \leq 1\}$  is compact and convex.*

*Proof.* Define  $N_{\mathbf{c}}(p) := \sqrt{\|\mathbf{c}(p)\|}$  for all  $p \in \mathbb{R}^2$ , where  $\|m\|$  denotes the spectral norm of a matrix  $m$ , which is also the largest eigenvalue if  $m \in \mathcal{S}_2^{++}$ . Since  $\mathbf{c}$  is a strictly elliptic Hooke tensor, as shown in Proposition 2.4, the function  $N_{\mathbf{c}}$  defines a norm over  $\mathbb{R}^2$ , by [DCC<sup>+</sup>21, Theorem 3.3]. As a result, the set  $B_{\mathbf{c}} := \{p \in \mathbb{R}^2; N_{\mathbf{c}}(p) \leq 1\}$  is compact and convex.

Since  $\mathbf{c}$  is separable, as shown in Corollary 2.8, one has  $B_{\mathbf{c}} = \text{CC}_0\{p \in \mathbb{R}^2; \det(\mathbf{c}(p) - \text{Id}) \geq 0\}$ , by [DCC<sup>+</sup>21, Proposition 3.7]. Recalling the identity (30) we conclude the proof.  $\square$

We present in Lemma 2.11 a simple criterion for the convexity of axisymmetric sets, which is applied to the slowness profile  $\mathcal{B}$  in Corollary 2.12, thus concluding the proof of Theorem 1.2.

**Lemma 2.11.** *Let  $E, F$  be normed vector spaces, and let  $K \subset \mathbb{R} \times F$  be convex and such that  $(-s, z) \in K$  for all  $(s, z) \in K$ . Then  $\{(x, z) \in E \times F; (|x|, z) \in K\}$  is convex.*

*Proof.* Let  $(x_1, z_1), (x_2, z_2) \in E \times F$ , and let  $t \in ]0, 1[$ . Define

$$s := \frac{|(1-t)x_1 + tx_2|}{(1-t)|x_1| + t|x_2|} \in [0, 1], \quad \alpha = (1-t)\frac{1+s}{2}, \quad \beta = (1-t)\frac{1-s}{2}, \quad \gamma = t\frac{1+s}{2}, \quad \delta = t\frac{1-s}{2},$$

choosing  $s \in [0, 1]$  arbitrarily if  $|x_1| = |x_2| = 0$ . Then  $\alpha, \beta, \gamma, \delta \geq 0$ ,  $\alpha + \beta + \gamma + \delta = 1$ , and

$$(|(1-t)x_1 + tx_2|, (1-t)z_1 + tz_2) = \alpha(|x_1|, z_1) + \beta(-|x_1|, z_1) + \gamma(|x_2|, z_2) + \delta(-|x_2|, z_2),$$

which establishes the announced convexity property.  $\square$

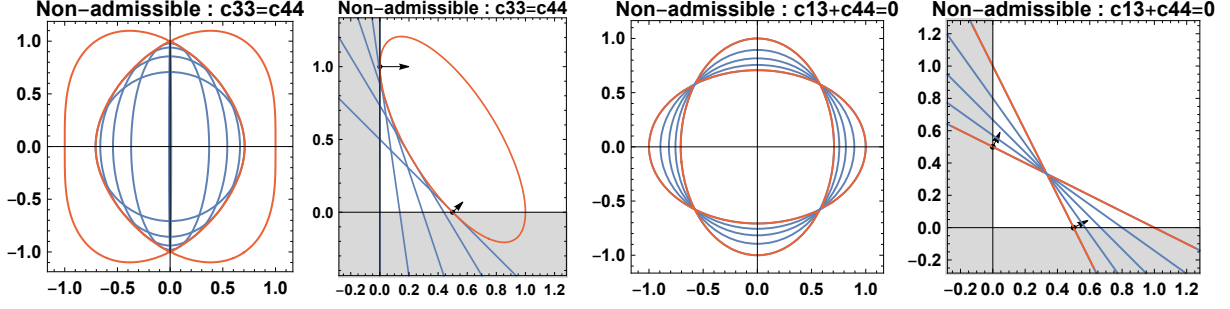


Figure 5: Examples where coefficients  $(c_{11}, c_{13}, c_{33}, c_{44})$  fail the admissibility conditions in such way that the Hooke elasticity tensor remains positive, but is not separable. As a result the inner slowness surface, associated with pressure waves, is non-smooth and intersects the outer slowness surface, associated with shear waves. Coefficients :  $(2, 0, 1, 1)$  for subfig. (i,ii), and  $(2, -1, 2, 1)$  for subfig. (iii,iv). Subfigures (i,iii): slowness surfaces (red) and tangent ellipsoids (blue). Subfigures (ii,iv): root domain with the conic  $\mathcal{C}$  (red), its tangent lines (blue), and normal vectors of Lemma 2.14 (black arrows).

**Corollary 2.12.** *The set  $\mathcal{B} := \text{CC}_0\{(p_x, p_y, p_z); \mathcal{Q}(p_x^2 + p_y^2, p_z^2) \leq 1\}$  is compact and convex.*

*Proof.* The closedness and boundedness of  $\mathcal{B}$ , hence compactness, follow immediately from the same properties of  $\mathcal{B}'$ , established in Proposition 2.10. Convexity follows from Lemma 2.11 applied to the set  $K = \mathcal{B}'$  from Proposition 2.10, choosing  $E = \mathbb{R}^2$  equipped with the Euclidean norm, and  $F = \mathbb{R}$ .  $\square$

**Remark 2.13** (Positivity without separability). *If one weakens the admissibility condition for the TTI coefficients (8), assuming only that  $(c_{11}, c_{13}, c_{33}, c_{44}) \in C_{\text{adm}}^1$ , see Proposition 2.4, then the Hooke tensor (28, left) remains positive but may not be separable. As a result, the P and SH slowness surfaces may intersect each other, see Fig. 5. Under these weaker assumptions, the open TTI unit ball  $\text{CC}_0\{(p_x, p_y, p_z) \in \mathbb{R}^3; \mathcal{Q}_\sigma(p_x^2 + p_y^2, p_z^2) < 1\}$  is bounded and convex but may have a non-smooth boundary, and likewise the solution  $u$  of the eikonal PDE (5) has lower regularity. Since no common geophysical material appears to fail the stronger  $C_{\text{adm}}$  conditions, see Appendix A, we limit our attention to those, eliminating a few mathematical technicalities in the process.*

The rest of this section is a preparation for the proof of Theorem 1.3, achieved in Section 2.3. In particular, the alternative between the (max) and (min) cases arises in Corollary 2.15 from the fact that a (connected component of a non-degenerate) conic curve has no inflexion point, and therefore has a convex side and concave side. We recall from Section 1.1 that

$$\mathcal{A} := \text{CC}_0\{(x, z) \in \mathbb{R}_+^2; \mathcal{Q}(x, z) \leq 1\}, \quad \mathcal{C} := \{(x, z) \in \mathbb{R}^2; \mathcal{Q}(x, z) = 1\}. \quad (33)$$

The set  $\mathcal{C}$  is a conic, in other words an algebraic set of degree two - which can thus be an ellipse, a hyperbola, a parabola, the union of two lines, etc, depending on the choice of  $\mathcal{Q}$ . A portion of this conic bounds the domain  $\mathcal{A}$ , which is the image of to the set  $\mathcal{B}'$  of Proposition 2.10 by a square root transformation.

Our first lemma describes two extremal points of the set  $\mathcal{A}$ , lying on the coordinate axes.

**Lemma 2.14.** *Define  $p_* := (1/c_{11}, 0)$  and  $p^* := (0, 1/c_{33})$ . Then  $p_*, p^* \in \mathcal{A} \cap \mathcal{C}$  and*

$$\begin{aligned} \nabla \mathcal{Q}(p_*) &= (c_{11}(c_{11} - c_{44}), (c_{13} + c_{44})^2 + (c_{11} - c_{44})c_{44}) / c_{11}, \\ \nabla \mathcal{Q}(p^*) &= ((c_{13} + c_{44})^2 + (c_{33} - c_{44})c_{44}, c_{33}(c_{33} - c_{44})) / c_{33}. \end{aligned}$$

It follows that  $\nabla \mathcal{Q}(p_*), \nabla \mathcal{Q}(p^*) \in ]0, \infty[^2$ .

*Proof.* By symmetry and w.l.o.g., we limit our attention to  $p_*$ . The polynomial  $1 - \mathcal{Q}(x, 0) = c_{11}c_{44}x^2 - (c_{11} + c_{44})x + 1$  admits the two roots  $1/c_{11}$  and  $1/c_{44}$ . Since  $c_{11} > c_{44}$ , by admissibility, one has  $\mathcal{Q}(x, 0) \leq 1$  iff  $x \in ]-\infty, 1/c_{11}] \cup [1/c_{44}, \infty[$  and thus  $(1/c_{11}, 0) \in \mathcal{A}$ . A direct computation yields the announced expression of  $\nabla \mathcal{Q}(p_*)$ , and the positivity of its components follows again from the admissibility conditions  $c_{11} > c_{44}$ ,  $c_{33} > c_{44}$  and  $c_{44} > 0$ .  $\square$

**Corollary 2.15.** *If  $\mathcal{I} \geq 0$  then  $\mathcal{A}$  is convex, and if  $\mathcal{I} \leq 0$  then  $\mathbb{R}_+^2 \setminus \mathcal{A}$  is convex, where*

$$\mathcal{I} := c_{11}c_{33} - c_{13}^2 - c_{11}c_{44} - 2c_{13}c_{44} - c_{33}c_{44}, \quad \mathcal{I} \propto \det(\nabla \mathcal{Q}(p_*), \nabla \mathcal{Q}(p^*)). \quad (34)$$

*Proof.* By Proposition 2.7 one has  $\mathcal{A} = \{r(\cos \theta, \sin \theta); 0 \leq \theta \leq \pi/2, 0 \leq r \leq R(\theta)\}$ . Therefore the boundary  $\partial \mathcal{A}$  is the union of the two segments  $[(0, 0), p_*]$  and  $[(0, 0), p^*]$ , and of the portion of conic  $\mathcal{A} \cap \mathcal{C} = \{R(\theta)(\cos \theta, \sin \theta); 0 \leq \theta \leq \pi/2\}$ . Likewise  $\partial(\mathbb{R}_+^2 \setminus \mathcal{A}) = [p_*, (\infty, 0)] \cup [p^*, (0, \infty)] \cup (\mathcal{A} \cap \mathcal{C})$  is the union of two half lines (with the obvious notation) and of the same portion of conic.

A direct computation yields the determinant of  $\nabla \mathcal{Q}(p_*)$  and  $\nabla \mathcal{Q}(p^*)$ , which are normal vectors to  $\mathcal{C}$  oriented outwards of  $\mathcal{A}$ , at the endpoints  $p_*$  and  $p^*$  of  $\mathcal{A} \cap \mathcal{C}$ . More precisely

$$\det(\nabla \mathcal{Q}(p_*), \nabla \mathcal{Q}(p^*)) = \mathcal{I}\mathcal{J}, \quad \text{where } \mathcal{J} := ((c_{13} + c_{44})^2 + c_{11}c_{33} - c_{44}^2) / (c_{11}c_{33}) > 0.$$

In order to establish the announced convexity properties, we distinguish two cases:

- Case of a degenerate conic  $\mathcal{C}$  (the union of two lines). Then  $\mathcal{A} \cap \mathcal{C} = [p_*, p^*]$  is a straight segment. Indeed, by Proposition 2.7, either the two lines are parallel, or their intersection lies outside  $[0, \infty]^2$ . As a result  $\mathcal{I} = 0$  (since the normal along a line is constant) and both  $\mathcal{A}$  and  $\mathbb{R}_+^2 \setminus \mathcal{A}$  are convex, as announced.
- Case of a non-degenerate conic (ellipse, hyperbola, parabola). Since a conic is a curve of degree two, it has no inflection point. Therefore the sign of the curvature is constant along  $\mathcal{C}$ , and thus either  $\mathcal{A}$  or  $\mathbb{R}_+^2 \setminus \mathcal{A}$  is convex, recalling the endpoint normals  $\nabla \mathcal{Q}(p_*), \nabla \mathcal{Q}(p^*) \in ]0, \infty[^2$ . Since the normal vectors along the boundary of a convex set are ordered trigonometrically in clockwise order, we obtain that  $\det(\nabla \mathcal{Q}(p_*), \nabla \mathcal{Q}(p^*)) \geq 0$  if  $\mathcal{A}$  is convex (resp.  $\det(\nabla \mathcal{Q}(p^*), \nabla \mathcal{Q}(p_*)) \leq 0$  if  $\mathbb{R}^2 \setminus \mathcal{A}$  is convex), which concludes the proof.  $\square$

### 2.3 Properties and computation of $\mu(\alpha)$

We establish Theorem 1.3 which describes the set  $\mathcal{A}$  as an intersection or a union of triangles, and thus  $\mathbb{B}$  as an intersection or a union of ellipses see Fig. 2, whose size is determined by a function  $\mu : [\alpha_*, \alpha^*] \rightarrow ]0, \infty[$ . We also prove that  $\mu$  is either convex or concave. The argument relies on Proposition 2.16 which is an elementary result on the support function of a convex set, see [BL10] for more detail on this rich subject. In the second part of this subsection, we establish that  $\mu$  is smooth and provide expressions of  $\mu, \alpha_*, \alpha^*$  suitable for numerical implementation.

**Proposition 2.16.** *The support function  $\mu_K : \mathbb{R}^d \rightarrow ]-\infty, \infty]$ , of a closed and convex set  $K \subset \mathbb{R}^d$ , is defined for all  $v \in \mathbb{R}^d$  as*

$$\mu_K(v) := \sup_{p \in K} \langle v, p \rangle.$$

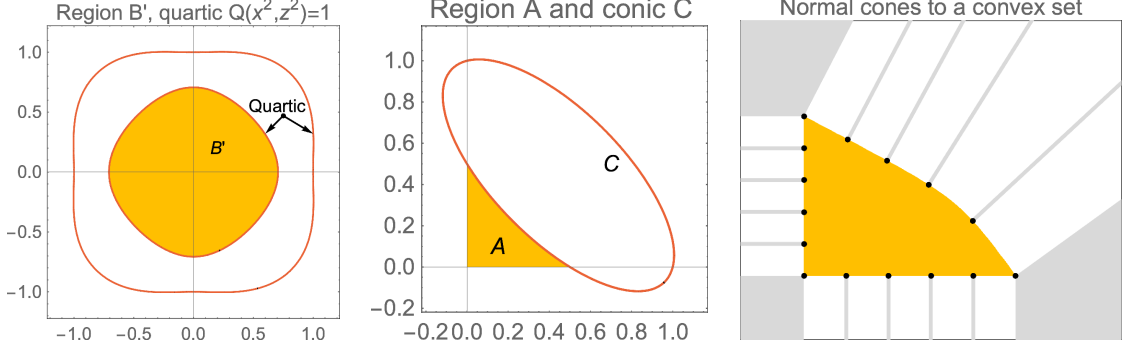


Figure 6: (Left) Quartic slowness surfaces, and  $xz$ -slice  $\mathcal{B}'$  of the anelliptic ball  $\mathcal{B}$ . (Center) Region  $\mathcal{A}$  and curve  $\mathcal{C}$  in the root domain. (Right) Normal cones, shown gray, to a convex set.

This function is convex and lower semi-continuous (l.s.c.), and furthermore

$$K = \{p \in \mathbb{R}^d; \forall v \in V, \langle v, p \rangle \leq \mu_K(v)\}, \quad (35)$$

provided the set  $V \subset \mathbb{R}^d$  contains a generator of each extreme ray of each normal cone to  $K$ .

The proof of Proposition 2.16 is postponed to the end of this section. We obtain in Corollaries 2.17 and 2.18 two descriptions of the set  $\mathcal{A}$ , announced in Theorem 1.3, concluding its proof except for the smoothness of the function  $\mu$  which follows from the explicit expression (43) below. They are deduced from the description (35) of convex sets as half-space intersections, and from the fact established in Corollary 2.15 that  $\mathcal{A}$  is either convex or the complement of a convex set. The endpoints  $p_*, p^*$  of the portion of conic  $\mathcal{A} \cap \mathcal{C}$  are defined in Lemma 2.14. We denote by  $\text{Cone}(E) = \{\sum_{i=1}^I \lambda_i e_i; I \geq 0, \lambda_1, \dots, \lambda_I \geq 0, e_1, \dots, e_I \in E\}$  the convex cone generated by non-negative linear combinations within a set of vectors  $E$ .

**Corollary 2.17.** Assume that  $\mathcal{A}$  is convex, which corresponds to the case (max). Define  $0 < \alpha_* \leq \alpha^* < 1$  by  $\nabla \mathcal{Q}(p_*) \propto (1 - \alpha_*, \alpha_*)$  and  $\nabla \mathcal{Q}(p^*) \propto (1 - \alpha^*, \alpha^*)$ , and let  $\mu(\alpha) := \mu_{\mathcal{A}}(1 - \alpha, \alpha)$ . Then  $\mu$  is convex and  $\mathcal{A} = \{p \in \mathbb{R}_+^2; \forall \alpha \in [\alpha_*, \alpha^*], \langle (1 - \alpha, \alpha), p \rangle \leq \mu(\alpha)\}$ .

*Proof.* Denote  $\nabla \mathcal{Q}_*(p_*) = (v_1, v_2)$ , and note that  $v_1, v_2 > 0$  by Lemma 2.14. Then the positive proportionality relation  $\nabla \mathcal{Q}_*(p_*) \propto (1 - \alpha_*, \alpha_*)$  admits the unique solution  $\alpha_* := v_2 / (v_1 + v_2) \in ]0, 1[$ . Likewise  $\alpha^* \in ]0, 1[$ , and furthermore by Corollary 2.15 we obtain as announced

$$0 \leq \det(\nabla \mathcal{Q}(p_*), \nabla \mathcal{Q}(p^*)) \propto [(1 - \alpha_*)\alpha^* - \alpha_*(1 - \alpha^*)] = \alpha^* - \alpha_*. \quad (36)$$

The function  $\mu(\alpha) := \mu_{\mathcal{A}}(1 - \alpha, \alpha)$  is convex since it is the composition of  $\mu_{\mathcal{A}}$ , which is convex by Proposition 2.16, with an affine mapping. Observing that  $\mathcal{A}$  is closed and convex, we obtain

$$\mathcal{A} = \{p \in \mathbb{R}^2; \forall v \in V, \langle v, p \rangle \leq \mu_{\mathcal{A}}(v)\}, \quad \text{with } V := \{(-1, 0), (0, -1)\} \cup \{(\alpha, 1 - \alpha); \alpha_* \leq \alpha \leq \alpha^*\}.$$

by Proposition 2.16, which implies the announced expression of  $\mathcal{A}$ . To show that  $V$  obeys the assumption of Proposition 2.16, we describe the normal cones to  $\mathcal{A}$ , illustrated on Fig. 6, for all points of the boundary  $\partial \mathcal{A} = [(0, 0), p_*] \cup [(0, 0), p^*] \cup (\mathcal{C} \cap \mathcal{A})$ . At the corners one has  $N_{\mathcal{A}}(0, 0) = \mathbb{R}_+^2 = \text{Cone}\{(-1, 0), (0, -1)\}$ ,  $N_{\mathcal{A}}(p_*) = \text{Cone}\{(-1, 0), (1 - \alpha_*, \alpha_*)\}$ , and  $N_{\mathcal{A}}(p^*) = \text{Cone}\{(-1, 0), (1 - \alpha^*, \alpha^*)\}$ . On the straight segments  $N_{\mathcal{A}}(p) = \text{Cone}\{(-1, 0)\}$  for all  $p \in ](0, 0), p_*[$ , and  $N_{\mathcal{A}}(p) = \text{Cone}\{(0, -1)\}$  for all  $p \in ](0, 0), p^*[$ . Finally for  $p \in (\mathcal{A} \cap \mathcal{C}) \setminus \{p_*, p^*\}$  one has  $N_{\mathcal{A}}(p) = \text{Cone}\{\nabla \mathcal{Q}(p)\} = \text{Cone}\{(1 - \alpha, \alpha)\}$  for some  $\alpha_* \leq \alpha \leq \alpha^*$ , since  $(1, 0) \preceq \nabla \mathcal{Q}(p_*) \preceq \nabla \mathcal{Q}(p) \preceq \nabla \mathcal{Q}(p^*) \preceq (0, 1)$  in the circular trigonometric ordering of vectors, by convexity of  $\mathcal{A}$ . The result follows.  $\square$

**Corollary 2.18.** *Assume that  $\mathbb{R}_+^2 \setminus \mathcal{A}$  is convex, which corresponds to case (min). Define  $0 < \alpha_* \leq \alpha^* < 1$  by  $\nabla \mathcal{Q}(p_*) \propto (1 - \alpha_*, \alpha_*)$  and  $\nabla \mathcal{Q}(p^*) \propto (1 - \alpha^*, \alpha^*)$ , and let  $\mu(\alpha) := -\mu_{\mathcal{A}^c}(-(1 - \alpha, \alpha))$  where  $\mathcal{A}^c := \mathbb{R}_+^2 \setminus \mathcal{A}$ . Then  $\mu$  is concave and  $\mathcal{A} = \{p \in \mathbb{R}_+^2; \exists \alpha \in [\alpha_*, \alpha^*], \langle (1 - \alpha, \alpha), p \rangle \leq \mu(\alpha)\}$ .*

*Proof.* Similarly to the proof of Corollary 2.17, one has  $\alpha_*, \alpha^* \in ]0, 1[$  by positivity of the gradient coordinates, see Lemma 2.14, and  $0 \geq \det(\nabla \mathcal{Q}(p_*), \nabla \mathcal{Q}(p^*)) \propto (\alpha_* - \alpha^*)$ . The function  $\mu$  is concave since it is the opposite of  $\mu_{\mathcal{A}^c}$ , which is convex by Proposition 2.16, composed with an affine mapping. In addition

$$\mathcal{A}^c = \{p \in \mathbb{R}^2; \forall v \in V, \langle v, p \rangle \leq \mu_{\mathcal{A}^c}(v)\}, \quad (37)$$

$$\mathbb{R}^2 \setminus \mathcal{A}^c = \{p \in \mathbb{R}^2; \exists v \in V, \langle v, p \rangle > \mu_{\mathcal{A}^c}(v)\}, \quad (38)$$

with  $V = \{(-1, 0), (0, -1)\} \cup \{-(1 - \alpha, \alpha); \alpha_* \leq \alpha \leq \alpha^*\}$ . Equation (37) follows from Proposition 2.16, where the assumption on  $V$  is checked as in Corollary 2.17 since  $\partial \mathcal{A}^c = [p_*, (\infty, 0) \cup [p^*, (0, \infty) \cup (\mathcal{A} \cap \mathcal{C})$ . Equation (38) is obtained by taking the complement. Noting that  $\mathcal{A} = \mathbb{R}_+^2 \cap (\mathbb{R}^2 \setminus \mathcal{A}^c)$  we conclude the proof.  $\square$

**Explicit formulas for implementation.** In the rest of this section, we obtain explicit formulas for the function  $\mu$  and the bounds  $\alpha_*$  and  $\alpha^*$ , suitable for implementing our numerical scheme, and announced below Theorem 1.3. For that purpose, we rewrite the quadratic function  $\mathcal{Q}$  defining the eikonal equation (3) in the following form:

$$\mathcal{Q}(p) = \langle l, p \rangle + \frac{1}{2} \langle p, Qp \rangle, \quad \text{hence } \nabla \mathcal{Q}(p) = l + Qp, \quad (39)$$

where  $l \in \mathbb{R}^2$  and  $Q \in S_2$  is a symmetric matrix. Indeed, our numerical implementation relies on the linear and quadratic forms defined by  $l$  and  $Q$ , rather than the coefficients  $(c_{11}, c_{13}, c_{33}, c_{44})$  which lead to expressions more complicated and of higher algebraic degree, and also restrict the generality. We nevertheless assume that all the guarantees derived previously apply.

By computing the smallest root  $x_*$  of the quadratic equation  $1 = \mathcal{Q}(x, 0) = l_1 x + Q_{11} x^2 / 2$ , we obtain the endpoint  $p_* = (x_*, 0)$  of  $\mathcal{A} \cap \mathcal{C}$ . Likewise we obtain the second endpoint  $p^*$ , and then  $\nabla \mathcal{Q}(p_*)$  and  $\nabla \mathcal{Q}(p^*)$  by (39, right). By computing the sign of  $\det(\nabla \mathcal{Q}(p_*), \nabla \mathcal{Q}(p^*))$  we distinguish between the cases (max) or (min) of Theorem 1.3. Also  $\alpha_*$  and  $\alpha^*$  are trivially obtained from  $\nabla \mathcal{Q}(p_*)$  and  $\nabla \mathcal{Q}(p^*)$ , as in Corollaries 2.17 and 2.18.

In the rest of this section, we fix  $\alpha \in ]\alpha_*, \alpha^*[$  and we denote  $\alpha := (1 - \alpha, \alpha)$ . Then one has

$$\mu(\alpha) = \max_{p \in \mathcal{A} \cap \mathcal{C}} \langle \alpha, p \rangle \quad \left( \text{resp. } \mu(\alpha) = \min_{p \in \mathcal{A} \cap \mathcal{C}} \langle \alpha, p \rangle \right), \quad (40)$$

in case (max) (resp. (min)). Indeed, the formula (40) is equivalent to the definitions presented in Corollaries 2.17 and 2.18 involving the support function of the set  $\mathcal{A}$  (resp.  $\mathbb{R}^2 \setminus \mathcal{A}$ ), because  $\alpha$  (resp.  $-\alpha$ ) is proportional their exterior normal at some point of the boundary  $\mathcal{A} \cap \mathcal{C}$ . In addition, since  $\alpha \in ]\alpha_*, \alpha^*[$ , the extremum (40) is attained at a point  $p \in \mathcal{A} \cap \mathcal{C}$  distinct from the endpoints  $p_*$  and  $p^*$ , hence to which the method of Lagrange multipliers is applicable (as opposed to the more complex KKT relations). As a result, and recalling that  $\mathcal{C} = \{\mathcal{Q} = 1\}$ , there exists a Lagrange multiplier  $\lambda \in \mathbb{R}$  such that

$$\nabla \mathcal{Q}(p) = \lambda \alpha, \quad \mathcal{Q}(p) = 1. \quad (41)$$

Rewriting (41, left) as  $l + Qp = \lambda\alpha$ , yields

$$p = Q^{-1}(\lambda\alpha - l). \quad (42)$$

We assume here that  $Q$  is an invertible matrix, which actually is not ensured by our admissibility assumptions (8). Further discussion of the case where  $Q$  is not invertible is postponed to the end of this section. Rewriting (9, right) as  $1 = \langle l, p \rangle + \frac{1}{2}\langle p, Qp \rangle$  and inserting (42) yields

$$1 = \langle l, Q^{-1}(\lambda\alpha - l) \rangle + \frac{1}{2}\langle (\lambda\alpha - l), Q^{-1}(\lambda\alpha - l) \rangle = \frac{1}{2}\lambda^2\langle \alpha, Q^{-1}\alpha \rangle - \frac{1}{2}\langle l, Q^{-1}l \rangle.$$

Therefore,

$$\lambda^2 = \frac{2 + \langle l, Q^{-1}l \rangle}{\langle \alpha, Q^{-1}\alpha \rangle}.$$

By (41, left), the scalar  $\lambda$  is the proportionality coefficient between the gradient  $\nabla\mathcal{Q}(p)$ , which has positive components for all  $p \in \mathcal{A} \cap \mathcal{C}$ , and the vector  $\alpha = (1 - \alpha, \alpha)$  which is likewise positive. Thus  $\lambda > 0$ , and the Lagrange multiplier  $\lambda$  is fully determined. Therefore

$$\mu(\alpha) := \langle \alpha, p \rangle = \lambda\langle \alpha, Q^{-1}\alpha \rangle - \langle \alpha, Q^{-1}l \rangle = \varepsilon\sqrt{\langle \alpha, Q^{-1}\alpha \rangle(2 + \langle l, Q^{-1}l \rangle)} - \langle \alpha, Q^{-1}l \rangle. \quad (43)$$

We denote by  $\varepsilon \in \{-1, 1\}$  the sign of  $\langle \alpha, Q^{-1}\alpha \rangle$ , which by (41) is also the sign of the expression

$$\langle \nabla\mathcal{Q}(p), Q^{-1}\nabla\mathcal{Q}(p) \rangle = \langle Qp + l, Q^{-1}(Qp + l) \rangle = \langle p, Qp \rangle + 2\langle p, l \rangle + \langle l, Q^{-1}l \rangle = 2 + \langle l, Q^{-1}l \rangle.$$

In particular,  $\varepsilon$  is independent of  $\alpha$ , and can be determined in advance from the coefficients  $Q$  and  $l$  of the PDE. From (43) we obtain that  $\mu$  has  $C^\infty$  regularity, as announced in Theorem 1.3, and is computable in a straightforward manner. We also recover the fact that it must be convex or concave, by an immediate application of the following lemma to the polynomial  $P(\alpha) := \langle \alpha, Q^{-1}\alpha \rangle$ .

**Lemma 2.19.** *Let  $P(t) := at^2 + bt + c$  be a second degree polynomial, with discriminant  $\Delta := b^2 - 4ac$ . If  $\Delta \geq 0$  (resp.  $\Delta \leq 0$ ) then  $\sqrt{P}$  is concave (resp. convex) on each connected component of  $\{t \in \mathbb{R}; P(t) > 0\}$ .*

*Proof.* This follows from a direct computation: assuming  $P(t) > 0$  one obtains

$$\frac{d^2}{dt^2}\sqrt{P(t)} = \frac{2P''(t)P(t) - P'(t)^2}{4P(t)^{\frac{3}{2}}} = \frac{2 \times 2a(at^2 + bt + c) - (2at + b)^2}{4P(t)^{\frac{3}{2}}} = -\frac{\Delta}{4P(t)^{\frac{3}{2}}}. \quad \square$$

**Case of a singular matrix  $Q$ .** We denote by  $R$  the adjugate matrix of  $Q$ , and let  $\delta := \det Q$  and  $\varepsilon' = \text{sign}(2\delta + \langle l, Rl \rangle)$ . Note that  $Q^{-1} = R/\delta$  if  $\delta \neq 0$ . We obtain from (43) and after straightforward manipulations (namely, the multiplication by the conjugate root) the following alternative expression of  $\mu$ . Its numerical evaluation is usually more stable than (43) when  $Q$  is singular or almost singular, since it does not involve  $Q^{-1}$

$$\mu(\alpha) = \frac{\det(\alpha, l)^2 + 2\langle v, Rv \rangle}{\varepsilon'\sqrt{\langle \alpha, R\alpha \rangle(2\delta + \langle l, Rl \rangle) + \langle \alpha, Rl \rangle}}. \quad (44)$$

One still needs to handle separately the degenerate case, at the intersection of the (max) and (min) cases, where  $\alpha_* = \alpha^*$  (in that case the conic  $\mathcal{C}$  is a union of two lines, and the TTI ball  $\mathcal{B}$  is an ellipsoid rather than a quartic surface).



*Proof of Proposition 2.16.* The function  $\mu_K$  is convex (resp. l.s.c.) since it is defined as the supremum of a family of linear functions, which are convex (resp. continuous hence l.s.c.) by definition. In the following, we denote by  $P_K : \mathbb{R}^d \rightarrow K$  the orthogonal projection, and by  $N_K(p_*) \subset \mathbb{R}^d$  the normal cone at a point  $p_* \in K$ , illustrated on Fig. 6, which admits the following equivalent characterizations:

$$v \in N_K(p_*) \Leftrightarrow \forall p \in K, \langle v, p_* - p \rangle \geq 0 \Leftrightarrow \mu_K(v) = \langle v, p_* \rangle \Leftrightarrow P_K(p_* + v) = p_*. \quad (45)$$

Denote (35, r.h.s.) by  $\tilde{K}$ , and note that  $K \subset \tilde{K}$  by definition of the support function. In the following, we consider  $p \notin K$ , and denote  $p_* := P_K(p)$  and  $v := p - p_*$ . For any  $q \in K$  one has  $\langle p - p_*, p_* - q \rangle \geq 0$  by general properties of the orthogonal projection, therefore  $\langle v, p \rangle \geq \|v\|^2 + \langle v, q \rangle$  by rearranging terms, and thus by taking the supremum over  $q \in K$

$$\langle v, p \rangle \geq \|v\|^2 + \mu_K(v) > \mu_K(v). \quad (46)$$

Since  $P_K(p_* + v) = P_K(p) = p_*$ , one has  $v \in N_K(p_*)$  by (45). By the Krein-Milman theorem, the cone  $N_K(p_*)$  is the convex hull of its extreme rays, and thus by assumption there exists  $\lambda_1, \dots, \lambda_N \geq 0$  and  $v_1, \dots, v_N \in V \cap N_K(p_*)$  such that  $v = \sum_{n=1}^N \lambda_n v_n$ , for some  $N \geq 0$ . Then, assuming for contradiction that  $p \in \tilde{K}$  we obtain:

$$\langle v, p \rangle = \sum_{1 \leq n \leq N} \lambda_n \langle v_n, p \rangle \leq \sum_{1 \leq n \leq N} \lambda_n \mu_K(v_n) = \sum_{1 \leq n \leq N} \lambda_n \langle v_n, p_* \rangle = \langle v, p_* \rangle = \mu_K(v). \quad (47)$$

We used successively (i) linearity, (ii) the assumption  $p \in \tilde{K}$ , (iii) the normal cone characterization (45), (iv) linearity again, and (v) again (45). Noting that (47) contradicts (46), we must have  $p \notin \tilde{K}$ . We have shown that  $p \notin K \Rightarrow p \notin \tilde{K}$ , which establishes the reverse inclusion  $K \supset \tilde{K}$ , and concludes the proof.  $\square$

### 3 Quasi-convexity or quasi-concavity of the update operator

We present two constructions of quasi-convex and quasi-concave functions in Section 3.1. By an adequate choice of parameters, they encompass the update operator associated to our finite differences discretization of the TTI eikonal PDE, which establishes Theorem 1.11. We study the primal metric associated to a TTI model using a similar strategy in Section 3.2, thus establishing Corollary 1.4. Interestingly, the proof differs in the (max) and (min) cases, a discrepancy also encountered in the convergence analysis Section 4.

#### 3.1 Two constructions of quasi-convex and quasi-concave functions.

We show that quasi-convex and quasi-concave functions, introduced in Definition 1.10, can be obtained as ratios of suitable functions in Lemma 3.1, and as implicit functions in Proposition 3.2.

For that purpose, we fix a convex subset  $A$  of a vector space, and recall from Definition 1.10 that a map  $f : A \rightarrow \mathbb{R}$  is quasi-convex iff  $\{x \in A; f(x) \leq \lambda\}$  is a convex set for all  $\lambda \in \mathbb{R}$ . Likewise we say that  $f$  is quasi-concave if  $-f$  is convex, equivalently iff  $\{x \in A; f(x) \geq \lambda\}$  is a convex set for all  $\lambda \in \mathbb{R}$ .

**Lemma 3.1.** *If  $f : A \rightarrow [0, \infty[$  is convex, and  $g : A \rightarrow ]0, \infty[$  is concave, then  $f/g$  is quasi-convex. Likewise if  $f : A \rightarrow ]0, \infty[$  is convex, and  $g : A \rightarrow [0, \infty[$  is concave, then  $g/f$  is quasi-concave.*

*Proof.* We only prove the first statement, since the second one is similar. Let  $\lambda \in \mathbb{R}$ . If  $\lambda < 0$ , then  $\{f/g \leq \lambda\} = \emptyset$  is convex. Otherwise  $\lambda \geq 0$  and  $\{f/g \leq \lambda\} = \{f - \lambda g \leq 0\}$  is convex since  $f - \lambda g$  is convex.  $\square$

**Proposition 3.2.** *Let  $F : A \times \mathbb{R} \rightarrow \mathbb{R}$  be such that (i)  $\alpha \in A \mapsto F(\alpha, \lambda)$  is quasi-convex (resp. quasi-concave) for all  $\lambda \in \mathbb{R}$ , and that (ii)  $\lambda \in \mathbb{R} \mapsto F(\alpha, \lambda)$  is non-decreasing for all  $\alpha \in A$ . Assume also that (iii)  $F(\alpha, \Lambda(\alpha)) = 0$  admits for all  $\alpha \in A$  a unique solution  $\Lambda(\alpha)$ , thus defining a mapping  $\Lambda : A \rightarrow \mathbb{R}$ . Then  $\Lambda$  is quasi-concave (resp. quasi-convex).*

*Proof.* We limit our attention to the case where  $\alpha \mapsto F(\alpha, \lambda)$  is quasi-convex, since the second case is similar. By (ii) and (iii) one obtains  $\Lambda(\alpha) \geq \lambda \Leftrightarrow F(\alpha, \lambda) \leq 0$ , for any  $\alpha \in A$ ,  $\lambda \in \mathbb{R}$ . Thus

$$\{\alpha \in A; \Lambda(\alpha) \geq \lambda\} = \{\alpha \in A; F(\alpha, \lambda) \leq 0\},$$

for any  $\lambda \in \mathbb{R}$ . Noting by (i) that the r.h.s. is a convex set, we obtain that  $\Lambda$  is quasi-concave, which concludes the proof.  $\square$

**Proof of Theorem 1.11, on the update operator quasi-convexity or quasi-concavity.**

We proceed to apply Proposition 3.2 to a function of the following form, defined in view of the expression (20) of the numerical scheme for the TTI eikonal PDE,

$$F(\alpha, \lambda) = \frac{1}{\mu(\alpha)} \sum_{1 \leq i \leq I} \rho_i(\alpha) \max\{0, \lambda - u_i\}^2. \quad (48)$$

Specifically using the notations of Proposition 1.7 and (18), we fix  $q \in h\mathbb{Z}^d$  and  $0 \leq k < K$ , define  $A = [\alpha_k, \alpha_{k+1}]$  which is a segment of  $\mathbb{R}$ , and let  $u_i = \min\{u(q + he_{ik}), u(q - he_{ik})\}$ , for all  $1 \leq i \leq I$  where  $u : h\mathbb{Z}^d \rightarrow ]-\infty, \infty]$  is the unknown of the finite difference scheme. (Recall that  $u$  is finite on  $\Omega_h$  and extended by  $+\infty$  elsewhere.) Then  $F(\alpha, \lambda) = \hat{\mathfrak{F}}^\alpha(q, [\lambda - u(r)]_{r \in \bar{X}})$  is the numerical scheme (20) with the base point value  $u(q)$  replaced with the unknown  $\lambda$ , consistently with the formulation of the update operator (15).

By Theorem 1.3 the function  $\mu : A \rightarrow ]0, \infty[$  is convex in case (max) (resp. concave in case (min)). By Proposition 1.7 the functions  $\rho_i : A \rightarrow [0, \infty[$  are affine. For any given  $\lambda \in \mathbb{R}$  the sum  $f(\alpha) := \sum_{i=1}^I \rho_i(\alpha) \max\{0, \lambda - u_i\}^2$  is thus non-negative and affine w.r.t.  $\alpha \in A$ , thus simultaneously convex and concave. Lemma 3.1 therefore yields that  $\alpha \in A \mapsto F(\alpha, \lambda)$  is quasi-concave in case (max) (resp. quasi-convex in case (min)).

The partial mapping  $\lambda \in \mathbb{R} \mapsto F(\alpha, \lambda)$  is non-decreasing, for any  $\alpha \in A$ , since the weights  $\rho_i(\alpha)$  are non-negative and since  $\lambda \mapsto \max\{0, \lambda\}^2$  is non-decreasing. As already observed in Section 1.2 there is a unique solution  $\Lambda : A \rightarrow \mathbb{R}$  to the equation  $F(\alpha, \Lambda(\alpha)) = 1$  (one has  $\Lambda(\alpha) := \Lambda^\alpha u(q)$  with the notations of Theorem 1.11). Applying Proposition 3.2 to  $F - 1$ , we obtain that the update operator  $\Lambda$  is quasi-convex in case (max) (resp. quasi-concave in case (min)), which concludes the proof of Theorem 1.11.

### 3.2 Expression of the norm value, gradient, and dual.

Given admissible TTI parameters  $\sigma \in \mathbb{R}^3$ , and a co-vector  $p = (p_x, p_y, p_z) \in \mathbb{R}^3 \setminus \{0\}$ , we obtain

$$\begin{aligned} \mathcal{F}_\sigma^*(p) &:= \min\{\nu > 0; p/\nu \in \mathcal{B}_\sigma\} \\ &= \min\{\nu > 0; (p_x^2 + p_y^2, p_z^2)/\nu^2 \in \mathcal{A}_\sigma\} \\ &= \min\{\nu > 0; \forall \alpha \in [\alpha_*, \alpha^*], (1 - \alpha)(p_x^2 + p_y^2) + \alpha p_z^2 \leq \nu^2 \mu(\alpha)\} \end{aligned} \quad (49)$$

$$= \max_{\alpha \in [\alpha_*, \alpha^*]} \sqrt{\frac{(1 - \alpha)(p_x^2 + p_y^2) + \alpha p_z^2}{\mu(\alpha)}}, \quad (50)$$

assuming case (max) of Theorem 1.3 in (49). We used successively (i) the norm definition (5, right), (ii) the definitions (9) and (4) of the sets  $\mathcal{A}_\sigma$  and  $\mathcal{B}_\sigma$ , (iii) Theorem 1.3 in case (max), and (iv) a direct algebraic computation. Alternatively, in case (min) of Theorem 1.3, the universal quantifier  $\forall$  of (49) is replaced with an existential quantifier  $\exists$ , and as a result the max operator in (50) is replaced with the min operator. The announced expression (11) of  $\mathcal{F}^*(p) := \mathcal{F}_\sigma^*(Rp)$  follows. The expression (12, left) of the gradient  $\nabla \mathcal{F}^*(p)$  then follows from the envelope theorem [Car01, Theorem 6.1], on the differentiation of functions defined as an extremum.

We turn to the computation of the dual norm (12, right), which is obtained as follows

$$\begin{aligned}
\frac{1}{2}\mathcal{F}(v)^2 &= \max_p \left( \langle p, v \rangle - \frac{1}{2}\mathcal{F}^*(p)^2 \right) \\
&= \max_p \overline{\text{mix}}_{\alpha \in [\alpha_*, \alpha^*]} \left( \langle p, v \rangle - \frac{1}{2\mu(\alpha)} \|p\|_{D(\alpha)}^2 \right) \\
&= \overline{\text{mix}}_{\alpha \in [\alpha_*, \alpha^*]} \max_p \left( \langle p, v \rangle - \frac{1}{2\mu(\alpha)} \|p\|_{D(\alpha)}^2 \right) \\
&= \overline{\text{mix}}_{\alpha \in [\alpha_*, \alpha^*]} \frac{\mu(\alpha)}{2} \|v\|_{D(\alpha)}^2.
\end{aligned} \tag{51}$$

We used successively, (i) Legendre-Fenchel duality, which is a generalization of norm duality, (ii) the explicit expression of  $\mathcal{F}^*$ , recalling that  $\{\overline{\text{mix}}, \overline{\text{mix}}\} = \{\min, \max\}$ , see (10), (iii) an interversion of the extremum operators max and  $\overline{\text{mix}}$ , discussed in more detail below, and (iv) the known explicit expression of the Legendre-Fenchel dual of the positive quadratic form  $p \mapsto \langle p, D(\alpha)p \rangle / \mu(\alpha)$ .

As announced, we discuss in more detail the interversion of max and  $\overline{\text{mix}}$  in (51), and in this occasion we need to distinguish the treatment of the (min) and (max) case, associated with Theorem 1.3. If  $\overline{\text{mix}} = \max$ , then (51) amounts to a maximization over the joint variable  $(p, \alpha) \in \mathbb{R}^3 \times [\alpha_*, \alpha^*]$ , hence the order of the maximizations is irrelevant and the result is proved. On the other hand, if  $\overline{\text{mix}} = \min$ , then we invoke Sion's minimax theorem [Kom88] to exchange the ordering of the min and max operators, whose assumptions are checked below. Define, for  $\alpha \in [\alpha_*, \alpha^*]$  and  $p \in \mathbb{R}^3$

$$F(\alpha, p) := \langle p, v \rangle - \frac{1}{2}F_0(\alpha, p), \quad \text{where } F_0(\alpha, p) := \frac{1}{\mu(\alpha)} \|p\|_{D(\alpha)}^2.$$

Then  $p \mapsto F_0(\alpha, p)$  is a positive quadratic form, hence is a convex function. On the other hand  $\alpha \mapsto F_0(\alpha, p)$  is quasi-concave by Lemma 3.1, since it is the ratio of the non-negative and affine (hence concave) function  $\|p\|_{D(\alpha)}^2 = (1 - \alpha)\|p\|_{D_0}^2 + \alpha\|p\|_{D_1}^2$ , divided by  $\mu(\alpha)$  which is convex by Theorem 1.3 in case (max) (recall that we assume  $\overline{\text{mix}} = \min$  here, and see (10)). The function  $F$  thus matches the assumptions of Sion's minimax theorem, as announced:  $F$  is quasi-convex w.r.t.  $\alpha$ , concave w.r.t.  $p$  (hence also quasi-concave), and in addition we note that  $F$  is continuous, that  $[\alpha_*, \alpha^*]$  is convex and compact, and that  $\mathbb{R}^3$  is convex. This completes the proof.

## 4 Convergence analysis

We prove that the solutions to our discretization of the TTI eikonal PDE, obey a Lipschitz regularity property in the (max) case, and a weaker growth estimate in the (min) case, from which we deduce their convergence as the grid scale is refined, see Theorem 4.1. This discrepancy between the (max) and (min) cases illustrates the fact that PDE operators presented in (generalized) Bellman form, i.e. as a maximum of simpler monotone operators, are usually more

easily amenable to analysis than those presented as a minimum. Such Bellman forms are at the foundation of multistencil fast marching methods [HF07], and of discretizations of second order PDEs with general coefficients [Kry05], as well as the special cases of the Monge-Ampere [BM21] or Pucci [BBM20] PDEs. In the case of the TTI eikonal PDE, we only obtain a Bellman form in the (max) case, and by Theorem 1.3 this depends on the model coefficients.

We believe that the difference between the (max) and (min) cases, both in the proof technique and in the obtained regularity results (53) and (54), is interesting since it departs from the symmetrical treatment of these two cases in the introduction. Nevertheless, we do establish in both cases the convergence to the viscosity solution of the TTI eikonal PDE as the grid scale is refined, at least in the simplified setting of null Dirichlet boundary conditions on  $\partial\Omega$  (as opposed to the point source and outflow boundary conditions often considered in applications), see Theorem 4.1. Let us also mention that, empirically, our numerical experiments Section 5 do not show a difference in behavior between the (max) and (min) cases.

We fix an open and bounded domain  $\Omega \subset \mathbb{R}^d$ , where  $d \in \{2, 3\}$ , with a smooth boundary. Given  $h > 0$  we let  $\Omega_h := \Omega \cap h\mathbb{Z}^d$ , and  $\partial\Omega_h := h\mathbb{Z}^d \setminus \Omega_h$ . The notation  $C = C(\Omega, \sigma, R)$  means that the constant  $C$  only depends on the specified parameters.

**Theorem 4.1.** *Consider continuous TTI coefficients  $\sigma \in C^0(\overline{\Omega}, \mathbb{R}^5)$  obeying the admissibility conditions (8) pointwise, and a continuous field of invertible matrices  $R \in C^0(\overline{\Omega}, \text{GL}_d(\mathbb{R}))$ . Then there exists a unique solution  $u : h\mathbb{Z}^d \rightarrow [0, \infty[$ , to*

$$\mathfrak{F}u(q) = 1, \forall q \in \Omega_h, \quad u(q) = 0, \forall q \in \partial\Omega_h, \quad (52)$$

where  $\mathfrak{F}$  stands for the proposed finite differences discretization of the TTI eikonal PDE (20). In addition, there exists a constant  $C = C(\Omega, \sigma, R)$  such that for any  $h > 0$  sufficiently small and for all  $q, r \in h\mathbb{Z}^d$ .

- If the parameters  $\sigma$  fall in the (max) case of Theorem 1.3 over the whole  $\overline{\Omega}$ , then

$$|u(q) - u(r)| \leq C|q - r|. \quad (53)$$

- For arbitrary parameters  $\sigma$ , possibly mixing of the (max) and (min) cases over  $\overline{\Omega}$ , one has

$$u(q) \leq \max\{u(r'); r' \in h\mathbb{Z}^d, |r' - r| \leq Ch\} + C|q - r|. \quad (54)$$

In both cases, one has  $\|u_h - \mathbf{u}\|_{L^\infty(\Omega_h)} \rightarrow 0$  as  $h \rightarrow 0$ , where  $u_h : \Omega_h \rightarrow \mathbb{R}$  denotes the discrete solution to (52), and  $\mathbf{u} : \overline{\Omega} \rightarrow \mathbb{R}$  denotes the unique viscosity solution of the TTI eikonal PDE

$$\mathcal{F}_q^*(\nabla \mathbf{u}(q)) = 1, \forall q \in \Omega, \quad \mathbf{u}(q) = 0, \forall q \in \partial\Omega, \quad (55)$$

where we denoted  $\mathcal{F}_q^*(v) := \mathcal{F}_{\sigma(q)}^*(R(q)v)$ , for all  $q \in \Omega$ .

Before turning to the proof, we recall the definition of a sub-solution or super-solution to a numerical scheme, whereas the corresponding PDE notions are briefly evoked below (70).

**Definition 4.2.** *We say that  $u : h\mathbb{Z}^d \rightarrow \mathbb{R}$  is a sub-solution to a scheme  $\mathfrak{F}$  on  $\Omega_h$  if*

$$\mathfrak{F}u(q) \leq 1, \forall q \in \Omega_h, \quad u(q) = 0, \forall q \in \partial\Omega_h. \quad (56)$$

Likewise, we define the notions of strict sub-solution, solution, super-solution, and strict super-solution, by replacing the comparison operator with  $<$ ,  $=$ ,  $\geq$ ,  $>$ , in (56, left) respectively.

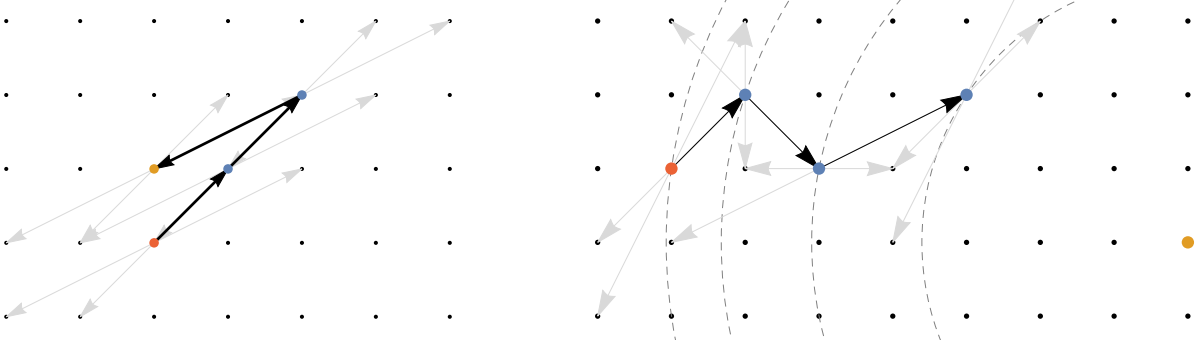


Figure 7: Left: In the (max) case, one can exploit consistency in the scheme stencils, and use the fact that they span  $\mathbb{Z}^d$ , so as to join two neighbor points using a chain of offsets. Right: In the (min) case, the active stencils at the different points may be uncorrelated. Their offsets can be used to move toward a given target, up to some radius, but not to reach it exactly in general.

**Notations.** Selling's decomposition is denoted  $D = \sum_{e \in \mathbb{Z}^d} \rho(e; D) e e^\top$ , consistently with Appendix B and (77); this notation avoids introducing an arbitrary indexing  $(e_i)_{i=1}^I$  of the active offsets  $\{e \in \mathbb{Z}^d; \rho(e; D) > 0\}$ , and is thus more convenient for discussing the regularity of the weights  $D \mapsto \rho(e; D)$ , see Proposition B.4. Throughout this section, the quantities associated in Theorem 1.3 to admissible TTI parameters define the following functions pointwise on  $\bar{\Omega}$ :

$$\alpha_*, \alpha^* \in C^0(\bar{\Omega}, ]0, 1[), \quad \mu \in C^0(\mathcal{A}, ]0, \infty[),$$

where  $\mathcal{A} := \{(\alpha, q) \in ]0, 1[ \times \bar{\Omega}; \alpha_*(q) \leq \alpha \leq \alpha^*(q)\}$ . For all  $\alpha \in ]0, 1[$  and  $q \in \bar{\Omega}$  we let

$$D(q, \alpha) := R(q)^\top \begin{pmatrix} 1 - \alpha & & \\ & 1 - \alpha & \\ & & \alpha \end{pmatrix} R(q). \quad (57)$$

#### 4.1 Lipchitz property in case (max)

The main result of this subsection is a Lipschitz regularity property for sub-solutions to the numerical scheme  $\mathfrak{F}^D$  discretizing the Riemannian eikonal PDE (18), see Proposition 4.4. The Lipschitz estimate (53) in case (max) of Theorem 4.1 is then deduced. Note that the weaker growth estimate (54), valid in all cases, suffices for the proof of convergence in Section 4.3. We nevertheless present the Lipschitz estimate since it is simple, expected, and since the proof exploits a number of properties of Selling's matrix decomposition, gathered in Proposition B.4, which is central in the method. A similar Lipschitz regularity result is proved in [Mir14a, Lemma 2.7] for a different discretization of the Riemannian eikonal PDE.

A preliminary technical lemma defines, between any two neighbor points on the grid, a chain whose length is bounded above, and such that successive points are connected by offsets of Selling's decomposition of a given matrix field, see Fig. 7 (left).

**Lemma 4.3.** *Given a field of symmetric positive definite matrices  $\mathcal{D} \in C^0(\bar{\Omega}, S_d^{++})$ , there exists  $h_0 > 0$ ,  $\rho_0 > 0$  and  $N_0$  such that the following holds. Let  $0 < h < h_0$ , and let  $q_*, q^* \in h\mathbb{Z}^d$  be such that  $|q_* - q^*| = h$ . Then there exists a chain  $q_0, \dots, q_N \in h\mathbb{Z}^d$  of length  $N \leq N_0$ , whose endpoints are  $q_0 = q_*$  and  $q_N = q^*$ , and signs  $\varepsilon_1, \dots, \varepsilon_N \in \{-1, 1\}$ , such that*

$$\rho((q_{n+1} - q_n)\varepsilon_n/h; \mathcal{D}(q_n)) \geq \rho_0, \text{ for any } 0 \leq n < N \text{ such that } q_n \in \Omega_h. \quad (58)$$

*Proof.* Since the matrix field  $\mathcal{D}$  is pointwise positive definite and continuous, it is bounded over the compact set  $\overline{\Omega}$ , as well as its inverse and condition number, which fits the assumptions of Proposition B.4 on the properties of Selling's decomposition. We can assume  $q_* \in \Omega_h$ , since otherwise the condition (58) is empty for  $n = 0$ , and the trivial chain of length  $N = 1$  complies.

By Proposition B.4 (spanning property), there exists a direct basis  $e_1, \dots, e_d$  of  $\mathbb{Z}^d$  such that  $\rho(e_i; \mathcal{D}(q_*)) \geq 2\rho_0$  for all  $1 \leq i \leq d$ , where  $\rho_0 = \rho_0(\mathcal{D})$ . By Proposition B.4 (Lipschitz weights), the functions  $q \in \overline{\Omega} \mapsto \rho(e_i; \mathcal{D}(q))$  are continuous, hence there exists  $r_0 = r_0(\mathcal{D}) > 0$  such that:

$$\rho(e_i; \mathcal{D}(q)) \geq \rho_0, \quad \text{for all } 1 \leq i \leq d, \text{ and all } q \in \overline{\Omega} \text{ s.t. } |q - q_*| \leq r_0. \quad (59)$$

By Proposition B.4 (bounded offsets), one has  $|e_i| \leq R_0$  for all  $1 \leq i \leq d$ , where  $R_0 = R_0(\mathcal{D})$ . Defining the matrix  $G := [e_1, \dots, e_d]$ , and noting that  $\det(G) = \det(e_1, \dots, e_d) = 1$ , we obtain that  $G^{-1}$  has integer coefficients bounded in absolute value by  $R_1 = R_1(\mathcal{D})$ . Denote  $e := (q^* - q_*)/h$ , recall that this vector or its opposite belongs to the canonical basis of  $\mathbb{R}^d$  by assumption, and let  $(\lambda_1, \dots, \lambda_d) = G^{-1}e$ , in such way that  $e = \lambda_1 e_1 + \dots + \lambda_d e_d$ ,  $|\lambda_1|, \dots, |\lambda_d| \leq R_1$  and  $\lambda_1, \dots, \lambda_d \in \mathbb{Z}$ . Assuming w.l.o.g. that  $\lambda_1, \dots, \lambda_d \geq 0$  we define for all  $n \leq N := \lambda_1 + \dots + \lambda_d$ :

$$q_n = q_* + h(\lambda_1 e_1 + \dots + \lambda_r e_r + \lambda_{r+1} e_{r+1}), \quad \text{where } n = \lambda_1 + \dots + \lambda_r + \lambda,$$

with  $\lambda$  integer such that  $1 \leq \lambda \leq \lambda_{r+1}$ . Observe that  $N \leq N_0$  where  $N_0 = N_0(\mathcal{D}) := dR_1$ , and that  $|q_n - q_*| \leq hN_0R_0$  is smaller than  $r_0$  provided  $h \leq h_0$  where  $h_0 = h_0(\mathcal{D}) = r_0/(N_0R_0)$ , for any  $0 \leq n \leq N$ . This construction of  $q_0, \dots, q_N$  satisfies in view of (59) the announced properties, which concludes the proof.  $\square$

**Proposition 4.4.** *Let  $\mathcal{D} \in C^0(\overline{\Omega}, S_d^{++})$ , and let  $u : h\mathbb{Z}^d \rightarrow [0, \infty[$  obey*

$$\mathfrak{F}^{\mathcal{D}(q)} u(q) \leq 1, \forall q \in \Omega_h, \quad u(q) = 0, \forall q \in \partial\Omega_h. \quad (60)$$

*Then  $|u(q) - u(r)| \leq C|q - r|$  for all  $q, r \in h\mathbb{Z}^d$ , where  $h > 0$  is small enough and  $C = C(\mathcal{D})$ .*

*Proof.* It suffices to prove that  $|u(q) - u(r)| \leq C|q - r|$  when  $|q - r| = h$  are neighbors on the grid  $h\mathbb{Z}^d$  (up to multiplying  $C$  by  $\sqrt{d}$ ). It also suffices to prove the one sided inequality  $u(q) \leq u(r) + C|q - r|$ , by symmetry.

Assumption (60, left) at a point  $q \in \Omega_h$  can be rewritten as

$$\sum_{e \in \mathcal{Z}} \rho(e; \mathcal{D}(q)) \max\{0, u(q) - u(q - he), u(q) - u(q + he)\}^2 \leq h^2,$$

in view of the Riemannian scheme definition (18). Therefore  $u(q) \leq u(q + he) + h\rho(\varepsilon e; \mathcal{D}(q))^{-\frac{1}{2}}$  for any  $e \in \mathbb{Z}^d$  and any sign  $\varepsilon \in \{-1, 1\}$ , with the convention  $0^{-\frac{1}{2}} = \infty$ . Let  $q = q_0, \dots, q_N = r$  be a chain as described in Lemma 4.3, joining the points of interest. Then

$$u(q_n) \leq u(q_{n+1}) + h\rho_0^{-\frac{1}{2}}, \quad (61)$$

for all  $0 \leq n < N$ . Indeed this follows from (58) and the previous estimate when  $q_n \in \Omega_h$ , and otherwise  $u(q_n) = 0$  by the boundary condition satisfies the bound since  $u$  is non-negative. Accumulating these inequalities we obtain  $u(q) \leq u(r) + hN_0\rho_0^{-\frac{1}{2}}$ , as announced.  $\square$

*Proof of (53) in Theorem 4.1, using Proposition 4.4.* Assume that  $\mathfrak{F}u(q) \leq 1$  for all  $q \in \Omega_h$ , and  $u(q) = 0$  for all  $u \in \partial\Omega_h$ . Define  $\mathcal{D}(q) := D(q, \alpha_*(q))/\mu(q, \alpha_*(q))$ , and note that  $\mathcal{D} \in C^0(\overline{\Omega}, S_d^{++})$ . Since the TTI scheme (20) is defined as a maximum of Riemannian schemes parameterized by  $\alpha \in [\alpha_*(q), \alpha^*(q)]$ , we obtain that  $1 \geq \mathfrak{F}u(q) \geq \mathfrak{F}^{\mathcal{D}(q)} u(q)$ . Thus Proposition 4.4 applies and (53) is proved.  $\square$

## 4.2 Growth estimate in case (min)

The main result of this subsection is a growth estimate for sub-solutions to a minimum of discretized eikonal PDEs, established in Proposition 4.6. The growth estimate (54) is then deduced. The proof strategy differs from Section 4.1 since we cannot exploit any local consistency between the active stencils of close discretization points. Instead, Lemma 4.5 below shows that we can use Selling offsets whose weights are positive to move in the general direction (62) of a given point, assumed to be far enough, see Fig. 7 (right).

**Lemma 4.5.** *Given a compact set  $\mathfrak{D} \subset S_d^{++}$ , there exists  $\rho_0 > 0$ ,  $r_0 > 0$  and  $R_0$  such that the following holds. For any  $D \in \mathfrak{D}$  and any  $v \in \mathbb{R}^d$  with  $|v| \geq R_0$ , there exists an offset  $e \in \mathbb{Z}^d$  and a sign  $\varepsilon \in \{-1, 1\}$  such that*

$$\rho(\varepsilon e; D) \geq \rho_0 \text{ and } |v - e| \leq |v| - r_0. \quad (62)$$

*Proof.* Since  $\mathfrak{D}$  is compact, its elements are bounded, and likewise their inverses and condition numbers. Let  $D \in \mathfrak{D}$  and  $v \in \mathbb{R}^d$ , then

$$\lambda_0 |v|^2 \leq \|v\|_D^2 = \sum_{e \in \mathbb{Z}^d} \rho(e; D) \langle v, e \rangle^2 \leq I \max_{e \in \mathbb{Z}^d} \rho(e; D) \langle v, e \rangle^2, \quad (63)$$

where we used successively (i) a lower bound  $\lambda_0 = \lambda_0(\mathfrak{D})$  on the eigenvalues of  $D \in \mathfrak{D}$ , (ii) Selling's formula (74), and (iii) the fact that Selling's decomposition involves at most  $I := d(d+1)/2$  positive weights. Thus there exists  $e \in \mathbb{Z}^d$ , a maximizer of (63, right), such that  $\rho(e; D) \langle v, e \rangle^2 \geq |v|^2 \lambda_0 / I$ . Observing that  $|e| \leq R_1 = R_1(\mathfrak{D})$  by Proposition B.4 (bounded offsets), we obtain  $\rho(e; D) \geq \rho_0$  where  $\rho_0 = \rho_0(\mathfrak{D}) := \lambda_0 / (R_1^2 I)$ .

On the other hand  $\rho(e; D) \leq \rho_1$  where  $\rho_1 = \rho_1(\mathfrak{D}) := \max\{\text{Tr}(D); D \in \mathfrak{D}\}$ , since  $\rho(D) \leq \rho(D) \|e\|^2 \leq \sum_{e \in \mathbb{Z}^d} \rho(e; D) \|e\|^2 = \text{Tr}(D)$ . Therefore  $|\langle v, e \rangle| \geq 2r_0 |v|$  with  $r_0 = r_0(\mathfrak{D}) = \frac{1}{2} \sqrt{\lambda_0 / I \rho_1}$ . Then, assuming w.l.o.g. that  $\langle v, e \rangle \geq 0$ , we obtain

$$\|v\| - \|v - e\| \geq \frac{\langle e, v \rangle}{|v|} - \frac{|e|^2}{2|v|} \geq 2r_0 - \frac{R_1^2}{2R_0},$$

using successively (i) [Mir19, Lemma 2.10], and (ii) the upper bounds on  $\langle v, e \rangle$  and  $|e|$ , and the lower bound on  $|v|$ . Defining  $R_0 = R_0(\mathfrak{D}) := R_1^2 / (2r_0)$  we conclude the proof.  $\square$

**Proposition 4.6.** *Let  $\mathfrak{D} \subset S_d^{++}$  be a compact set, and let  $u : h\mathbb{Z}^d \rightarrow [0, \infty[$  obey*

$$\min_{D \in \mathfrak{D}} \mathfrak{F}^D u(q) \leq 1, \forall q \in \Omega_h, \quad u(q) = 0, \forall q \in \partial\Omega_h. \quad (64)$$

*Then for any  $q, r \in h\mathbb{Z}^d$ , one has with  $R = R(\mathfrak{D})$  and  $C = C(\mathfrak{D})$*

$$u(q) \leq \max\{u(r'); |r' - r| \leq Rh\} + C|q - r|. \quad (65)$$

*Proof.* We claim that the announced result holds with the constants  $R = R_0$  and  $C = r_0^{-1} \rho_0^{-\frac{1}{2}}$ , where  $R_0$ ,  $r_0$  and  $\rho_0$  are from Lemma 4.5. For that purpose, we fix the point  $r \in h\mathbb{Z}^d$ , and prove the announced result for all  $q \in h\mathbb{Z}^d$ , by induction on  $|(q - r)/h|^2$  which is a positive integer. Note that if  $q \notin \Omega_h$ , then  $u(q) = 0$  by the boundary condition, and (65) holds. Also (65) clearly holds if  $|q - r| \leq Rh$ .

The assumption (64, left) at  $q \in \Omega_h$  such that  $|q - r| \geq Rh$ , can be rewritten as

$$\sum_{e \in \mathbb{Z}} \rho(e; D) \max\{0, u(q) - u(q - he), u(q) - u(q + he)\}^2 \leq h^2, \quad (66)$$

for some  $D \in \mathfrak{D}$ , in view of the Riemannian scheme definition (18). Denoting  $v := (q - r)/h$ , and noting that  $|v| \geq R$ , we find by Lemma 4.5 an offset  $e \in \mathbb{Z}^d$  such that  $q' := q - he$  satisfies,

$$u(q) \leq u(q') + h\rho_0^{-\frac{1}{2}}, \quad |q' - r| \leq |q - r| - hr_0,$$

using (66) for the first estimate. The announced result follows by induction.  $\square$

*Proof of the growth estimate (54).* Define the set of positive definite matrices

$$\mathfrak{D} := \{D(q, \alpha)/\mu(q, \alpha); q \in \overline{\Omega}, \alpha \in [\alpha_*(q), \alpha^*(q)]\}, \quad (67)$$

which is compact since the functions  $D, \mu, \alpha_*, \alpha^*$  are continuous over a compact domain. Then (52, left) implies (64, left), and the announced growth estimate (54) is established in (65).  $\square$

### 4.3 Proof of convergence

We follow a standard proof strategy [BR06] to establish the uniform convergence of the solutions to the proposed discretization scheme of the TTI eikonal PDE, henceforth denoted  $\mathfrak{F}_h$  where  $h > 0$  is the grid scale, towards the continuous solution as  $h \rightarrow 0$ . Note that alternative proof strategies exist which may yield stronger quantitative results, including convergence rates, see Remark 4.8.

As a first step, in Lemma 4.7, we establish the existence of a sub-solution and of a super-solution to  $\mathfrak{F}_h$ , which are bounded independently of  $h$ , as well as an approximation property of super-solutions by strict super-solutions.

**Lemma 4.7.** *The proposed discretization scheme  $\mathfrak{F}_h$  of the TTI eikonal PDE (20) satisfies:*

- (Explicit sub-solution) *The null function  $\bar{u} = 0$  satisfies  $\mathfrak{F}_h \bar{u} = 0$  identically on  $h\mathbb{Z}^d$ .*
- (Explicit super-solution) *Let  $a \in \mathbb{R}$ ,  $b \in \mathbb{R}^d$  be such that  $a + \langle b, q \rangle > 0$  for all  $q \in \overline{\Omega}$ , and  $\mathcal{F}_q^*(b) \geq 1$  for all  $q \in \overline{\Omega}$ . Let  $\underline{u}(q) := a + \langle b, q \rangle$  on  $\Omega$ , and  $\underline{u} = 0$  on  $\mathbb{R}^d \setminus \Omega$ . Then  $\mathfrak{F}_h \underline{u} \geq 1$  on  $\Omega_h$ , for any sufficiently small  $h > 0$ .*
- (Approximation of super-solutions) *One has  $\mathfrak{F}_h[\lambda u(q)] = \lambda^2 \mathfrak{F}_h u(q)$  for any  $u : h\mathbb{Z}^d \rightarrow \mathbb{R}$ ,  $\lambda \geq 0$ , and  $q \in \Omega_h$ . In particular, if  $u$  is a super-solution of  $\mathfrak{F}_h$ , then  $(1 + \varepsilon)u$  is a strict super-solution for any  $\varepsilon > 0$ , converging to  $u$  as  $\varepsilon \rightarrow 0$ .*

*Proof.* The 2-homogeneity property, announced in the last point, is obvious in view of the definition of the Riemannian (18) and TTI (20) schemes. The points (Explicit sub-solution) and (Perturbation of sub-solution) follow; for instance if  $\mathfrak{F}_h u \geq 1$  then  $\mathfrak{F}_h[(1 + \varepsilon)u] = (1 + \varepsilon)^2 \mathfrak{F}_h u > 1$  for any  $\varepsilon > 0$ .

The constants  $a \in \mathbb{R}$ ,  $b \in \mathbb{R}^d$ , of the second point exist by compactness of  $\overline{\Omega}$  and continuity and definiteness of the norms  $\mathcal{F}_q^*$ ,  $q \in \overline{\Omega}$ . Define  $v(q) = a + \langle b, q \rangle$  on  $\mathbb{R}^d$ , and note that  $v(q + he) \geq 0$  for any point  $q \in \Omega$  and offset  $\|e\| \leq R_0$ , where  $R_0$  is a bound on the scheme stencil radius, by continuity and provided the discretization scale  $h$  is small enough. It follows under these conditions that  $\underline{u}(q + he) \leq v(q + he)$ . Then for any  $q \in \Omega_h$

$$1 \leq \mathcal{F}_q^*(b) = \mathfrak{F}_h v(q) \leq \mathfrak{F}_h \underline{u}(q), \quad (68)$$

using successively (i) the assumption on  $b$ , (ii) the scheme consistency (21), and (iii) the DDE property of the scheme, see Definition 1.6, and the observation that  $\underline{u}(q + he) \leq v(q + he)$  for all offsets  $e$  of the scheme stencil, whereas  $\underline{u}(q) = v(q)$  since  $q \in \Omega$ . The result follows.  $\square$



By [Mir19, Theorem 2.3] there exists a unique solution  $u_h : h\mathbb{Z}^d \rightarrow \mathbb{R}$  to the scheme  $\mathfrak{F}_h$ . For context, uniqueness is established using the *comparison principle*, whereas existence is proved by *Perron's method* (maximal sub-solution). Both of these classical techniques require a scheme obeying the DDE property, see Definition 1.6. Since in addition the scheme  $\mathfrak{F}_h$  is causal, see again Definition 1.6, its solution may be computed using the single pass FMM on  $\Omega_h$ , see Algorithm 1.

The scheme solution  $u_h : \Omega_h \rightarrow \mathbb{R}$  is bounded above and below by any super- and sub-solution, hence choosing those of Lemma 4.7 we obtain the bounds  $\bar{u} \leq u_h \leq \underline{u}$  on  $\Omega_h$  which are independent of  $h$ . This allows us to consider the lower and upper limits  $\bar{\mathbf{u}}, \underline{\mathbf{u}} : \bar{\Omega} \rightarrow \mathbb{R}$ , defined for all  $q \in \bar{\Omega}$  as

$$\bar{\mathbf{u}}(q) := \liminf_{h \rightarrow 0, q_h \rightarrow q} u_h(q_h), \quad \underline{\mathbf{u}}(q) := \limsup_{h \rightarrow 0, q_h \rightarrow q} u_h(q_h), \quad (69)$$

where implicitly  $q_h \in h\mathbb{Z}^d$ . By construction  $0 = \bar{u} \leq \bar{\mathbf{u}} \leq \underline{\mathbf{u}} \leq \underline{u}$  on  $\bar{\Omega}$ , and  $\bar{\mathbf{u}}$  is lower semi-continuous whereas  $\underline{\mathbf{u}}$  is upper semi-continuous. By the growth estimate (54) one has  $0 \leq \bar{\mathbf{u}}(q) \leq \underline{\mathbf{u}}(q) \leq C \text{dist}(q, \partial\Omega)$ , hence  $\bar{\mathbf{u}}$  and  $\underline{\mathbf{u}}$  obey the null Dirichlet boundary condition on  $\partial\Omega$ . By the DDE property of the scheme  $\mathfrak{F}_h$ , and by consistency (21), passing to the limit one obtains that in the sense of viscosity solutions<sup>6</sup> [BCD08]

$$\mathcal{F}_q^*(\nabla \bar{\mathbf{u}}(q)) \geq 1, \quad \mathcal{F}_q^*(\nabla \underline{\mathbf{u}}(q)) \leq 1, \quad (70)$$

for all  $q \in \Omega$ , with the notation  $\mathcal{F}_q^*$  of (55). By the continuous comparison principle [BCD08, Theorem 5.9], we obtain that  $\underline{\mathbf{u}} \leq \bar{\mathbf{u}}$ . Hence  $\mathbf{u} = \bar{\mathbf{u}} = \underline{\mathbf{u}}$  is a viscosity solution to (55), and one has  $u_h \rightarrow \mathbf{u}$  uniformly as  $h \rightarrow 0$  by (69). This concludes the proof of Theorem 4.1.

**Remark 4.8.** A quantitative convergence rate  $\|u_h - \mathbf{u}\|_{L^\infty(\Omega_h)} = \mathcal{O}(\sqrt{h})$  as  $h \rightarrow 0$ , improving on the uniform convergence result of Theorem 4.1, can likely be established by following the same scheme of proof as [Mir19, §2.1], and by assuming Lipschitz regularity for the TTI parameters  $\sigma$  and  $R$ . However this would introduce a number of technicalities, such as the doubling of variables argument [Eva10], that we have chosen to avoid here since they are not specifically related to the models of interest.

## 5 Numerical experiments

In this section, we present numerical experiments on three-dimensional test cases so as to evaluate the cost and accuracy of our TTI eikonal solver.

First, we consider a TTI medium with a semi-analytical solution to determine the convergence order and computation time of our numerical scheme. In the numerical experiments, we compare the two versions of our scheme to solve the underlying 1D-optimization problem: quasi-convex optimization (with CPU implementation), and optimization by grid search (with GPU implementation), see Section 1.3. In the latter case, we also study the influence of the sampling rate, denoted  $K$  in (23), on the solution accuracy. We find that the GPU implementation is fifty times faster than the CPU implementation in this test-case.

We then consider two alternative eikonal solvers, able to handle speed propagation profiles either (i) *less* or (ii) *more* general than TTI anisotropy. (i) A standard isotropic fast marching solver [Set96], enhanced with source factorization and second order finite differences, addresses isotropic (spherical) speed profiles. (ii) The state of the art CPU eikonal solver [DCC<sup>+</sup>21], referred to as the “general scheme”, handles anisotropy associated with a full Hooke tensor, of

<sup>6</sup>In the sense of viscosity solutions, (70) should be understood as follows: let  $\varphi \in C^2(\Omega)$  be arbitrary. If  $\bar{u} - \varphi$  attains its minimum at  $q \in \Omega$ , then  $\mathcal{F}_q^*(\nabla \varphi(q)) \geq 1$ . If  $\varphi - \underline{u}$  attains its minimum at  $q \in \Omega$ , then  $\mathcal{F}_q^*(\nabla \varphi(q)) \leq 1$ .

which TTI anisotropy is a special case, see Appendix A. The comparison is done on a medium with orthorhombic anisotropy, with an analytical solution to the eikonal equation, and its projections to the closest TTI medium and isotropic medium. This experiment allows to quantify and compare the *discretization error*, associated to the grid scale, with the *consistency error*, related to the approximation of the anisotropic speed propagation profile. We also compare the computation time and accuracy of the different schemes.

Last, we consider an application to a realistic synthetic test-case which comes from the homogenization of an isotropic medium. The resulting anisotropy is fully general, but is expected to be close to TTI anisotropy because of the sedimentary structure of the medium. We verify this assumption by considering a projection of the general medium to a TTI medium and to an isotropic medium, and compare the results of the general scheme, TTI scheme and isotropic scheme. We also compare the solution to the eikonal equation with the solution to the elastic wave equation, to verify that the solution to the eikonal equation indeed is consistent with the first-arrival traveltimes of the wave propagation.

All the computations presented here have been performed on the Univ. Grenoble Alpes HPC perform. For the numerical scheme in the CPU case as well as for the general scheme, the computation has been performed on one Intel node equipped with a Xeon Skylake Gold processor, with a core clocked 2.1 GHz and 192 GO of RAM. For the numerical scheme in the GPU case as well as for the isotropic scheme, the computation has been performed using an Nvidia Tesla V100 with 5120 CUDA cores and 96 GO of RAM.

The fast marching method only uses a single CPU core due to its intrinsic sequential nature. The massively parallel solver presently uses a single GPU card, but its multi-GPU extension is an opportunity for future work, possibly along the lines of [HJ16] in the isotropic setting.

## 5.1 Numerical application on a synthetic case obtained from the conformal transformation of a TTI metric

We consider a semi-analytical test case, so as to investigate the convergence rate and numerical error of our numerical scheme. It is a non-trivial heterogeneous TTI metric, obtained from a conformal diffeomorphic transformation of a homogeneous TI metric. By design, the exact solution to the eikonal equation is known and easily evaluated numerically to machine precision.

Conformal transformations are helpful to create non-trivial media with known solutions: indeed, the Jacobian of a conformal transformation  $\phi : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$  is a scaled rotation, namely  $\text{Jac } \phi(x) = \alpha(x)R(x)$  with scaling parameter  $\alpha(x) > 0$  and rotation matrix  $R(x)$ . Our test medium is obtained as the pull-back of a homogeneous TI medium by  $\phi$ , so that the associated TTI eikonal PDE (1) parameters take the form  $(\frac{a}{\alpha(x)^4}, \frac{b}{\alpha(x)^4}, \frac{c}{\alpha(x)^4}, \frac{d}{\alpha(x)^2}, \frac{e}{\alpha(x)^2})$  and  $R_0 R(x)$  pointwise, where  $(a, b, c, d, e)$  and  $R_0$  are fixed. One can likewise define the pull back by a conformal transformation of an eikonal equation which is isotropic, or whose anisotropy is defined by a general Hooke tensor [DCC<sup>+</sup>21, Appendix A].

Three dimensional conformal transformations include dilations, translations, rotations, the inversion  $x \in \mathbb{R}^3 \setminus \{0\} \mapsto x/\|x\|^2$ , and compositions of these. In our experiments we use a “special conformal transformation”, defined by:  $\phi(x) := \frac{x - b\|x\|^2}{1 - 2\langle b, x \rangle + \|b\|^2\|x\|^2}$ . It is smooth except for a singularity at  $b/\|b\|^2$ , where  $b \in \mathbb{R}^3$  is a parameter. It is obtained as the composition of an inversion, a translation by  $-b$ , and another inversion. We choose  $b := (1/6, 1/9, 1/18)$  and let  $\tilde{\Omega} := ]-1, 1[^3$  with seed at the origin, so that the singular point  $b/\|b\|^2 \notin \tilde{\Omega}$ , and the image domain  $\Omega := \phi(\tilde{\Omega})$  is star shaped with respect to the origin, see Figure 8.

We consider the homogeneous TI metric from the mica medium [BC91], defined by:

$$\begin{pmatrix} 178 & 42.4 & 14.5 & 0 & 0 & 0 \\ 42.4 & 178 & 14.5 & 0 & 0 & 0 \\ 14.5 & 14.5 & 54.9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 12.2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 12.2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 12.2 \end{pmatrix}, \quad \rho = 2.79\text{g/cm}^3,$$

which we rotate by  $3\pi/5$  with Euler axis  $(2, 1, 3)$ , in such way that the transverse isotropy plane is in a generic position rather than axis aligned (the latter may unfairly advantage eikonal solvers based on a Cartesian discretization grid such as ours).

The TTI metric from the mica corresponds to a maximization case, see (2). We also consider the same setting with different materials, such as the stishovite medium [BC91] which corresponds to a minimization case. Remarkably, in our numerical experiments, the two different cases yield completely similar computation time and error convergence, despite the difference in the formulation of the numerical scheme and in the mathematical proof of convergence Section 4. As the results are very close, we only illustrate the case of the mica medium. A cross-section of the solution to the eikonal equation is presented in Fig 9, which shows how the solution  $u$  relative to the constant metric on a transformed domain (right figure) translates to a solution  $\tilde{u}$  relative to a non-trivial metric on the regular cube domain (left figure) with the conformal transformation.

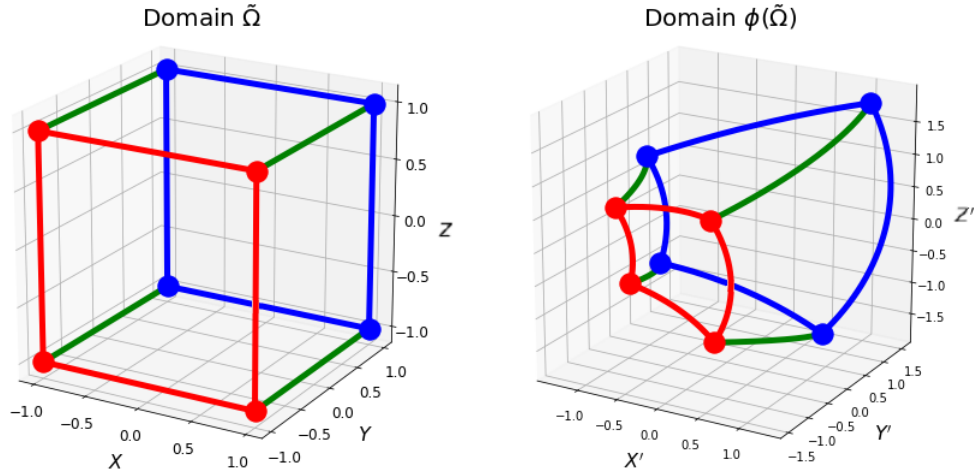


Figure 8: Edges of the domain  $\tilde{\Omega} = ]-1, 1[^3$  (a cube) and of its image  $\Omega = \phi(\tilde{\Omega})$  by a special conformal transformation.

In the CPU implementation of our TTI eikonal solver, the 1D optimization problem underlying our scheme is solved up to machine precision, by taking advantage of its quasi-convexity, see Theorem 1.11. As a result, the numerical accuracy is directly related to the scale  $h$  of the Cartesian discretization grid, and thus to the number of points  $N \approx |\Omega|/h^3$ . Second order  $\mathcal{O}(h^2)$  convergence rates are observed on Fig. (10, right), as expected since we use second order finite differences, see Appendix C.

In the GPU implementation, there is an additional source of approximation, related to the sampling parameter  $K$  in the optimization by grid search of the 1D optimization problem underlying our scheme, see (23). As illustrated on Fig. 1, this amounts to approximating the slowness

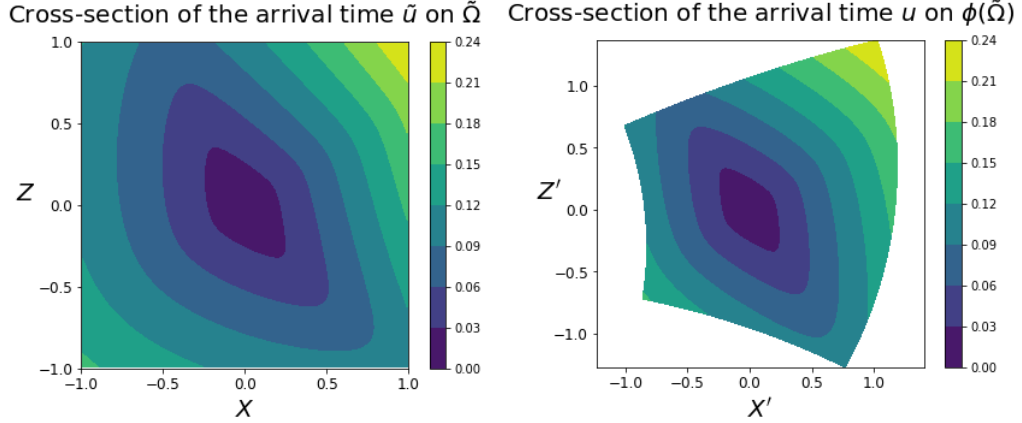


Figure 9: Cross-section at  $Y = 0$  of the solution to the eikonal equation for a non-trivial TTI metric on  $\tilde{\Omega}$  (left), which corresponds to a constant metric on the transformed domain  $\Omega = \phi(\tilde{\Omega})$  (right).

surface of the pressure wave with the union or the intersection of  $k := K + 1$  ellipses. When  $k$  is fixed, the numerical error of the scheme first decreases as the grid scale is refined, until a plateau is reached. This could be expected from the  $\mathcal{O}(h^2 + k^{-2})$  consistency error of the scheme, with second order finite differences as here, see Proposition 1.9. The plateau occurs for a number of ellipses  $k$  approximately proportional to the inverse grid scale  $h^{-1}$ . This scheme exhibits second-order convergence for small domain sizes, provided  $k$  is large enough, but the convergence rate then slightly degrades for the finest grid scales  $h$ ; from the theoretical standpoint, convergence is guaranteed by Theorem 4.1, but not a specific rate.

In our numerical experiment, we consider from 10 to 26 ellipses. The size of the medium also goes from  $39 \times 39 \times 39$  to  $217 \times 217 \times 217$ . This upper limit on the size of the domain comes from a memory limit on the GPU, rather than a limit on computation time. The computation time is quasi-linear w.r.t the total number of points for both the CPU and GPU eikonal solvers, and increases with the number  $k$  of ellipses in the latter case. The CPU implementation is about twice more accurate than the GPU implementation, for the examples considered in Fig. 10, but comes at the cost of a computation time more than 10 times larger.

**Remark 5.1** (Elliptic approximation). *For computational efficiency purposes, one can be tempted to approximate the TTI eikonal equation with a Riemannian eikonal equation, and thus the algebraic  $P$ -wave slowness surface with a single ellipsoid. This corresponds to a special case of our scheme, with a single ellipse ( $k = 1$ ), and thus a trivial grid search. For the test case considered here, the numerical error almost immediately reaches a plateau, as in Fig. 10, and the scheme converges towards an erroneous solution, which shows the importance of properly taking into account a TTI anisotropy compared with elliptic anisotropy.*

## 5.2 Projection error for the orthorhombic anisotropy

Orthorhombic anisotropy is a more general type of anisotropy compared with the TTI anisotropy, and can be found in some crystalline structures. An orthorhombic medium does not exhibit the rotational symmetry that is found in TTI media, and corresponds to a Hooke tensor with nine independent elastic parameters, compared with five for the TTI anisotropy. The eikonal equation

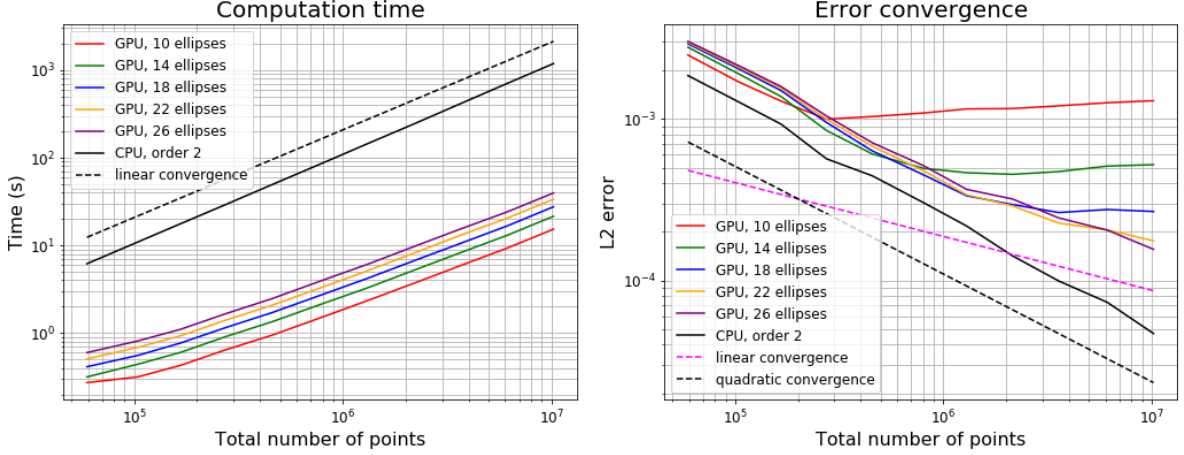


Figure 10: Computation time and error of the numerical scheme, depending on the number of ellipses for the GPU implementation. For visual interpretation, dotted lines illustrate linear convergence  $\mathcal{O}(h)$  and quadratic convergence  $\mathcal{O}(h^2)$ .

with orthorhombic anisotropy can be solved by the anisotropic variant of the Fast Marching method presented in [DCC<sup>+</sup>21], which we refer to as the “general scheme”.

We consider the projection of an orthorhombic Hooke tensor to the closest Hooke tensor with TTI anisotropy. We have two goals in mind: first, we want to compare the computation time and the accuracy of our numerical scheme with the general scheme, and second, we want to quantify the projection error caused by the approximation in the anisotropy. Likewise, we consider the projection from the orthorhombic to an isotropic Hooke tensor, which we solve with an isotropic Fast Marching method implemented on GPU.

We consider the orthorhombic anisotropy defined from the olivine medium [BC91]:

$$\begin{pmatrix} 323.7 & 66.4 & 71.6 & 0 & 0 & 0 \\ 66.4 & 197.6 & 75.6 & 0 & 0 & 0 \\ 71.6 & 75.6 & 235.1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 64.6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 78.7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 79.0 \end{pmatrix}, \quad \rho = 3.311 \text{g/cm}^3,$$

Similarly to the previous subsection, we use a conformal transformation to create a non-trivial heterogeneous metric with a known solution. The projection of the Hooke tensor and the pull-back of the metric by the conformal transformation can be done in any order, so the TTI and isotropic settings also correspond to a conformal transformation of a homogeneous metric, with known solution.

We illustrate on Fig 11 the slowness surfaces related to the olivine (orthorhombic medium) and its TI and isotropic projections. Contrarily to TI and isotropic metrics, the orthorhombic anisotropy does not possess a rotational symmetry, and so we show two cross-sections of the corresponding slowness surfaces. For the eikonal equation, we are only interested in the inner surface, related to the fastest speed.

We show in Table 1 the results of the different numerical schemes on the different media. The domain size is  $77 \times 77 \times 77$ . For the GPU case, we consider 10 ellipses for the optimization by grid search.

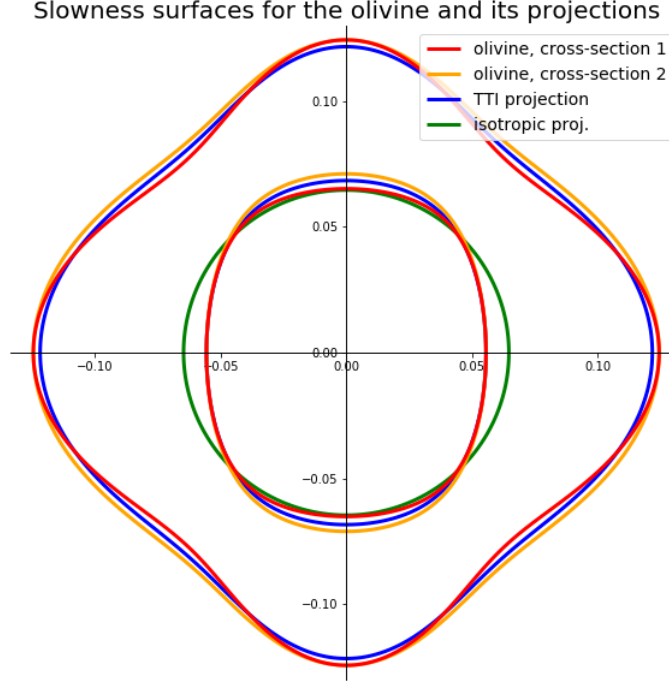


Figure 11: Slowness surfaces for the olivine and its TTI and isotropic projections. The TTI and isotropic projections have a rotational symmetry along the vertical axis, but not the olivine which is an orthorhombic medium, so we present two cross-sections of the slowness surface for the olivine along two vertical plans. Only the inner surfaces are of interest for the eikonal equation, related to the fastest speed.

We observe that the general scheme and the TTI scheme with CPU have a similar computation time, and that the TTI scheme with GPU is approximately fifty times faster. The isotropic scheme with GPU is also fifty times faster compared with the TTI scheme with GPU. The error due to the numerical scheme is much smaller than the error due to the approximation in the anisotropy of the medium: indeed, we observe that the numerical  $L^2$ -error and the exact  $L^2$ -error are almost identical. Besides, the  $L^2$ -error is more than two times bigger when comparing the TTI projection to the isotropic projection, going from an error of 2.27% to 5.31%.

With this example, we see that we need to make a choice between the accuracy of the numerical scheme, the accuracy of the anisotropy model, and the computation time required to solve the corresponding eikonal equation. It can be interesting to consider media with fully general Hooke tensors to better encompass the anisotropy of a geophysical medium, especially for orthorhombic media, as we see that the TTI projection leads to errors. However, for use-cases in seismic imaging which do not deviate too much from TTI anisotropy, such as the one presented in the next section, the TTI scheme is a better choice: the computation time is greatly improved as it is approximately fifty times faster than the general scheme, which can open the door to efficient applications in seismic imaging.

### 5.3 Anisotropic media coming from the homogenization of an isotropic medium

In the case of a medium with a sedimentary structure, the local invariance by rotation around the normal axis to the layers is expected, which leads to TTI anisotropy. Even if the materials of the medium are intrinsically isotropic, the medium can be represented with TTI anisotropy

	Comp. time (s)	$L^2$ -error (num.)	$L^2$ -error (exact)
General scheme on orth. med.	22.6	0.000109	0
TTI scheme (CPU) on TTI med.	10.4	0.0227	0.0226
TTI scheme (GPU) on TTI med.	0.545	0.0227	0.0226
Isotr. scheme (GPU) on isotr. med.	0.0125	0.0531	0.0546

Table 1: Computation on a synthetic test-case of size  $77 \times 77 \times 77$ , with orthorhombic anisotropy and its projection to TTI anisotropy and isotropy. The  $L^2$ -error (num.) corresponds to the difference between the numerical solution on the corresponding medium and the exact solution on the orthorhombic medium. The  $L^2$ -error (exact) corresponds to the difference between the exact solution on the corresponding medium and the exact solution on the orthorhombic medium. All the  $L^2$ -errors are normalized by the  $L^2$  norm of the exact solution on the orthorhombic medium. The general scheme is the numerical scheme from [DCC<sup>+</sup>21], used with second-order accuracy and “cut-cube” setting. The TTI scheme with GPU is used with 10 ellipses. The isotropic scheme is a Fast Marching scheme with second-order accuracy using GPU.

if the typical wavelength of seismic waves is larger than the typical size of the layers caused by the sedimentation, and this can be done through the homogenization process.

In this subsection, we study the application of our numerical scheme on a realistic dataset, which comes from the homogenization of an isotropic medium into a model with fully general anisotropy: the Hooke tensor has 21 independent elastic parameters. The isotropic model is the SEG/EAGE overthrust model, see [ABK97] and Fig. 12, and information on the homogenized model can be found on [CMA<sup>+</sup>20]. We consider the projection of the general medium into a TTI medium, with the projection being made on the Hooke tensor at each point of the domain, see [CMA<sup>+</sup>20]. We then study the relevance of this TTI projection by comparing the solution in the medium with general anisotropy to the solution in the medium with TTI anisotropy, by using the scheme from [DCC<sup>+</sup>21] which can handle general anisotropy.

We also consider the projection of the homogenized medium to an isotropic medium, which is a way to study the strength of the anisotropy coming from the homogenization process in this medium. We use the Fast Marching isotropic scheme with GPU to solve the corresponding eikonal equation.

The medium is discretized on a  $107 \times 534 \times 534$  grid, corresponding to a real medium of dimensions  $4 \text{ km} \times 20 \text{ km} \times 20 \text{ km}$ . The source point is placed on the point  $(0, 267, 267)$ , which corresponds to the middle of the medium on the surface. In order to use the multi-scale source factorization described in Appendix C, we need a finer discretization near the source. For that purpose, we interpolate the value of the general metric near the source with a trilinear interpolation on each coefficients of the Hooke tensor, and then we use the projection to TTI metric again.

First, we use the scheme from [DCC<sup>+</sup>21], which can solve the eikonal equation with a Hooke tensor of general anisotropy, and use it with second-order precision. With this scheme, we get the solution to the eikonal equation on both the general metric and the TTI metric, and we can consider these two results as the closest we have to the exact solutions. We also compute the solution on the TTI metric with our numerical scheme, with both CPU and GPU implementations. For the GPU case, we use 10 ellipses for the optimization by grid search, and checked that using a higher number of ellipses does not significantly change the result.

On Fig 12, we show a superposition of the solution to the eikonal equation with the solution to the wave equation in the same medium. The elastic wave propagation problem is solved using the spectral-element based modeling and inversion code SEM46 [TBM<sup>+</sup>19, CBM20]. We

observe that the isochrones computed from our eikonal solver properly follows the wavefront of the solution to the wave equation at the corresponding time, as expected.

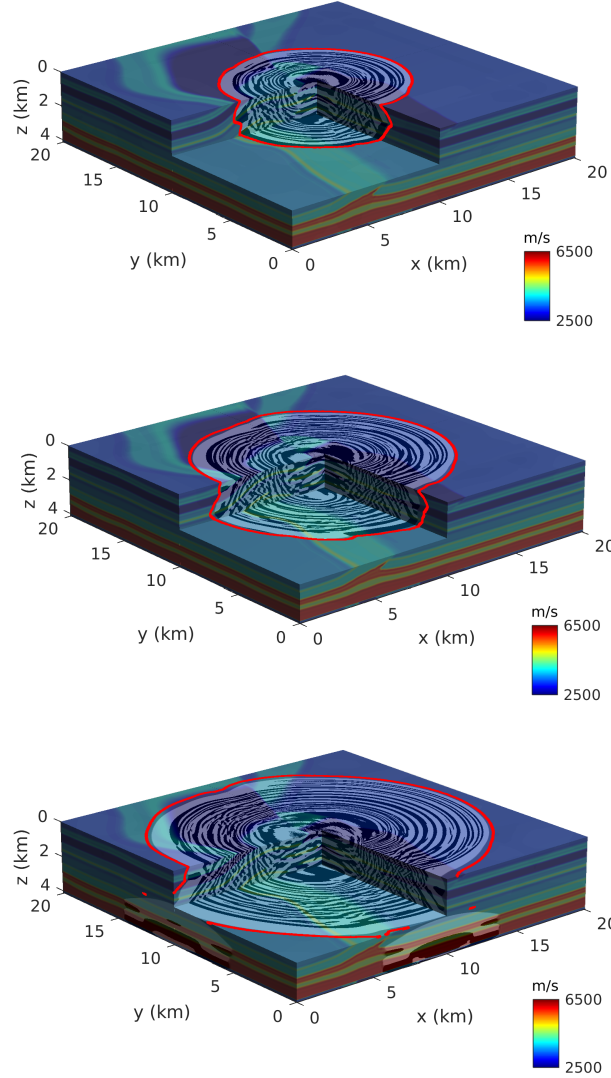


Figure 12: Elastic wavefield (black and white) computed in the 3D TTI medium coming from the homogenization of the SEG/EAGE overthrust model. The background corresponds to the P-wave velocity of this model. The red contour corresponds to the isochrone computing through our fast marching eikonal solver (with GPU implementation). The different snapshots are obtained at  $t = 1.5$  s (top),  $t = 2$  s (middle) and  $t = 2.5$  s (bottom).

The error due to the TTI projection can be estimated by the  $L^2$ -error between the general scheme on general medium and the general scheme on TTI medium as shown in Table 2, and is around 0.062%. For comparison, it is more than thirty times less than the projection error in the orthorhombic setting presented Section 5.2. The proposed TTI scheme, and the general scheme [DCC<sup>+</sup>21], rely on completely different discretization principles, and for this reason they produce slightly different numerical solutions even when applied to the *same* TTI eikonal PDE. We observe on Table 2 that the  $L^2$  error between the TTI scheme and the general scheme on a general medium is around 0.68%, which is ten times larger than the error associated with the



	Computation time (s)	$L^2$ -error
General scheme on general medium	1571	-
General scheme on TTI medium	1569	0.00062
TTI scheme with CPU on TTI medium	478	0.0068
TTI scheme with GPU on TTI medium	13	0.0056
Isotr. scheme with GPU on isotr. med.	0.28	0.011

Table 2: Computation time on a realistic synthetic test-case. The  $L^2$ -error corresponds to the difference between the numerical solution and the solution computed by the general scheme on general medium, normalized by the  $L^2$  norm of the solution computed by the general scheme on general medium. The general scheme is the scheme from [DCC<sup>+</sup>21], used with second-order accuracy. The TTI scheme with GPU is used with 10 ellipses. The isotropic scheme is a Fast Marching scheme with second-order accuracy using GPU.

TTI projection alone, observed with the general scheme on the TTI medium. This validates the assumption that the anisotropy in the homogenized model is very close to TTI anisotropy, and that the associated projection error is well below the discretization error, related to the grid scale and scheme design. Besides, we observe that the computation time is greatly improved by the TTI scheme with GPU implementation compared with the general scheme, with a computation time a hundred times faster. On the other hand, the isotropic projection of the homogenized model leads to an error of around 1.1%, which is seventeen times higher than the projection error due to the TTI scheme, with a computation again fifty times faster.

As a conclusion, the anisotropy in this dataset is close to TTI anisotropy, as is expected from the homogenization process in an isotropic medium with sedimentary structure: the normal axis to the layers is a natural axis of symmetry. In the case of seismics faults or complex interactions between the layers, the anisotropy can become more complex and lose this symmetry axis, but in this realistic instance, the TTI anisotropy seems to be enough to explain the general anisotropy coming from the homogenization process. Contrast this with the Olivine medium considered Section 5.2, which is an orthorhombic crystal system with no such rotation invariance, and whose TTI projection error is significant. Therefore, the TTI scheme is adapted to efficiently compute the first arrival traveltimes in a realistic medium with sedimentary structure and no intrinsic anisotropy coming from inner crystal structure.

## 6 Conclusion

We presented a discretization for the eikonal equation with anisotropy coming from a TTI Hooke tensor. The scheme is monotone and causal, hence solvable in a single pass using the fast-marching method, but also has a simple Eulerian structure, hence fits massively parallel architectures as well; using classical enhancements such as source factorization, we achieve second order accuracy. Two implementations have been proposed, one for CPU and one for GPU. The GPU implementation features an additional parameter which must be correctly tuned, namely the number of ellipses whose envelope approximates the  $P$ -slowness surface, but performs much faster compared with the CPU implementation. The scheme is more than fifty times faster with a small loss in accuracy compared with the scheme from [DCC<sup>+</sup>21], which is a state-of-the-art scheme able to handle media with Hooke tensors of general anisotropy. In addition, the scheme from [DCC<sup>+</sup>21] suffers from a limit on the strength of anisotropy it can tackle (defined as the ratio between the highest velocity and the lowest velocity at a given position) even for TTI

media, whereas our present scheme does not exhibit such a restriction on TTI media.

Future research will be devoted to applications to seismic imaging by tomographic inversion. Besides, an extension of the method to orthorhombic Hooke tensors seems possible, since those are TTI in every cross section. The generalization of our present scheme would involve a two dimensional - rather than one dimensional in the present TTI setting - minimization, maximization, or min-max saddle point optimization problem at each grid point.

## A Thomsen parameters and Hooke tensor symmetry

In this section, we briefly describe how the TTI eikonal PDE (1) is related to classical descriptions of an elastic medium, based either on the Hooke elasticity tensor, or on the Thomsen parameters. A 3D geological medium is described by a fourth-order elasticity tensor, referred to as the Hooke tensor and denoted  $\mathbf{c} = (c_{ijkl})$ , where  $i, j, k, l \in \{1, 2, 3\}$ , and by the density  $\rho$  of the medium. The Hooke tensor is subject to the symmetry relations  $c_{ijkl} = c_{jikl} = c_{jkij}$ , allowing it to be represented as a  $6 \times 6$  matrix  $\mathfrak{C}$  using Voigt's notation, see Section 2.1 and (25).

Some additional symmetries are often considered for a geological medium. A transversely isotropic medium is a geological medium whose local elasticity properties are invariant by rotation around a specific axis. It is called *vertically transversely isotropic* (VTI) in case of invariance around the vertical axis, and *tilted transversely isotropic* (TTI) otherwise. In the case of VTI symmetry, the Hooke tensor (in Voigt notation) only has 5 independent elastic parameters and can be written as [Tho86]:

$$\mathfrak{C}^{VTI} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & 0 & 0 & 0 \\ c_{12} & c_{11} & c_{13} & 0 & 0 & 0 \\ c_{13} & c_{13} & c_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{c_{11}-c_{12}}{2} \end{pmatrix}.$$

A Hooke tensor with TTI symmetry can be obtained from a Hooke tensor with VTI symmetry  $\mathfrak{C}^{VTI}$  and a rotation matrix  $R$ , through the usual change of variables formula

$$\mathfrak{c}_{i'j'k'l'}^{TTI} = \sum_{i,j,k,l \in \{1,2,3\}} \mathfrak{c}_{ijkl}^{VTI} R_{ii'} R_{jj'} R_{kk'} R_{ll'}. \quad (71)$$

Conversely, a (non-convex) projection procedure allows to reconstruct a VTI tensor and a rotation  $R$  from a given Hooke tensor [CMA<sup>+</sup>20], up to some accuracy loss if the latter only has approximate TTI symmetry.

We now present the eikonal equation related to a TTI Hooke tensor, starting from the Christoffel equation (obtained as a high-frequency approximation of the elastic wave equation) [Sla03]:

$$\det \left[ \sum_{j,l \in \{1,2,3\}} c_{ijkl} \frac{\partial u}{\partial x_j} \frac{\partial u}{\partial x_l} - \rho \delta_{ik} \right] = 0. \quad (72)$$

Between brackets is a  $3 \times 3$  matrix, of indices  $i, k$ , obtained as the difference of (i) a contraction of the Hooke tensor by the partial derivatives of the arrival time function  $u$ , and (ii) the identity matrix scaled by the density  $\rho$  of the medium. For simplicity we assume in the following that  $\rho = 1$ , up to considering the reduced tensor  $\mathbf{c}/\rho$ .

Define the *slowness vector*  $(p_x, p_y, p_z) := R\nabla u$ , and let  $p_r^2 := p_x^2 + p_y^2$ . Then the Christoffel equation (72) for a TTI symmetry can be algebraically factored as follows:

$$0 = \left( \frac{c_{11} - c_{12}}{2} p_r^2 + c_{44} p_z^2 - 1 \right) \times \\ (c_{11} c_{44} p_r^4 + c_{33} c_{44} p_z^4 - (2c_{13} c_{44} + c_{13}^2 - c_{11} c_{33}) p_r^2 p_z^2 - (c_{11} + c_{44}) p_r^2 - (c_{33} + c_{44}) p_z^2 + 1).$$

The first factor of this equation characterizes the arrival time of the SH (horizontal shear wave) propagation. This factor defines a Riemannian eikonal equation, which can be solved numerically [Mir14a, Mir19], but is of not interest for the computation of the first travel time. The second factor corresponds to the coupling P-SV, between the qP (quasi-pure pressure wave) and the qSV (quasi-pure vertical shear wave), and is the factor we need to consider for the first-arrival time. The P-SV equation for a TTI symmetry is a non-Riemannian anisotropic eikonal equation of degree four, mathematically more complex than the SH equation, which is reproduced in (6) and studied in this paper. Interestingly, the parameter  $c_{12}$  only appears in the SH equation, and the four relevant parameters  $(c_{11}, c_{13}, c_{33}, c_{44})$  for the P-SV equation can be organized in a 2-dimensional Hooke tensor (28).

The P-SV equation, henceforth referred to as the TTI eikonal equation, is summarized as

$$ap_r^4 + bp_z^4 + cp_r^2 p_z^2 + dp_r^2 + ep_z^2 = 1, \quad \text{where } (p_x, p_y, p_z) = R\nabla u \text{ and } p_r^2 := p_x^2 + p_y^2, \quad (73)$$

with coefficients  $(a, b, c, d, e)$  derived from the Hooke tensor as above. For the Hooke tensors considered in geophysics, the TTI equation has two distinct solutions. In the  $(p_x, p_y, p_z)$  coordinate system, these solutions are called *slowness surfaces*, and are invariant by the rotation  $R$ . The inner surface corresponds to the slowness of the P wave (that is, the inverse of its velocity), whereas the outer surface corresponds to the slowness of the S wave. They are illustrated on Figs. 1 and 2, and studied in detail in Section 2.

**Remark A.1.** *Thomsen's elastic parameters  $(V_p, V_s, \epsilon, \delta)$  define another approach to obtain the TTI eikonal equation (73), with the conversion formula [Tho86]:*

$$V_p = \sqrt{\frac{c_{33}}{\rho}}, \quad V_s = \sqrt{\frac{c_{44}}{\rho}}, \quad \epsilon = \frac{c_{11} - c_{33}}{2c_{33}}, \quad \delta = \frac{(c_{13} + c_{44})^2 - (c_{33} - c_{44})^2}{2c_{33}(c_{33} - c_{44})}.$$

*The Thomsen parameters have physical interpretations in a weakly anisotropic setting: in particular,  $V_p$  approximates the speed of the P-wave, and  $V_s$  of the S-wave. Nevertheless this is only an approximation in a special asymptotic setting, and in general both the P and S slowness surfaces depend on the four Thomsen parameters. For this reason we do not use here the convention  $V_s = 0$ , which has sometimes been considered to simplify the PDE (73) when one is only interested in the first travel time computation, corresponding to the P-wave.*

**Remark A.2.** *In this paper, and in our numerical method, we only require the matrix  $R$  to be invertible in the definition (73) of the eikonal equation. This may be surprising since the TTI formalism (71) and (72), based on physical considerations, makes the stronger assumption that  $R$  is a rotation. Our motivation for allowing non-rotations is that the computational domain is often the image of the physical domain by a diffeomorphism, e.g. to take into account the topography of the surface. In that case the equivalent eikonal PDE in the computational domain involves a matrix  $R$  defined as the product of the original rotation  $R_0$ , associated with the TTI model, and of the Jacobian of the diffeomorphism, which is usually not a rotation.*

Finally, we want to create a criterion based on the coefficients of a TTI metric, to quantify its anellipticity. We suggest the criterion:  $c_{anel} := \alpha^* - \alpha_*$ , where  $0 < \alpha_* \leq \alpha^* < 1$  are defined in Theorem 1.3. It characterizes the difference between the two most extreme ellipses when doing the envelope of the TTI metric, see Figure 13. In the case of an elliptic metric, we have:  $c_{anel} = 0$ .

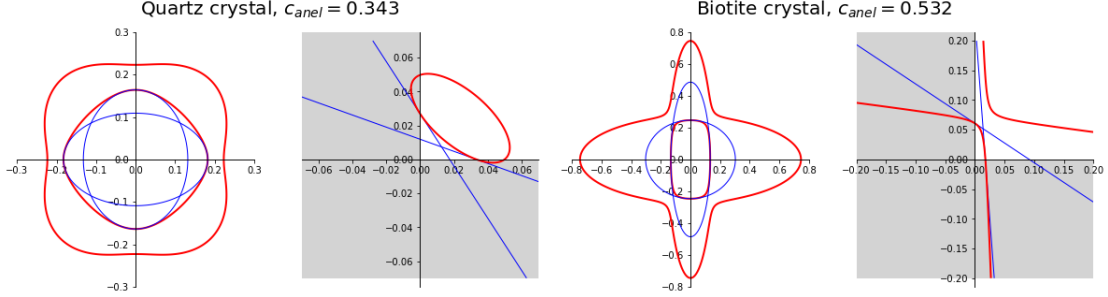


Figure 13: Example of two TTI media from the article [Tho86], represented as in 2, and with corresponding  $c_{anel}$ . In blue are shown the two most extreme ellipses in the optimization problem, which we use to define  $c_{anel}$ .

In order to evaluate the update operator  $\Lambda$  with the optimization by fine sampling, we consider a sampling of an interval over  $K + 1$  elements, with  $K$  chosen by the user: the parameter  $K$  could reasonably be chosen depending on the criterion  $c_{anel}$ , as the accuracy of our approach should depend on how far the TTI metric is from an elliptic metric.

For illustrative purposes, we consider the article [Tho86], in which there are 58 examples of TTI metrics, corresponding to real and hypothetical materials, and we show the histogram of the corresponding  $c_{anel}$  in Figure 14. The two media with the highest  $c_{anel}$  correspond to crystallographic media, which are not usual in geophysics.

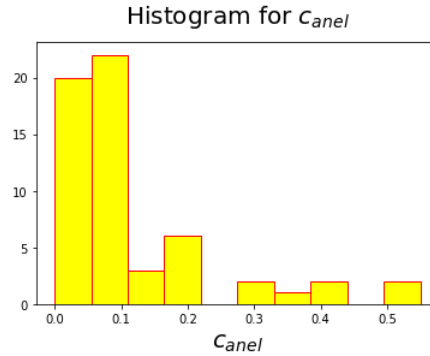


Figure 14: Histogram of  $c_{anel}$  for TTI examples from the article [Tho86].

From this study, we can reasonably assume that the usual TTI metrics do not exhibit a very strong anellipticity, apart from some crystallographic media. Therefore, for most real test-cases, the optimization by fine sampling should give a good result even if we use a small amount of ellipses in the sampling.

## B Selling's decomposition

This section is devoted to Selling's decomposition of positive definite matrices, and to the proof of Proposition 1.7. Selling's decomposition is a tool originating from the field of lattice geometry [Sel74, CS92], which has recently found a number of applications in the design of anisotropic PDE discretizations, see [FM14, Mir19, BBM20, BM21] and Section 1.2. In the following, we present its definition, basic properties, and practical construction. For that purpose we introduce the concept of superbase, which is a slightly redundant coordinates system in  $\mathbb{Z}^d$ .

**Definition B.1.** A superbase of  $\mathbb{Z}^d$  is a family  $(v_0, \dots, v_d) \in (\mathbb{Z}^d)^{d+1}$  such that  $v_0 + \dots + v_d = 0$  and  $|\det(v_1, \dots, v_d)| = 1$ .

Given a superbase  $(v_0, \dots, v_d)$  of  $\mathbb{Z}^d$  and a matrix  $D \in S_d^{++}$ , we have Selling's formula

$$D = - \sum_{0 \leq i < j \leq d} \langle v_i, Dv_j \rangle e_{ij} e_{ij}^\top, \quad (74)$$

where the offsets are defined by the linear relations  $\langle e_{ij}, v_k \rangle = \delta_{ij} - \delta_{ik}$  for all  $0 \leq i < j \leq d$  and  $0 \leq k \leq d$ . This decomposition involves  $I = \binom{d+1}{2} = d(d+1)/2$  terms, and the offsets have integer coordinates by construction. They admit simple expressions in small dimension: if  $d = 2$  and  $\{i, j, k\} = \{0, 1, 2\}$  then  $v_{ij} = \pm e_k^\perp$ , and if  $d = 3$  and  $\{i, j, k, l\} = \{0, 1, 2, 3\}$  then  $v_{ij} = \pm v_k \times v_l$  (where  $\times$  denotes the cross product). For a proof of Selling's formula (74), see [BBM21a, Lemma B.2].

Selling's decomposition of  $D \in S_d^{++}$ ,  $d \in \{2, 3\}$ , is defined as Selling's formula (74) applied to a  $D$ -obtuse superbase, defined below, in such way that the weights  $-\langle v_i, Dv_j \rangle$  are non-negative.

**Definition B.2.** A superbase  $b = (v_0, \dots, v_d)$  is said  $D$ -obtuse, where  $D \in S_d^{++}$ , if  $\langle v_i, Dv_j \rangle \leq 0$  for all  $0 \leq i < j \leq d$ . We let  $S_b := \{D \in S_d^{++}; b \text{ is } D\text{-obtuse}\}$ .

Using Pauli matrices in dimension  $d = 2$ , one obtains a linear parametrization

$$D(x, y) = \begin{pmatrix} 1+x & y \\ y & 1-x \end{pmatrix}, \quad x^2 + y^2 < 1, \quad (75)$$

of the set of symmetric positive definite matrices of trace two, by the Euclidean unit ball. Figure 3 (left) illustrates the anisotropy defined by  $D(a, b)$ . The domains  $S_b$  associated to superbases  $b$  of  $\mathbb{Z}^2$  appear as triangles in this parametrization, and together they define an infinite triangulation of the open unit ball  $\{x^2 + y^2 < 1\}$ , see Fig. 3 (center).

In order to conclude the proof of the first part of Proposition 1.7, one needs to show that a  $D$ -obtuse superbase exists, which is the purpose of Selling's algorithm.

**Proposition B.3** (Selling's algorithm, [Sel74] or [BBM21a, Proposition B.3]). Let  $b = (v_0, \dots, v_d)$  be a superbase of  $\mathbb{Z}^d$ , where  $d \in \{2, 3\}$ , and let  $D \in S_d^{++}$ . If  $b$  is not  $D$ -obtuse, permute it so that  $\langle v_0, Dv_1 \rangle > 0$  and update it as follows

$$b \leftarrow (-v_0, v_1, v_0 - v_1) \text{ if } d = 2, \quad b \leftarrow (-v_0, v_1, v_2 + v_0, v_3 + v_0) \text{ if } d = 3. \quad (76)$$

Repeating this operation yields a  $D$ -obtuse superbase in finitely many steps.

In order to establish the second part of Proposition 1.7, we normalize Selling's decomposition as in [BBM21a, Appendix A], up to replacing some offsets with their opposites:

$$D = \sum_{e \in \mathcal{Z}^d} \rho(e; D) e e^\top, \quad \text{where } \mathcal{Z}^d := \{e \in \mathbb{Z}^d; e \succ_{\text{lex}} 0\}, \quad (77)$$

where  $\succ_{\text{lex}}$  stands for the lexicographic ordering. (Note that exactly one of  $e \succ_{\text{lex}} 0$  or  $-e \succ_{\text{lex}} 0$  holds for each  $e \in \mathbb{Z}^d \setminus \{0\}$ , and that  $ee^\top = (-e)(-e)^\top$ .) The weights  $[\rho(e; D)]_{e \in \mathcal{Z}^d}$  are known as Selling parameters [CS92], and depend on  $D$  but *not* on the choice of  $D$ -obtuse superbase, see e.g. [BBM21b, Remark 2.13] for a proof. In view of Selling's formula (74), there exists at most  $d(d+1)/2$  offsets  $e \in \mathcal{Z}^d$  such that  $\rho(e; D) \neq 0$ , for any given  $D \in S_d^{++}$ . With these notations, we summarize in the next result some properties of Selling's decomposition: the offsets are bounded, the weights are locally Lipschitz, and a basis of  $\mathbb{Z}^d$  can be extracted. We denote by  $\mu(D) := \sqrt{\|D\| \|D^{-1}\|}$  the square root of the condition number of a matrix  $D \in S_d^{++}$ .

**Proposition B.4** (Propositions B.4, B.5 and B.8 in [BBM21a]). *The following holds in dimension  $d \in \{2, 3\}$ , denoting  $C = 2$  if  $d = 2$  (resp.  $C = 2\sqrt{3}$  if  $d = 3$ ), and for some absolute constant  $c > 0$ :*

- (Offset boundedness) For any  $e \in \mathcal{Z}^d$ ,  $D \in S_d^{++}$  s.t.  $\rho(e; D) \neq 0$ , one has  $\|e\| \leq 2C\mu(D)$ .
- (Lipschitz weights) For any  $e \in \mathcal{Z}^d$ , the mapping  $D \in S_d^{++} \mapsto \rho(e; D)$  is locally Lipschitz with constant  $C^2\mu(D)^2$ .
- (Spanning property) For any  $D \in S_d^{++}$ , there exists  $e_1, \dots, e_d \in \mathcal{Z}^d$  such that

$$\det(e_1, \dots, e_d) = 1, \quad \min_{1 \leq i \leq d} \rho(e_i; D) \geq c\|D^{-1}\|^{-1}.$$

For each superbase  $b = (v_0, \dots, v_d)$  of  $\mathbb{Z}^d$ , and each offset  $e \in \mathcal{Z}$ , the mapping  $D \in S_b \mapsto \rho(e; D)$  is *linear*, where  $S_b$  is defined in Definition B.2. Indeed, in view of Selling's formula (74), either  $e = \pm e_{ij}$  for some  $0 \leq i < j \leq d$  and thus  $\rho(e; D) = -\langle v_i, Dv_j \rangle$  is a linear function of  $D$ , or  $\rho(e; D) = 0$  identically on  $S_b$ . This linearity property, already used in the design of the PDE schemes [BBM20, BM21], allows here to conclude the proof of Proposition 1.7.

*Proof of the second part of Proposition 1.7.* For concreteness and w.l.o.g. we can assume that  $\alpha_* = 0$ ,  $\alpha^* = 1$ , and thus  $D(\alpha) := (1 - \alpha)D_0 + \alpha D_1$ ,  $\alpha \in [0, 1]$ , for some given  $D_0, D_1 \in S_d^{++}$ . Note that  $\mu(D(\alpha)) \leq \max\{\mu(D_0), \mu(D_1)\}$  for all  $\alpha \in [0, 1]$ .

Denote by  $B$  be the collection of all superbases of  $\mathbb{Z}^d$  whose elements are bounded by  $2C \max\{\mu(D_0), \mu(D_1)\}$ , where  $C$  is from Proposition B.4 (offset boundedness). Then any  $D(\alpha)$ -obtuse superbase belongs to  $B$ , for any  $\alpha \in [0, 1]$ .

Given a superbase  $b = (v_0, \dots, v_d) \in B$  the set  $I_b = \{\alpha \in [0, 1]; b \text{ is } D(\alpha)\text{-obtuse}\}$  is defined by linear inequalities:  $\langle v_i, [(1 - \alpha)D_0 + \alpha D_1]v_j \rangle \leq 0$  for all  $0 \leq i < j \leq d$ . Therefore  $I_b$  is closed and convex, hence either  $I_b = \emptyset$ , or  $I_b = [\alpha_b^-, \alpha_b^+]$  is a segment with  $0 \leq \alpha_b^- \leq \alpha_b^+ \leq 1$ . The weights of Selling's decomposition (74) are affine functions of the parameter  $\alpha \in I_b$ , with the general form  $\alpha \mapsto -\langle v_i, [(1 - \alpha)D_0 + \alpha D_1]v_j \rangle$ , whereas the offsets  $e_{ij}$  are constant over  $I_b$ , as announced. Noting that  $[0, 1] = \cup_{b \in B} I_b$  is a finite union of such segments, we establish that Selling's decomposition is piecewise affine (17) which concludes the proof of Proposition 1.7.  $\square$

For concreteness and implementation purposes, we present in Algorithm 2 (without proof) a variant of Selling's algorithm (Proposition B.3), which can be regarded as a constructive implementation of the above proof of the piecewise affine nature of Selling's decomposition. For notational simplicity and w.l.o.g. we assume again that  $D(\alpha) := (1 - \alpha)D_0 + \alpha D_1$  is parametrized over the interval  $[\alpha_*, \alpha^*] = [0, 1]$ . This algorithm produces some breakpoints  $0 = \alpha_0 < \dots \leq \alpha_K = 1$ , and corresponding superbases  $b_0, \dots, b_{K-1}$  of  $\mathbb{Z}^d$ , such that  $b_k$  is  $D(\alpha)$ -obtuse for all  $\alpha \in [\alpha_k, \alpha_{k+1}]$ ,  $0 \leq k < K$ . Thus Selling's decomposition (74) of  $D(\alpha)$  is affine on each of these intervals, as required.

---

**Algorithm 2** A modification of Selling’s algorithm, producing  $0 = \alpha_0 \leq \dots \leq \alpha_K = 1$  and  $b^0, \dots, b^{K-1}$  superbases, such that  $b_k$  is  $[(1 - \alpha)D_0 + \alpha D_1]$ -obtuse  $\forall \alpha \in [\alpha_k, \alpha_{k+1}]$ ,  $0 \leq k < K$ .

---

**Input :**  $D_0, D_1 \in S_d^{++}$ .

**Initialization :** set  $\alpha_0 = 0$ ,  $k = 0$ , and compute a  $D_0$ -obtuse superbase  $b^0$  using Proposition B.3.

**Repeat :**

Denote  $b^k = (v_0, \dots, v_d)$ , and let  $\alpha_{k+1} \in [\alpha_k, \infty]$  be the smallest  $\alpha$  such that  
 $\exists 0 \leq i < j \leq d$ ,  $\langle v_i, (D_1 - D_0)v_j \rangle > 0$  and  $\langle v_i, [(1 - \alpha)D_0 + \alpha D_1]v_j \rangle = 0$ .  
 (Up to permuting  $b^k$ , we assume that  $i = 0$  and  $j = 1$  are the active indices.)

**If**  $\alpha_{k+1} \geq 1$  **then** set  $K := k + 1$  and **exit**.

Define  $b^{k+1}$  by applying Selling’s update (76) to  $b^k$ .

Set  $k \leftarrow k + 1$  and proceed to the next iteration.

---

## C Scheme enhancements for higher accuracy

We describe in this subsection some algorithmic enhancements to the finite differences discretization (2) of the TTI eikonal PDE, meant to improve its accuracy, and we discuss of their relevance and applicability to the CPU and GPU implementations. The improvements are validated by a consistency analysis and by numerical experiments in Section 5.1, but not by a formal convergence analysis. The proposed scheme variants are adapted from the literature, hence are not original in themselves: various approaches to source factorization are presented in [LQ12], second order accurate fast marching is introduced in [Set99], and the use of several discretization scales and coordinate systems depending on the distance from the source is documented in [WFNBZ20].

The discussion applies to any finite differences scheme of the following form, including the discretizations of the Riemannian (18) and TTI (20) eikonal PDEs:

$$\mathfrak{F}u(q) = \hat{\mathfrak{F}}(q, [\delta_h^e u(q)]_{e \in E}), \quad \delta_h^e u(q) := \frac{u(q) - u(q - he)}{h}, \quad (78)$$

where the scheme unknown  $u$  is defined over a subset of the Cartesian grid  $h\mathbb{Z}^d \ni q$  of scale  $h > 0$ , and where  $E \subset \mathbb{Z}^d$  denotes the scheme stencil. We assume that the scheme  $\mathfrak{F}$  obeys (i) the DDE (discrete degenerate ellipticity) property and (ii) the causality property, see Definition 1.6. In other words  $\mathfrak{F}u(q)$  is (i) a non-decreasing function of the finite difference  $\delta_h^e u(q)$ , and (ii) only depends on the positive part  $\max\{0, \delta_h^e u(q)\}$ , for any  $e \in E$ . Since some considered scheme modifications may unfortunately break these properties, we discuss beforehand the extent to which our numerical eikonal solvers need them.

- The fast marching method, our CPU eikonal solver, requires a causal and DDE scheme. However a higher order non-DDE scheme can be used in the optional post-processing step of Algorithm 1, line 3., so as to improve the accuracy of the accepted value  $u(q)$  before it is frozen. Following the principles of the high accuracy fast marching method (HAFMM) [Set99], this modification is only applied if it is small enough. Such post-processing is guaranteed not to degrade the convergence order of the method, by the comparison principle see [DCC<sup>+</sup>21, Proposition D.5], and in practice appears to improve it.
- The iterative GPU eikonal solver, only requires a DDE scheme. On the positive side, the causality property is not needed. On the negative side there is no clear opportunity to introduce a higher order non-DDE scheme, without the risk to create instabilities and to

compromise the convergence of the solver. A basic fix to restore monotony along the solver iterations, without guarantee but with often good empirical results see Section 5.1, is to accept an update value only if it is smaller than the previous one.

The consistency of a finite differences scheme (78) with the corresponding eikonal PDE, such as (2) with the TTI equation (5), is ultimately based on the finite differences approximation:

$$|\delta_h^e u(q) - \langle \nabla u(q), e \rangle| \leq \frac{1}{2} \|\nabla^2 u(r)\| h, \quad (79)$$

for sufficiently smooth  $u$  and where  $r \in [q, q + he]$ . This follows from a Taylor expansion, at a point  $q \in \Omega$ , in a direction  $e \in \mathbb{R}^d$  of differentiation with  $\|e\| = \mathcal{O}(1)$ , and with finite difference scale  $h > 0$ . Each of the scheme variants discussed below is based on introducing in (78, left) a modified finite difference operator  $\tilde{\delta}_h^e$  whose consistency with  $\langle \nabla u(q), e \rangle$  is improved. These scheme variants are easily combined, and all three together are required to achieve second order convergence rates in Section 5.1.

**Source factorization.** The solution to the eikonal equation (5) has a singularity at the source point  $q_0$ , which degrades the accuracy of the finite difference approximation (79) since  $\|\nabla^2 u(q)\| = \mathcal{O}(1/\|q - q_0\|)$  explodes as  $q \rightarrow q_0$ . Source factorization techniques [LQ12] rely on the computation of an equivalent  $u_*$  of the solution near the singularity, which can be easily evaluated and differentiated to machine precision. Typically one uses  $u_*(q) = \mathcal{F}_{q_0}(q - q_0)$ , which is the exact solution in the case of a constant metric, see Remark 1.5. As a result  $\|\nabla^2 u(q) - \nabla^2 u_*(q)\| = \mathcal{O}(1)$  as  $q \rightarrow q_0$ , for a metric of suitable regularity, leading to a corresponding improvement in the scheme consistency (79). Following the principle of additive source factorization [LQ12], we introduce the modified finite differences operator

$$\tilde{\delta}_h^e u(q) := \frac{u(q) - u(q - he)}{h} + \omega_h^e, \quad \text{where } \omega_h^e := \left( \langle \nabla u_*(q), e \rangle - \frac{u_*(q) - u_*(q - he)}{h} \right). \quad (80)$$

The additional corrective term  $\omega_h^e = \mathcal{O}(h^2/\|q - q_0\|)$  preserves the DDE property, since  $\tilde{\delta}_h^e u(q) = \delta_h^e u(q) + \omega_h^e$  is a non-decreasing function of  $\delta_h^e u(q)$ , but breaks the causality property, since  $\max\{0, \tilde{\delta}_h^e u(q)\} = \max\{0, \delta_h^e u(q) + \omega_h^e\}$  is not a function of  $\max\{0, \delta_h^e u(q)\}$  when  $\omega_h^e > 0$ . As a result, this modification fits well in the iterative GPU solver, but introduces slight errors in the FMM CPU solver. (Note that there exists a stronger quantitative variant of the causality property, referred to as  $\delta_1$ -causality where  $\delta_1 > 0$ , which is preserved under source factorization, see [DCC<sup>+</sup>21, Proposition D.4]. However the scheme proposed in this paper is not  $\delta_1$ -causal.)

**Multiscale computation.** This technique features a preliminary run of the eikonal solver in a neighborhood  $\Omega^1 \subset \Omega$  of the source point  $q_0$  [WFNBZ20], using a smaller grid size  $h_1 = h/k_1$  where  $k_1 \geq 2$  is an integer. In essence, the pre-computation uses the modified finite difference operator

$$\tilde{\delta}_h^e := \frac{u(q) - u(q - h_1 e)}{h_1} = \frac{u(q) - u(q - (h/k_1)e)}{h/k_1},$$

which is more accurate by virtue of the smaller grid scale, reducing the consistency error (79) by an approximate factor  $k_1$ . The dimensions of  $\Omega^1$  are chosen (at least)  $k_1$  times smaller than  $\Omega$ , in such way that the refined computational domain around the source  $\Omega_h^1 = \Omega^1 \cap h_1 \mathbb{Z}^d$  has comparably many points as the global computational domain  $\Omega_h = \Omega \cap h \mathbb{Z}^d$ . In principle, this approach can be implemented recursively using an increasing sequence of subdomains  $q_0 \in \Omega^N \subset \dots \subset \Omega^1 \subset \Omega$  and of grid scales  $h_N | \dots | h_1 | h$ , but a basic two scales approach with  $h_1 = h/4$  was found to be appropriate in our experiments Section 5.1.



**Second order finite differences.** The modified finite difference operator

$$\tilde{\delta}_h^e u(x) = \frac{u(q) - u(q - he)}{h} + \frac{u(q) - 2u(q - he) + u(q - 2he)}{2h},$$

is upwind, second order accurate, and is often used in the post-processing step of fast marching methods [Set99, Mir19], see Algorithm 1, line 3. The consistency error  $\mathcal{O}(\|\nabla^3 u(q)\| h^2)$  is more favorable than (79) in regions where the solution is smooth. This variant is therefore mostly useful far from the source point, in contrast with the previous two modifications. The DDE property fails however, because  $\tilde{\delta}_h^e u(q) = 2\delta_h^e u(q) - \frac{1}{2}\delta_h^{2e} u(q)$  is *not* a non-decreasing function of  $\delta_h^e u(q)$  and  $\delta_h^{2e} u(q)$ , and for safety the second order correction is thus rejected if it is too large.

## References

- [ABK97] F. Aminzadeh, J. Brac, and T. Kunz. *3-D Salt and Overthrust models*. SEG/EAGE 3-D Modeling Series No.1, 1997.
- [BBM20] Joseph Bonnans, Guillaume Bonnet, and Jean-Marie Mirebeau. Monotone and second order consistent scheme for the two dimensional Pucci equation. 2020.
- [BBM21a] Frédéric Bonnans, Guillaume Bonnet, and Jean-Marie Mirebeau. A linear finite-difference scheme for approximating Randers distances on Cartesian grids. 2021.
- [BBM21b] J Frederic Bonnans, Guillaume Bonnet, and Jean-Marie Mirebeau. Second order monotone finite differences discretization of linear anisotropic differential operators. *Mathematics of computation*, 90(332):2671–2703, 2021.
- [BC91] Vladislav Babuska and Michel Cara. *Seismic anisotropy in the Earth*, volume 10. Springer Science and Business Media, 1991.
- [BCD08] Martino Bardi and Italo Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Springer Science and Business Media, 2008.
- [BL10] Jonathan Borwein and Adrian S Lewis. *Convex analysis and nonlinear optimization: theory and examples*. Springer Science and Business Media, 2010.
- [BM21] Guillaume Bonnet and Jean-Marie Mirebeau. Monotone discretization of the monge-ampère equation of optimal transport. 2021.
- [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.
- [BST83] Iulian Beju, Eugen Soós, and Petre P Teodorescu. *Euclidean tensor calculus with applications*. CRC Press, 1983.
- [Car01] Michael Carter. *Foundations of mathematical economics*. MIT Press, 2001.
- [CBM20] J. Cao, R. Brossier, and L. Métivier. 3d acoustic-(visco) elastic coupled formulation and its spectral-element implementation on a cartesian-based hexahedral mesh. In *SEG Technical Program Expanded Abstracts 2020*, pages 2643–2647. Society of Exploration Geophysicists, 2020.

- [CHK13] Marcel Campen, Martin Heistermann, and Leif Kobbelt. Practical Anisotropic Geodesy. *Computer Graphics Forum*, 32(5):63–71, August 2013.
- [CMA<sup>+</sup>20] Paul Cupillard, Wim Mulder, Pierre Anquez, Antoine Mazuyer, and J Barthélémy. The Apparent Anisotropy of the SEG-EAGE Overthrust Model. In *82nd EAGE Annual Conference and Exhibition*, pages 1–5. European Association of Geoscientists and Engineers, 2020.
- [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.
- [CWW13] Keenan Crane, Clarisse Weischedel, and Max Wardetzky. Geodesics in heat: A new approach to computing distance based on heat flow. *ACM Transactions on Graphics (TOG)*, 32(5):152, 2013.
- [DCC<sup>+</sup>21] François Desquilbet, Jian Cao, Paul Cupillard, Ludovic Métivier, and Jean-Marie Mirebeau. Single pass computation of first seismic wave travel time in three dimensional heterogeneous media with general anisotropy. 2021.
- [Eva10] Lawrence C Evans. *Partial Differential Equations*. American Mathematical Soc., 2010.
- [FKW13] Zhisong Fu, Robert M Kirby, and Ross T Whitaker. A fast iterative method for solving the eikonal equation on tetrahedral domains. *SIAM Journal on Scientific Computing*, 35(5):C473–C494, 2013.
- [FM14] Jérôme Fehrenbach and Jean-Marie Mirebeau. Sparse non-negative stencils for anisotropic diffusion. *Journal of Mathematical Imaging and Vision*, 49(1):123–147, 2014.
- [HF07] M Sabry Hassouna and Aly A Farag. Multistencils fast marching methods: A highly accurate solution to the eikonal equation on cartesian domains. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(9):1563–1574, 2007.
- [HJ16] Sumin Hong and Won-Ki Jeong. A multi-gpu fast iterative method for eikonal equations using on-the-fly adaptive domain decomposition. *Procedia Computer Science*, 80:190–200, 2016.
- [JW08] Won-Ki Jeong and Ross T Whitaker. A fast iterative method for eikonal equations. *SIAM Journal on Scientific Computing*, 30(5):2512–2534, July 2008.
- [Kom88] Hidetoshi Komiya. Elementary proof for Sion’s minimax theorem. *Kodai mathematical journal*, 11(1):5–7, 1988.
- [Kry05] Nicolai V Krylov. The rate of convergence of finite-difference approximations for Bellman equations with Lipschitz coefficients. *Applied Mathematics and Optimization*, 52(3):365–399, 2005.
- [KS98] R Kimmel and James A. Sethian. Computing geodesic paths on manifolds. *Proceedings of the National Academy of Sciences*, 95(15):8431–8435, July 1998.

- [LBBMV17] P Le Bouteiller, M Benjemaa, L Métivier, and J Virieux. An accurate discontinuous Galerkin method for solving point-source Eikonal equation in 2-D heterogeneous anisotropic media. *Geophysical Journal International*, 212(3):1498–1522, 2017.
- [LQ12] Songting Luo and Jianliang Qian. Fast sweeping methods for factored anisotropic eikonal equations: multiplicative and additive factors. *Journal of Scientific Computing*, 52(2):360–382, 2012.
- [MGB<sup>+</sup>21] Jean-Marie Mirebeau, Lionel Gayraud, Rémi Barrère, Da Chen, and François Desquilbet. Massively parallel computation of globally optimal shortest paths with curvature penalization. 2021.
- [Mir14a] Jean-Marie Mirebeau. Anisotropic Fast-Marching on cartesian grids using Lattice Basis Reduction. *SIAM Journal on Numerical Analysis*, 52(4):1573–1599, January 2014.
- [Mir14b] Jean-Marie Mirebeau. Efficient fast marching with Finsler metrics. *Numerische Mathematik*, 126(3):515–557, 2014.
- [Mir18] Jean-Marie Mirebeau. Fast-marching methods for curvature penalized shortest paths. *Journal of Mathematical Imaging and Vision*, 60(6):784–815, 2018.
- [Mir19] Jean-Marie Mirebeau. Riemannian fast-marching on cartesian grids, using voronoi’s first reduction of quadratic forms. *SIAM Journal on Numerical Analysis*, 57(6):2608–2655, 2019.
- [MP19] Jean-Marie Mirebeau and Jorg Portegies. Hamiltonian fast marching: A numerical solver for anisotropic and non-holonomic eikonal pdes. *Image Processing On Line*, 9:47–93, 2019.
- [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.
- [PTVF07] W H Press, S A Teukolsky, W T Vetterling, and B P Flannery. Numerical recipes: the art of scientific computing. 3rd. 2007.
- [Sel74] Eduard Selling. Ueber die binären und ternären quadratischen Formen. *Journal fur die Reine und Angewandte Mathematik*, 77:143–229, 1874.
- [Set96] James A. Sethian. A fast marching level set method for monotonically advancing fronts. *Proceedings of the National Academy of Sciences*, 93(4):1591–1595, 1996.
- [Set99] James A. Sethian. *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*. Cambridge University Press, 1999.
- [Sla03] Michael A Slawinski. *Seismic waves and rays in elastic media*, volume 34. Elsevier, 2003.
- [SV03] James A. Sethian and Alexander Boris Vladimirsky. Ordered upwind methods for static Hamilton-Jacobi equations: theory and algorithms. *SIAM Journal on Numerical Analysis*, 41(1):325–363, 2003.

- [TBM<sup>+</sup>19] P. T. Trinh, R. Brossier, L. Métivier, L. Tavard, and J. Virieux. Efficient 3D time-domain elastic and viscoelastic Full Waveform Inversion using a spectral-element method on flexible Cartesian-based mesh. *Geophysics*, 84(1):R75–R97, 2019.
- [Tho86] Leon Thomsen. Weak elastic anisotropy. *Geophysics*, 51(10):1954–1966, 1986.
- [WFNBZ20] Malcolm CA White, Hongjian Fang, Nori Nakata, and Yehuda Ben-Zion. PyKonal: a Python package for solving the eikonal equation in spherical and Cartesian coordinates using the fast marching method. *Seismological Research Letters*, 91(4):2378–2389, 2020.
- [Zha05] Hongkai Zhao. A fast sweeping method for eikonal equations. *Mathematics of computation*, 74(250):603–627, 2005.